

# BMJ Open Surgical versus conservative treatment for lumbar disc herniation: a prospective cohort study

Marinella Gugliotta,<sup>1</sup> Bruno R da Costa,<sup>2</sup> Essam Dabis,<sup>3</sup> Robert Theiler,<sup>3,4</sup> Peter Jüni,<sup>5,6</sup> Stephan Reichenbach,<sup>7,8</sup> Hans Landolt,<sup>1</sup> Paul Hasler<sup>3</sup>

**To cite:** Gugliotta M, da Costa BR, Dabis E, *et al*. Surgical versus conservative treatment for lumbar disc herniation: a prospective cohort study. *BMJ Open* 2016;**6**:e012938. doi:10.1136/bmjopen-2016-012938

► Prepublication history for this paper is available online. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2016-012938>).

MG and BRdC contributed equally and share first authorship.

Received 9 June 2016  
Revised 27 September 2016  
Accepted 30 September 2016



CrossMark

For numbered affiliations see end of article.

**Correspondence to**  
Professor Paul Hasler;  
[paul.hasler@ksa.ch](mailto:paul.hasler@ksa.ch)

## ABSTRACT

**Objectives:** Evidence comparing the effectiveness of surgical and conservative treatment of symptomatic lumbar disc herniation is controversial. We sought to compare short-term and long-term effectiveness of surgical and conservative treatment in sciatica symptom severity and quality of life in patients with lumbar disc herniation in a routine clinical setting.

**Methods:** A prospective cohort study of a routine clinical practice registry consisting of 370 patients. Outcome measures were the North American Spine Society questionnaire and the 36-Item Short-Form Health Survey to assess patient-reported back pain, physical function, neurogenic symptoms and quality of life. Primary outcomes were back pain at 6 and 12 weeks. Standard open discectomy was assessed versus conservative interventions at 6, 12, 52 and 104 weeks. We filled in missing outcome variable values with multiple imputation, accounted for repeated measures within patients with mixed-effects models and adjusted baseline group differences in relevant prognostic indicators by inverse probability of treatment weighting.

**Results:** Surgical treatment patients reported less back pain at 6 weeks than those receiving conservative therapy ( $-0.97$ ; 95% CI  $-1.89$  to  $-0.09$ ), were more likely to report  $\geq 50\%$  decrease in back pain symptoms from baseline to 6 weeks (48% vs 17%, risk difference: 0.34; 95% CI 0.16 to 0.47) and reported less physical function disability at 52 weeks ( $-3.7$ ; 95% CI  $-7.4$  to  $-0.1$ ). The other assessments showed minimal between-group differences with CIs, including the null effect.

**Conclusions:** Compared with conservative therapy, surgical treatment provided faster relief from back pain symptoms in patients with lumbar disc herniation, but did not show a benefit over conservative treatment in midterm and long-term follow-up.

## INTRODUCTION

Sciatica is one of the most debilitating types of pain emanating from the low back, with a lifetime incidence of  $\sim 30\%$ .<sup>1 2</sup> Sciatica is a disorder caused by pressure on or irritation of the nerve root. Main symptoms and signs

## Strengths and limitations of this study

- We included in the present study consecutively sampled patients from a routine clinical practice registry who were followed-up until 2 years after treatment. Thus, as opposed to randomised controlled trials of head-to-head comparisons between surgical and conservative treatment for lumbar disc herniation, which are hampered by a large number of crossovers and dropouts, our results are directly generalised to routine clinical settings.
- By using inverse probability weighting, we were able to minimise the risk of confounding by indication to mimic results of a randomised controlled experiment, while maintaining the generalisability of observational studies.
- Because of the observational nature of our investigation, results of the present investigation must be carefully interpreted because of the risk of residual confounding by indication.

include unilateral leg pain that is worse than concomitant low back pain, pain radiating beyond the knee, decreased muscle strength in a myotomal distribution and sensory deficits in a dermatomal distribution.<sup>3 4</sup> Compared with patients with localised low back pain only, those with sciatica generally have more persistent and severe pain, worse prognosis, consume more healthcare resources and are disabled and absent from work for a longer period of time.<sup>1</sup>

