Improving prediction of “inevitable” surgery during non-surgical treatment of sciatica

Wilco C. Peul a,b,*, Ronald Brand c, Ralph T.W.M. Thomeer a, Bart W. Koes d

a Department of Neurosurgery, Leiden University Medical Center/Medical Center Haaglanden, Albimusdreef 2, P.O. Box 9600, 2300 RC Leiden, The Netherlands
b Department of Neurosurgery, Medical Center Haaglanden, The Hague, The Netherlands
c Department of Medical Statistics, Leiden University Medical Center, Leiden, The Netherlands
d Department of General Practice, University Medical Center Rotterdam, Rotterdam, The Netherlands

Received 31 October 2007; received in revised form 19 January 2008; accepted 6 February 2008

Abstract

A randomized trial showed that surgery speeds up recovery of patients suffering sciatica for 6 weeks but prolonged conservative care yields similar results at one year. However 39% of this conservative care group ultimately underwent surgery after a mean period of 18.7 weeks. We evaluated variables to improve prediction of surgery in the conservatively treated cohort. Baseline data on 142 patients enrolled in the conservative treatment arm of a randomized trial were analyzed to select those that could contribute to the prediction of surgery. The actual occurrence of surgery was used as dependent outcome of interest. Variables measured at baseline included neurological examination results, the visual analogue scale for pain (VAS) and the Roland disability questionnaire (RDQ). Higher pain intensity and functional limitations at baseline were associated with an increased likelihood of surgery during follow-up. Mutually adjusted Odds Ratios of 1.7 (95% CI; 1.1–2.7) per 20 mm incremental intensification of pain on the VAS score and 1.8 (95% CI; 1.2–2.9) per 3 points of deterioration of the RDQ score quantify the increasing chance of undergoing delayed surgery. Despite maximal efforts to the contrary, surgery could not be prevented for a considerable proportion of patients in a conservatively treated cohort. Compared to those with lower scores initially, patients with more intense leg pain or higher disability scores were at higher risk to undergo delayed surgery. The individual surgical decision process is facilitated by the use of pain and disability scales complemented by patient preferences.

Keywords: Leg pain; Prolonged conservative care; Roland disability questionnaire; Sciatica; Visual analogue scale

1. Introduction

In general lumbar disc surgery for sciatica is performed to speed up recovery of leg pain and disability [4]. It ranks high among the most frequently performed surgical procedures for neuropathic and musculoskeletal disorders [2,17]. The timing of surgery appears to vary greatly between different countries [2]. A recently published study revealed evidence that for the majority of patients both prolonged conservative care and early surgery resulted in complete recovery at one year but the conservative arm took twice as long to reach recovery [12]. Delayed surgery, however, had been necessary after all for 39% of the patients assigned to the prolonged conservative treatment group, during the first year after randomization. Although the intention of prolonged conservative care reduced the number of surgical interventions substantially, those patients who underwent...
surgery at a later date suffered from pain and disability for quite a prolonged period, up to 6 months. It would be of great value to patients opting for a strategy of prolonged conservative care, if determinants available early in the course of sciatica could be found that would improve the prediction of surgery during follow-up. Therefore we evaluated the predictive value of clinical, demographic, and radiological variables for surgery actually performed in the conservative treatment cohort of a recently conducted randomized trial.

2. Methods

A multicenter prospective randomized trial of patients with short-term severe sciatica was conducted to determine whether early surgery resulted in a more effective outcome during the first year compared to a strategy of prolonged conservative treatment (including, if indicated, delayed surgery). The medical ethics committee at each of the nine participating hospitals approved the protocol. Written informed consent was obtained from all the patients. Details of the design and study protocol were published previously [12,13].

2.1. Patients

Eligible patients were 18–65 years of age, with radiological confirmation of a clinically expected disc herniation causing an incapacitating lumbosacral radicular syndrome lasting between 6 and 12 weeks as documented by the attending neurologist. The neurologist referred patients with persistent sciatica and a clear indication for surgery. The neurosurgeon confirmed the indication for surgery if a Magnetic Resonance Imaging (MRI) of the lumbar spine presented a herniated disc with nerve root compression. At the time of enrolment and randomization an independent research nurse again verified persistence of complaints and the surgical indication. Cauda equina syndrome or severe paresis (MRC < 3) was excluded as were identical complaints in the past twelve months, a history of spinal surgery, bony stenosis, spondylolisthesis, pregnancy or severe comorbidity.