Lumbar disc herniation is considered to be one of the main causes of sciatica, and lumbar discectomy is the most popular surgical procedure performed in patients with sciatica in the USA.<sup>1 2 5-7</sup> Lumbar disc herniation also occurs in asymptomatic patients, and often spontaneously regresses without surgery.<sup>8-10</sup> Conservative treatment, including physical therapy, pharmacological treatment and infiltrations, is an alternative approach for symptomatic patients; 90% of sciatica cases due to lumbar disc herniation

resolve with conservative measures.<sup>5–9</sup> Conservative treatment of lumbar disc herniation has a lower risk of complications than surgery and is preferred by the vast majority of patients.<sup>11</sup>

Several studies have compared the effectiveness of surgical and conservative treatment in patients with lumbar disc herniation associated with sciatica, but methodological aspects limit interpretation of their results. Observational cohort studies have typically differed in important baseline prognostic indicators between treatment group and their results were thus more prone to confounding.<sup>12–14</sup> Randomised controlled trials (RCTs) are less prone to generate confounded results. However, in RCTs comparing surgical with conservative treatment, a large proportion of patients randomly allocated to conservative treatment actually received surgical treatment right after randomisation or after an initial period of conservative treatment (26–54%).<sup>7 15–17</sup> Therefore, RCTs are actually mainly comparing early surgery with conservative treatment and delayed surgery in selected patients, as was referred to by Peul *et al.*<sup>18 19</sup> In addition, some researchers have questioned whether patients willing to participate in RCTs of surgery versus conservative treatment are representative of patients commonly seen in clinical practice.<sup>20–23</sup>

In order to present results that are more representative of routine clinical care while minimising the risk of confounded results, we conducted a properly sized observational cohort study in a routine clinical setting using consecutive sampling, in which baseline differences in prognostic indicators were accounted for in an analysis with inverse probability weighting closely mimicking an RCT, with the aim of comparing the effect of surgical and conservative treatment on sciatica symptom severity and quality of life in patients with lumbar disc herniation.

## METHODS

### Study design

This was a prospective observational cohort study based on the routine clinical practice registry of the Neurosurgery and Rheumatology Departments of the Cantonal Hospital in Aarau, Switzerland. All eligible patients were consecutively invited to participate in the study. Recruitment took place from May 2003 to December 2007.

### Ethics and consent

This study is part of a quality management programme on anonymised patients; therefore, no institutional review board approval is required in Switzerland.

### Patient population

Patients were considered eligible if they were at least 18 years old, were diagnosed with symptomatic low-back pain due to lumbar disc herniation and associated radicular pain and showed signs of nerve root irritation

(positive straight leg raise or femoral nerve tension tests) and/or neurological deficits (asymmetrical depressed reflexes or motor or sensory deficits in corresponding myotomal or dermatomal distribution) requiring hospitalisation. Diagnosis was verified by advanced spinal imaging (MRI or CT) with disk herniation at a level and side corresponding to the clinical symptoms and physical findings. The study population included all patients willing to participate in a standardised clinical follow-up programme comprising consultations and patient-based outcome measures, and who had outcome data available at 6 or 12-week follow-up measurement.

### Treatment interventions

The assignment to treatment interventions was decided by physicians based on patients' clinical indications. Surgical treatment was a standard open discectomy as described by Delamarter and McCulloch and by Spengler, with examination of the involved nerve root performed using a microscope, with the patient under general anaesthesia and in the knee–chest position.<sup>24 25</sup> After a midline incision, the paraspinal muscles were reflected and the interlaminar space was entered.<sup>24</sup> If necessary, the medial border of the superior facet was removed to provide an unobstructed view of the involved nerve root. Using a small annular incision, the herniated disc fragment was removed, the spinal canal was inspected and the foramen and recessus probed for residual disk or bony pathology.<sup>25</sup> The nerve root was then decompressed, with the purpose of leaving it freely mobile.

The conservative treatment consisted of ergonomic instruction, active physical therapy, education/counseling with instructions for home-based exercise, and non-steroidal anti-inflammatory drugs if tolerated. Patients with insufficient analgesic response were prescribed additional opioids. Those with an inadequate response to opioids were offered epidural infiltrations, CT-guided periradicular infiltrations<sup>26</sup> and, in the case of continued inadequate response or recurrence, CT-guided pulsed radiofrequency therapy of the affected nerve root.<sup>27</sup> If conservative treatment failed, which was ascertained on a case-by-case basis, surgery was provided as an option.

### Outcome measures

Sciatica symptom severity was assessed using the North American Spine Society (NASS) questionnaire, and quality of life was assessed using the 36-Item Short-Form Health Survey (SF-36). The primary outcome measures were the changes in score from baseline to weeks 6 and 12 as assessed by the back pain subscale of the NASS questionnaire using a scale from 0 to 10.<sup>28</sup> Secondary outcome measures included the NASS neurogenic symptoms subscale (which addresses leg or foot pain, numbness and tingling on a scale from 0 to 30), the NASS function subscale (which addressed disability because of pain on a scale from 0 to 45), the SF-36 V.1 physical and mental subscales (scale from 0 to 100)<sup>29</sup> and the

proportion of patients responding to treatment, defined as a 50% reduction in baseline scores of the NASS pain subscale. Lower scores indicate a better outcome for the NASS questionnaire, and a worse outcome for the SF-36. The validated German language versions of the NASS<sup>30</sup> and SF-36<sup>31</sup> questionnaires were used in an audiovisual touchscreen version (Qualitouch, Zürich, Switzerland).<sup>32</sup>

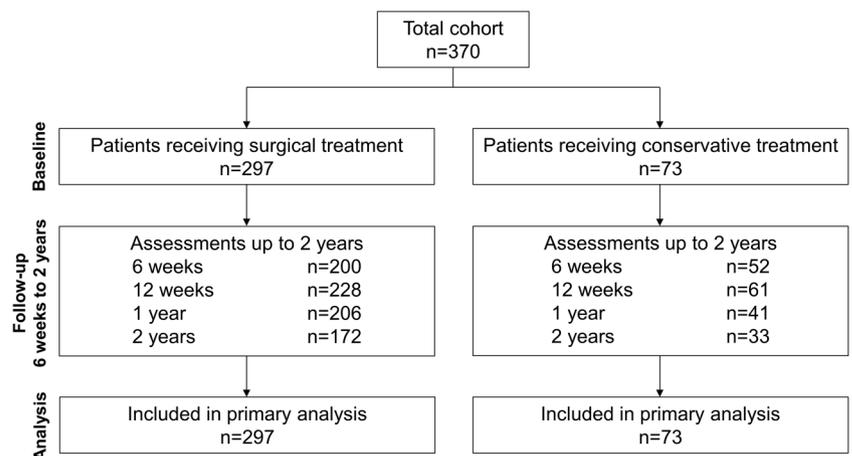
All outcomes were prospectively assessed at baseline and at 6, 12, 52 and 104 weeks. Outcome measures were specified prior to statistical analysis.

### Statistical analysis

Baseline and efficacy data are presented as counts and percentages for dichotomous variables and as means

and SDs for continuous variables. Between-group comparisons of baseline data were performed using Pearson's  $\chi^2$  for dichotomous variables and Student's t-test for continuous variables. Only patients with complete primary outcome data, that is, NASS back pain assessed at 6 or 12-week follow-up, were considered in the analysis. We accounted for missing data by using multiple imputation with baseline efficacy variables, age, body mass index, gender, social status, employment status, country of origin and treatment group as explanatory variables in the imputation model, to create 20 imputed data sets. For each patient, we estimated propensity scores for receiving surgical treatment using a probit model that included baseline efficacy variables,