For the purpose of the present analysis, the patients who originally were allocated at random to conservative care were selected as the study cohort.

2.2. Treatment

Prolonged conservative management was provided by the general practitioner. Ample information was provided about the favorable prognosis. If necessary the prescription of pain medication was adjusted according to existing clinical guidelines [16]. Opiates were frequently prescribed, but epidural or periradicular corticosteroids were not routinely injected. If there were considerable fear of movement, the help of a physiotherapist was recommended. Treatment was aimed mainly at resumption of daily activities. However if sciatica were still present at 6 months after randomization, microdiscectomy was considered if a repeat MRI still showed the disc herniation. The most important criteria to perform a repeated MRI were no or only slight clinical non-relevant improvements in leg pain intensity and functional disability. The intensity of sciatic complaints, neurological deficit and repeated MRI were again evaluated subsequently by the neurologist and the neurosurgeon. Increasing drug-resistant leg pain and progressive neurological deficit were reasons for performing surgery even before 6 months. When patients requested surgery, they were again evaluated by their treating physician and the assigned research nurse, who had to confirm that recovery had not occurred and that the repeated MRI showed an unresolved disc herniation. Subsequently the neurosurgeon was consulted by the patient who requested surgery under the premise that further recovery was not to be expected in the next few months. If surgery were performed, the herniated part of the disc would be removed together with as much degenerated nuclear material as possible [1]. The removal of bone to gain access to the disc space was minimized while total disc excision was never performed.

2.3. Variables

Prognostic determinants were selected on the basis of classical physiological hypotheses or inferred from earlier studies [6,7,18]. These socio-demographic, anamnestic, and neurological variables were collected before randomization was performed. Besides these variables the baseline intensity of leg pain, functional disability, the MRI presence of a sequestrated herniated disc and MRI-level of nerve root compression were used as independent variables of interest.

2.4. Outcomes

The occurrence of surgery performed during the course of prolonged conservative care was the event of interest. Functional outcome as indicated by means of the Roland disability questionnaire for sciatica (RDQ) [8,11] and intensity of leg or back pain on a 100 mm visual analogue scale for leg pain (VAS-leg and VAS-back) [3] was assessed at 0, 2, 4, 8, 12, 26, 38 and 52 weeks, respectively. For the present analysis only data from the baseline measurements were used. Furthermore a 7-point Likert scale evaluated the recovery. In a dichotomized form “satisfactory” outcome is similar to “complete” or “near-complete” recovery.

2.5. Statistical analysis

Data collection and quality checks were performed using the ProMISe data-management system of the Department of Medical Statistics & BioInformatics of the LUMC. All statistical analyses were carried out using SPSS, version 14.0 [15]. All variables were recoded before being used in a logistic regression analysis as described previously in the protocol [13]. Since the VAS-leg pain and RDQ are both continuous variables and the Odds Ratios of these variables correspond by definition to the effect of an increase in 1 unit of the underlying scale, these were rescaled (without any effect on their significance) for reporting purposes to let the Odds Ratios reflect the effect of an increment of 20 mm and 3 points, respectively, both corresponding to the estimated Minimal Clinical Important Difference [10,14].

The predictive effect of variables was analyzed by constructing a multivariable logistic regression model with “the occurrence of surgery during the first twelve months” as the
outcome of interest. The model was obtained using a stepwise backward elimination process with threshold values of 0.10 and 0.05 for removal and inclusion of predictors. Since follow-up observation was complete, there was no censoring and a survival modelling approach was not necessary, the outcome of interest simply being the probability of having undergone surgery by month 12 after being randomized to conservative care.

Finally, after having constructed the model that retained all significant predictors, the estimated probabilities were tabulated using some typical values of the predictors and the estimated odds ratio were tabulated with their respective confidence intervals. Whether prediction is actually possible in a reliable way for an individual patient can be seen in a classification table under the assumption of allocating the patient to “surgery” if the probability of surgery is more than 50% (as is common in, for example, diagnostic tests). In view of the restricted size of the study population we focused solely on identifying significant risk factors for surgery. No attempts were made to perform a more refined analysis with an ROC curve or a training/validation subset approach. However, we did a linear discriminant analysis (almost identical to a logistic regression model in this case) to obtain a classification table which is usually shown with a diagnostic test context with cross-validated percentages of correctly classified cases.