**Figure 1** Flow of patients through the different stages of the trial.



**Table 1** Patient characteristics at baseline

|                                   | Surgical (n=297) | Conservative (n=73) | p Value | Adjusted p Value* |
|-----------------------------------|------------------|---------------------|---------|-------------------|
| Age (years, mean±SD)              | 50.4±13.5        | 49.8±12.7           | 0.66    | 0.69              |
| BMI (kg/m <sup>2</sup> , mean±SD) | 26.7±4.5         | 27.3±4.4            | 0.27    | 0.79              |
| Male gender, n (%)                | 168 (56.6)       | 43 (58.9)           | 0.72    | 0.74              |
| Social class, n (%)               |                  |                     | 0.065   | 0.96              |
| Blue collar                       | 132 (44.4)       | 38 (52.1)           |         |                   |
| White collar                      | 74 (24.9)        | 12 (16.4)           |         |                   |
| Not working                       | 33 (11.1)        | 14 (19.2)           |         |                   |
| Retired                           | 58 (19.5)        | 9 (12.3)            |         |                   |
| Country of origin, n (%)          |                  |                     | ≤0.001  | 0.94              |
| Switzerland                       | 255 (85.9)       | 46 (63.0)           |         |                   |
| Foreign                           | 42 (14.1)        | 27 (37.0)           |         |                   |
| Social, n (%)                     |                  |                     | 0.84    | 0.92              |
| Family                            | 247 (83.2)       | 60 (82.2)           |         |                   |
| Single                            | 50 (16.8)        | 13 (17.8)           |         |                   |
| NASS, mean±SD                     |                  |                     |         |                   |
| Pain†                             | 7.8±2.2          | 7.7±2.7             | 0.64    | 0.88              |
| Neurogenic symptoms‡              | 18.8±6.7         | 17.3±8.3            | 0.098   | 0.84              |
| Function§                         | 24.7±8.4         | 25.8±9.3            | 0.32    | 0.72              |
| SF-36, mean±SD                    |                  |                     |         |                   |
| Physical¶                         | 26.5±7.0         | 27.3±6.8            | 0.39    | 1.00              |
| Mental¶                           | 48.7±12.5        | 46.8±14.5           | 0.26    | 0.93              |

\*p Value after adjustment with inverse probability weighting.

†Possible NASS subscale scores range from †0 to 10, ‡0 to 30 and §0 to 45; lower scores indicate lower severity of symptoms.

¶Possible score range for both SF-36 subscales is 0 to 100; higher scores indicate lower severity of symptoms.

BMI, body mass index; NASS, North American Spine Society lumbar spine outcome assessment instrument; SF-36, 36-Item Short-Form Health Survey.

age, body mass index, gender, social status, employment status and country of origin as explanatory variables. Propensity scores were then used to derive inverse probability of treatment weights, with the inverse of the propensity score as analytic weights for patients in the surgical group and the inverse of 1 minus the propensity score for patients in the conservative group.<sup>33–34</sup> To account for repeated measures within patients across multiple follow-up assessments, we used linear or logistic mixed-effects models adjusted for the inverse probability of treatment weighting to derive, for each outcome measure at each follow-up time, group-specific means or proportions with 95% CIs and between-group differences in means or proportions with 95% CI. Statistical analyses were performed with STATA release 12.1 (Stata Corp, College Station, Texas, USA). All p values are two-sided.

## RESULTS

### Study flow and patient characteristics

Three hundred and seventy patients were consecutively sampled and assigned to surgical (n=297) or

conservative (n=73) treatment (figure 1). Table 1 shows baseline clinical characteristics; patients receiving surgical treatment tended to have more severe neurogenic symptoms at baseline (p=0.098) and were more likely to be Swiss citizens (p≤0.001) from a higher social class (p=0.065). Adjusted p values in table 1 indicate that there is no evidence of significant differences between groups for all variables at baseline after adjustment using inverse probability weighting (p≥0.72).

### NASS pain

Table 2 shows that 6 weeks after the end of treatment, patients in the surgical group had less pain than patients in the conservative group (−1.0, 95% CI −1.9 to −0.1). However, we observed a constant decrease in between-group differences in pain scores in all subsequent follow-up assessments, with CIs overlapping the null effect (table 2 and figure 2). Similarly, 34% (95% CI 16% to 47%) more patients in the surgical group responded to treatment at 6 weeks after the end of treatment, but the 95% CI for between-group comparison in

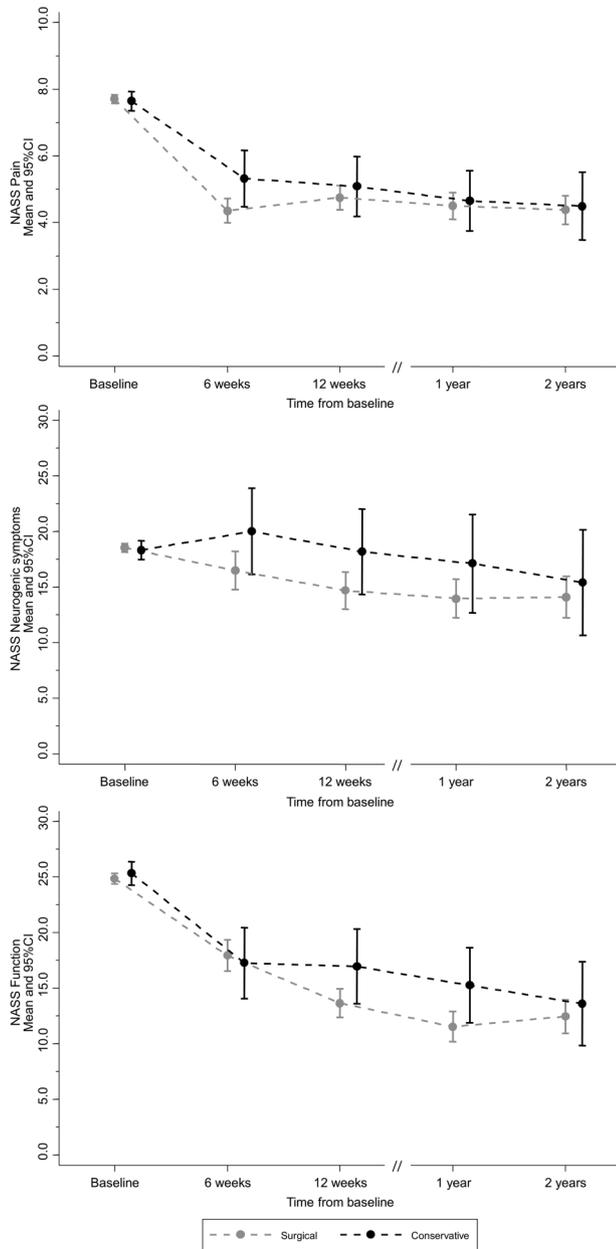
**Table 2** Primary and secondary outcomes in all follow-up assessments