3. Results

Of 142 patients assigned to receive prolonged conservative care, 55 (39%) underwent surgery after a mean period of 18.7 (95% CI 14.3–23.0) and median 14.6 (Interquartile range; 6.4–26.0) weeks. Before randomization the mean period of sciatic complaints was 9.5 (SD; 2.11) weeks for all patients treated conservatively initially. The mean Roland disability score for the 55 surgical patients was 15.0 (95% CI; 13.3–16.8) shortly before surgery, while the mean visual analogue scale at that time was 54 (95% CI; 46.2–61.8) mm. Repeated surgery within the first year was performed in a single case. Mean baseline and one-year outcome scores for those eventually undergoing surgery in the conservative arm were not significantly different from the scores for those treated without surgical intervention; the same holds for the proportion of recovered patients (Table 1).

Univariate logistic regression models with surgery as the event of interest did not reveal a significant association with the classical anamnestic, neurological and radiological variables (Table 2). Univariately significant odds ratios were obtained for “initially recorded leg pain intensity” (VAS; \( p = 0.03 \)) and the “Kemp neuroforamen compression test” (\( p = 0.006 \)), as well as for “magnitude of preference for surgery” (\( p = 0.04 \)). After entering these variables into one multivariable logistic regression model and performing a backward stepwise analysis, VAS-leg pain intensity and severity of sciatica-specific disability were retained as significant factors. The adjusted Odds ratios were 1.7 (95% CI; 1.1–2.7) per VAS 20 mm incremental increase in pain and 1.8 (95% CI; 1.2–2.9) per 3 points deterioration on the RDQ score (Table 2).

Predictability of the risk to undergo delayed surgery seems high by the estimated odds ratios of leg pain intensity and disability, but the absolute risk never exceeds levels higher than 80% (Table 3). Hence the sensitivity of the combined information of the two scores at intake and randomization is only around 43% and the specificity is around 85% with a total probability of correctly classifying surgery being estimated at 70% (67% cross-validated).

Since odds ratios describe relative effect sizes only, the estimated absolute risks of surgery as a function of combinations of pain and disability scores of the study patients are presented (Fig. 1) to illustrate the estimated magnitude of the problem.

4. Discussion

Delayed surgery did not lead to any differences in patient outcome at one year when compared to those

Table 1
Baseline scores and outcomes per treatment group at 12 months in patients initially started with a conservative treatment regimen\(^a\)

<table>
<thead>
<tr>
<th></th>
<th>Conservative ((n = 87))</th>
<th>Late surgery ((n = 55))</th>
<th>(p)-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roland baseline(^b)</td>
<td>15.8 ± 3.8</td>
<td>17.1 ± 3.9</td>
<td>&gt;0.05</td>
</tr>
<tr>
<td>12 months</td>
<td>3.9 ± 5.2</td>
<td>4.0 ± 6.2</td>
<td>&gt;0.05</td>
</tr>
<tr>
<td>VAS-leg pain baseline (mm)(^c)</td>
<td>64.0 ± 21.3</td>
<td>65.4 ± 21.3</td>
<td>&gt;0.05</td>
</tr>
<tr>
<td>12 months</td>
<td>11.4 ± 17.5</td>
<td>13.4 ± 25.3</td>
<td>&gt;0.05</td>
</tr>
<tr>
<td>VAS-back pain baseline (mm)(^c)</td>
<td>30.2 ± 25.0</td>
<td>32.2 ± 32.3</td>
<td>&gt;0.05</td>
</tr>
<tr>
<td>12 months</td>
<td>17.4 ± 21.3</td>
<td>14.6 ± 24.8</td>
<td></td>
</tr>
<tr>
<td>Satisfactory recovery %(^d)</td>
<td>79</td>
<td>85</td>
<td>&gt;0.05 (Fisher’s exact test)</td>
</tr>
</tbody>
</table>

\(^a\) Means and standard deviations.

\(^b\) The Roland disability questionnaire for sciatica is a disease-specific disability scale that measures functional status in patients with pain in the leg or back. Scores range from 0 to 23, with higher scores indicating worse functional status.

\(^c\) The intensity of pain was indicated on a horizontal 100 mm visual analogue scale, with 0 representing no pain and 100 the worst pain ever experienced.