|   | Surgical (n=297)    | Control (n=73)      | Difference          |
|---|---------------------|---------------------|---------------------|
| NASS back pain*, mean (95% CI)          |                     |                     |                     |
| 6 weeks                                 | 4.4 (4.0 to 4.7)    | 5.3 (4.5 to 6.2)    | −1.0 (−1.9 to −0.1) |
| 12 weeks                                | 4.8 (4.4 to 5.1)    | 5.1 (4.2 to 6.0)    | −0.3 (−1.3 to 0.6)  |
| 1 year                                  | 4.5 (4.1 to 4.9)    | 4.7 (3.8 to 5.6)    | −0.2 (−1.2 to 0.8)  |
| 2 years                                 | 4.4 (4.0 to 4.8)    | 4.5 (3.5 to 5.5)    | −0.1 (−1.2 to 1.0)  |
| Response to treatment, % (95% CI)       |                     |                     |                     |
| 6 weeks                                 | 48 (36 to 60)       | 17 (8 to 25)        | 34 (16 to 47)       |
| 12 weeks                                | 35 (24 to 46)       | 24 (13 to 35)       | 9 (−6 to 27)        |
| 1 year                                  | 42 (30 to 55)       | 35 (15 to 54)       | 3 (−15 to 30)       |
| 2 years                                 | 44 (30 to 58)       | 49 (21 to 77)       | −17 (−36 to 26)     |
| NASS neurogenic†, mean (95% CI)         |                     |                     |                     |
| 6 weeks                                 | 16.5 (14.8 to 18.2) | 20.0 (16.2 to 23.9) | −3.5 (−7.7 to 0.7)  |
| 12 weeks                                | 14.7 (13.0 to 16.4) | 18.2 (14.3 to 22.0) | −3.5 (−7.7 to 0.7)  |
| 1 year                                  | 14.0 (12.2 to 15.7) | 17.1 (12.7 to 21.5) | −3.1 (−7.9 to 1.6)  |
| 2 years                                 | 14.1 (12.2 to 16.0) | 15.4 (10.6 to 20.2) | −1.3 (−6.3 to 3.7)  |
| NASS function‡, mean (95% CI)           |                     |                     |                     |
| 6 weeks                                 | 17.9 (16.5 to 19.4) | 17.3 (14.1 to 20.5) | 0.7 (−2.8 to 4.2)   |
| 12 weeks                                | 13.7 (12.4 to 14.9) | 17.0 (13.6 to 20.3) | −3.3 (−6.9 to 0.3)  |
| 1 year                                  | 11.5 (10.2 to 12.9) | 15.3 (11.9 to 18.6) | −3.7 (−7.4 to −0.1) |
| 2 years                                 | 12.5 (10.9 to 14.0) | 13.6 (9.8 to 17.4)  | −1.1 (−5.2 to 2.9)  |
| SF-36 physical function§, mean (95% CI) |                     |                     |                     |
| 6 weeks                                 | 33.2 (31.8 to 34.6) | 36.3 (33.4 to 39.1) | −3.1 (−6.4 to 0.1)  |
| 12 weeks                                | 38.0 (36.7 to 39.3) | 37.8 (34.4 to 41.2) | 0.2 (−3.5 to 3.8)   |
| 1 year                                  | 41.0 (39.4 to 42.5) | 40.3 (36.7 to 43.8) | 0.7 (−3.2 to 4.6)   |
| 2 years                                 | 42.1 (40.3 to 44.0) | 42.8 (39.1 to 46.4) | −0.6 (−4.7 to 3.5)  |
| SF-36 mental function§, mean (95% CI)   |                     |                     |                     |
| 6 weeks                                 | 47.9 (45.8 to 49.9) | 48.2 (44.3 to 52.2) | −0.4 (−4.9 to 4.1)  |
| 12 weeks                                | 49.3 (47.6 to 51.0) | 47.5 (43.6 to 51.3) | 1.8 (−2.3 to 6.0)   |
| 1 year                                  | 50.3 (48.7 to 51.9) | 47.9 (43.9 to 51.8) | 2.4 (−1.9 to 6.7)   |
| 2 years                                 | 49.1 (47.0 to 51.1) | 47.7 (43.5 to 52.0) | 1.3 (−3.5 to 6.1)   |

Possible NASS subscale scores range from \*0 to 10, †0 to 30 and ‡0 to 45; lower scores indicate lower severity of symptoms.

§Possible score range for both SF-36 subscales is 0 to 100; higher scores indicate lower severity of symptoms.

NASS, North American Spine Society lumbar spine outcome assessment instrument; SF-36, 36-Item Short-Form Health Survey.

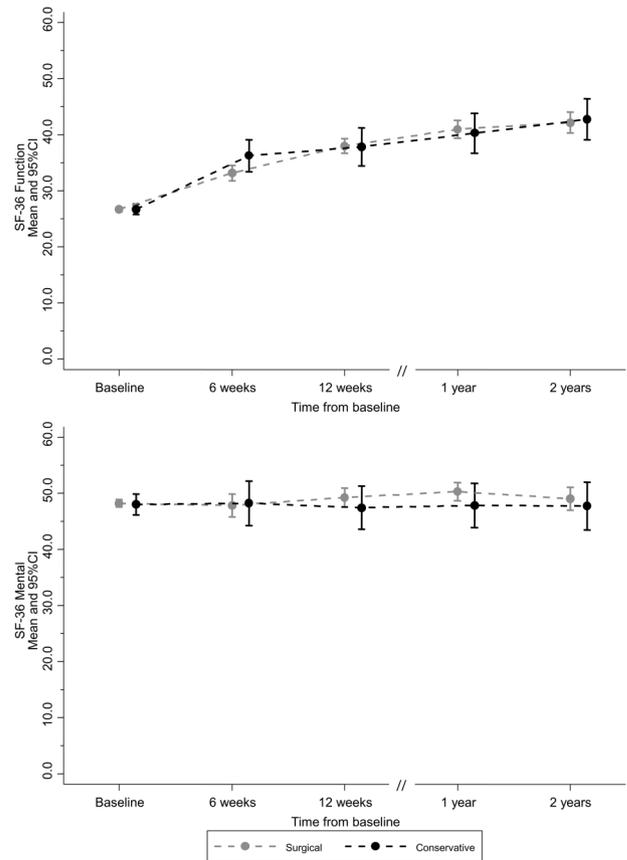


**Figure 2** Mean scores with 95% CI of the subscales of the North American Spine Society (NASS) questionnaire at baseline and subsequent follow-up assessments.

all subsequent follow-up assessments included the null effect (table 2).

### NASS neurogenic symptoms and NASS function

The neurogenic symptoms of patients in the surgical group tended to improve faster (6 and 12 weeks:  $-3.5$ , 95% CI  $-7.7$  to  $0.7$ ), but we observed no difference in the long term (2 years:  $-1.3$ , 95% CI  $-6.3$  to  $3.7$ ). There was no difference between groups in physical function in the first follow-up assessment at 6 weeks ( $0.7$ , 95% CI  $-2.8$  to  $4.2$ ). Patients in the surgical group reported lower functional impairment at 1 year ( $-3.7$ , 95% CI  $-7.4$  to  $-0.1$ ), but this difference was not sustained at



**Figure 3** Mean scores with 95% CI of the subscales of the 36-Item Short-Form Health Survey (SF-36) at baseline and subsequent follow-up assessments.

the 2-year follow-up assessment ( $-1.1$ , 95% CI  $-5.2$  to  $2.9$ ) (table 2 and figure 2).

### 36-Item Short-Form Health Survey

There was little evidence of a difference in quality of life between groups throughout the study. Patients in the surgical group tended to score better on the SF-36 physical subscale in the short term (6 weeks:  $-3.1$ , 95% CI  $-6.4$  to  $0.1$ ), but the difference was minimal in the long term (2 years:  $-0.6$ , 95% CI  $-4.7$  to  $3.5$ ) (table 2). Scores of the SF-36 mental subscale were similar in both groups in all follow-up assessments (table 2 and figure 3).

### DISCUSSION

We found no evidence that surgical treatment, when compared with conservative treatment, reduced the severity of sciatica symptoms or improved the quality of life of patients with lumbar disc herniation in the medium or long term. Pain was relieved more quickly in patients who received surgical treatment (evident at the 3-week follow-up), but the difference between groups was no longer present after 3 months. Patients in the surgical group did report less physical impairment at the 1-year follow-up, but not in previous or subsequent assessments. Surgery was not more effective for the

treatment of neurogenic symptoms or the improvement of quality of life over the course of the study.