\(^d\) Likert global perceived recovery is defined by a 7-point scale “Worse” to “Complete” recovery. Satisfactory recovery is defined as complete or nearly complete recovery using the Likert 7-point scale.
Table 2
Uni- and multivariate analysis of baseline variables to determine predictive value on delayed surgery

<table>
<thead>
<tr>
<th>Variable</th>
<th>Comparison (%)</th>
<th>Univariate analysis</th>
<th>Multivariate analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>OR</td>
<td>95% CI</td>
</tr>
<tr>
<td>Gender</td>
<td>Female (32) vs male (68)</td>
<td>0.80</td>
<td>0.37–1.70</td>
</tr>
<tr>
<td>Age (years)</td>
<td>≥40 (63) vs &lt;40 (37)</td>
<td>1.04</td>
<td>0.51–2.13</td>
</tr>
<tr>
<td>Job†</td>
<td>Mentally (69) vs physical (31)</td>
<td>0.56</td>
<td>0.29–1.09</td>
</tr>
<tr>
<td>Housewife</td>
<td>Yes (10) vs no (90)</td>
<td>2.87</td>
<td>0.93–8.81</td>
</tr>
<tr>
<td>Spouse/partner</td>
<td>Yes (78) vs no (22)</td>
<td>1.38</td>
<td>0.58–3.28</td>
</tr>
<tr>
<td>Children</td>
<td>Yes (68) vs no (32)</td>
<td>1.70</td>
<td>0.78–3.69</td>
</tr>
<tr>
<td>Smoking</td>
<td>Yes (39) vs no (61)</td>
<td>1.00</td>
<td>0.49–2.04</td>
</tr>
<tr>
<td>Queetele index</td>
<td>Per 1 point increments</td>
<td>1.10</td>
<td>0.98–1.23</td>
</tr>
<tr>
<td>Start sciatica</td>
<td>Slowly (67) vs acute (33)</td>
<td>1.66</td>
<td>0.80–3.43</td>
</tr>
<tr>
<td>Sciatica provoked by sitting</td>
<td>Yes (75) vs no (25)</td>
<td>1.27</td>
<td>0.56–2.86</td>
</tr>
<tr>
<td>Coughing, sneeze</td>
<td>Yes (72) vs no (28)</td>
<td>1.13</td>
<td>0.52–2.46</td>
</tr>
<tr>
<td>Straight leg raising b</td>
<td>Positive &lt;60° (74) vs negative ≥60° (26)</td>
<td>1.52</td>
<td>0.66–3.49</td>
</tr>
<tr>
<td>Crossed leg raising b</td>
<td>Positive (57) vs negative (43)</td>
<td>0.61</td>
<td>0.30–1.21</td>
</tr>
<tr>
<td>Kemp’s sign c</td>
<td>Positive (56) vs Negative (44)</td>
<td>2.75</td>
<td>1.31–5.79</td>
</tr>
<tr>
<td>Finger-ground in cm</td>
<td>&gt;30 (59) vs ≤30 (41)</td>
<td>1.55</td>
<td>0.75–3.23</td>
</tr>
<tr>
<td>Bragard’s test d</td>
<td>Positive (28) vs negative (72)</td>
<td>1.52</td>
<td>0.68–3.40</td>
</tr>
<tr>
<td>Sensory disturbance</td>
<td>Yes (91) vs no (9)</td>
<td>2.80</td>
<td>0.59–13.4</td>
</tr>
<tr>
<td>MRI-level herniation</td>
<td>L5/S1 (66) vs L4/L5 and L3/L4 (34)</td>
<td>0.95</td>
<td>0.49–1.83</td>
</tr>
<tr>
<td>MRI-sequester</td>
<td>Yes (42) vs no (58)</td>
<td>0.84</td>
<td>0.43–1.63</td>
</tr>
<tr>
<td>Preference surgery</td>
<td>Mild (66) vs strong (34)</td>
<td>0.48</td>
<td>0.23–0.98</td>
</tr>
<tr>
<td>Roland disability (0–23) f</td>
<td>Per 3 points increments</td>
<td>2.08</td>
<td>1.24–3.49</td>
</tr>
<tr>
<td>VAS-leg pain (0–100 mm) f</td>
<td>Per 20 mm increments</td>
<td>1.65</td>
<td>1.05–2.6</td>
</tr>
<tr>
<td>Roland difference (T2w–T0) f</td>
<td>Worse or equal (37) vs improved (63)</td>
<td>1.51</td>
<td>0.60–3.81</td>
</tr>
<tr>
<td>VAS difference (T2W–T0) f</td>
<td>Worse or equal (35) vs improved (65)</td>
<td>1.37</td>
<td>0.56–3.36</td>
</tr>
</tbody>
</table>

Bold values represent significance p < 0.10 to retain variables in the logistic regression model.
* Intellectual or physical jobs were defined by the patient.
† Lasegue’s sign was defined positive if the examiner observed a typically dermatomal area of pain reproduction and pelvic muscle resistance during unilateral provocative straight leg raising below an angle of 60 degrees, and crossed positive if the same experience was noted raising the other leg below 90 degrees.
‡ Kemp sign, also called neuroforamen compression test, is positive if the radicular leg pain is provoked by ipsilateral passive lateroflexion and extension of the lumbar spine by the research nurse.
§ The test of Bragard, indicative of nerve root compression is considered to be positive if the patient experiences aggravation of the provoked radicular pain during maximal straight leg raising.
' The Roland disability questionnaire for sciatica is a disease specific disability scale that measures functional status in patients with pain in the leg or back. Scores range from 0 to 23, with higher scores indicating worse functional status.
* The intensity of pain was indicated on a horizontal 100 mm visual analogue scale, with 0 representing no pain and 100 the worst pain ever experienced.

High initial pain and disability scores were found to be predictive of a higher chance on delayed surgery in this study. However, these indicators are not yet used for the regular care of sciatica patients [11] but may be valuable in the decision process to opt for early surgery or for prolonged conservative care. Indeed, if initial scores after 6–12 weeks of persistent sciatica correspond to severe disability plus high pain intensity and do not regress after a few more weeks of “wait-and-see”, one may infer that the risk of surgery at a later stage is high. These patients might consider surgery without further delay to reduce the period of suffering and absence from work.

Previously, we described the lack of interaction between initial pain intensities and the allocated timing of surgery on the speed of recovery [12]. These analyses, however, were bound to an Intent-To-Treat methodology and the current 55 surgical patients were, thus, part of the prolonged conservative treatment arm as it was a pragmatic randomized controlled trial comparing two different timing-of-surgery strategies. The current analysis describes the predictive value of pain intensity for surgery performed at a later stage, instead of the possible interaction effects on the speed of recovery or outcome per se.

The current results are clear but some restrictions in study design must be considered. Patients were recruited from neurological outpatient clinics after the usual refer-
ral by primary care physicians who stated that their patient had persistent sciatica and requested for surgery. As these patients had a strong or moderate preference for surgery this selective study accrual resulted in an absence of patients who preferred conservative care. This imbalance and lack of contrast between preferences for both treatment options might have underestimated the real effect and possibly led to the fact that this variable was not retained in the final model.

One may concur that disability and pain are measured by subjective questionnaires which, except for study purposes, are not yet used for the daily care of spine patients. However, due to a lack of diagnostic, prognostic and outcome properties of neurological and radiological signs, these validated low back disease-specific questionnaires might be the best tools we have today to fulfil the request of society to measure quality of care. The restricted population size prohibited the performance of a more refined analysis with the goal to define threshold scores of pain intensity or disability, above which surgery is inevitable. Future studies might address these critical cut-off points to anticipate failure or success of prolonged conservative care.

So far, this is the first study that thoroughly analyzed variables which possibly affect the risk to undergo surgery in a conservative treatment regime for patients with 6–12 weeks of severe sciatica [9]. Although our findings may not be surprising for most physicians, we do not use these instruments for the regular care of sciatica patients. Obviously it is important to quantify the influence of pain and disability on the timing of surgery. Since timing of surgery did not influence outcome at 1 year, the main indication for early surgery is to shorten the period of suffering. High pain intensities and disability scores complemented by personal preferences are valid arguments in support of the choice of surgery [5].

Despite maximal efforts of patients and physicians, surgery seems inevitable for a considerable proportion of a conservatively treated cohort. Compared to those with “tolerable” pain and disability, patients, who experience more intense leg pain and worse disability scores, run a higher risk of prolonged suffering and undergo delayed surgery and therefore might urge the spine surgeon to opt for earlier surgery to shorten their period of illness. Obviously we still cannot reliably estimate exactly which patient will receive surgery during the follow-up period although the prediction is significantly improved when using these scales.

5. Contributors

All authors contributed to the interpretation of the results, drafting and revision of the article. Additionally: W.C. Peul contributed to the design, funding application, management of the trial, data collection and analysis; R. Brand contributed to data-management and analysis; R.T.W.M. Thomeer contributed to the design and funding application; B.W. Koes contributed to the design and analysis.

6. Conflict of interest statement

We declare that we have no conflict of interest.

Acknowledgments

This study was funded by grants from the Netherlands Organisation for Health Research and Develop-
References