Faster improvement in pain symptoms with surgical treatment is a common finding in comparisons with conservative treatment in patients with lumbar disc herniation. Previous observational studies have also found that back pain is reduced more quickly with surgical treatment.<sup>12–14 35</sup> Findings regarding neurogenic symptoms, physical function, and quality of life, however, are not as consistently reported by other observational studies. As opposed to our findings, previous observational studies have found benefits of surgical treatment in these outcomes at short-term and long-term follow-up. The discrepancy between our findings and those of previous studies may be due to differences in eligibility criteria and methods of outcome assessment, a more effective control intervention and a different approach to statistical analysis to control for confounding by indication.

Interestingly, results of our observational cohort conducted in a routine care setting more closely resemble those reported by previous RCTs.<sup>7 15–19</sup> RCTs have also typically reported quicker pain reduction in patients who received surgery, but no clear benefit of surgery over conservative treatment at long-term assessments of neurogenic symptoms, physical function or quality of life. However, in the classic trial by Weber,<sup>15</sup> the beneficial effect of surgical treatment lasted longer than in other trials; treatment effects of surgical and conservative treatment only became similar after 4 years of follow-up, and remained similar until the 10-year final follow-up.

The observational nature of our investigation limits our ability to interpret its findings. In observational clinical studies, results are likely to be influenced by confounding by indication. Patients with a worse prognosis at baseline are more likely to be allocated by physicians to surgical intervention, and indeed, this was the case in our study, in which patients in the surgical group showed a trend towards worse neurogenic symptoms at baseline ( $p=0.098$ ). However, the methods we used for statistical analysis allowed us to mimic a randomised controlled experiment.<sup>29</sup> This method of analysis, that is, inverse probability weighting, assumes that the probability of being allocated to surgical or conservative treatment depends mainly on the prognostic indicators we included in our analysis. Although this assumption may be inaccurate in some cases, our results are remarkably similar to those reported in previous RCTs. Surgical RCTs are commonly criticised for lack of generalisability because patients who agree to be randomised in these trials may not be representative of those seen in clinical practice. The results of the present investigation do not suffer from this limitation, because no randomisation took place. Moreover, a significant number of patients dropped out of our study due to loss to follow-up, especially by latter time points. We conducted multiple imputation as an attempt to include in our analysis patients with missing outcome data; however, no

statistical technique is likely to completely solve the problem of missing data, and it is always better to have observed data as opposed to imputed data for all patients included in the analysis.

Surgical and conservative treatments had long-term beneficial effects on sciatica symptoms in patients with lumbar disc herniation. Compared with conservative treatment, surgical treatment relieved back pain faster, but no relevant clinical difference was observed after 3 months. Surgical treatment may thus be attractive to patients with debilitating pain symptoms who seek quick relief, or who did not experience satisfactory improvement with conservative treatment.

#### Author affiliations

<sup>1</sup>Department of Neurosurgery, Cantonal Hospital Aarau, Aarau, Switzerland

<sup>2</sup>Institute of Primary Health Care (BIHAM), University of Bern, Bern, Switzerland

<sup>3</sup>Department of Rheumatology, Cantonal Hospital Aarau, Aarau, Switzerland

<sup>4</sup>Department of Rheumatology, Stadtspital Triemli, Zürich, Switzerland

<sup>5</sup>Applied Health Research Centre (AHRC), University of Toronto, Toronto, Ontario, Canada

<sup>6</sup>Department of Medicine, University of Toronto, Toronto, Ontario, Canada

<sup>7</sup>Institute of Social and Preventive Medicine (ISPM), University of Bern, Bern, Switzerland

<sup>8</sup>Department of Rheumatology, Immunology and Allergology, Inselspital Bern University Hospital, Bern, Switzerland

**Contributors** PH, HL, RT and MG conceived and designed the experiments; MG, ED, RT, HL and PH performed the experiments; BRdC, PJ and SR analysed the data; MG, ED, RT, HL, PH, BRdC, PJ and SR contributed to the writing of the manuscript.

**Funding** This study was funded by the Hugo and Elsa Isler Foundation, Aarau, Switzerland.

**Competing interests** PJ has received research grants to the institution from Astra Zeneca, Biotronik, Biosensors International, Eli Lilly and The Medicines Company, and serves as unpaid member of the steering group of trials funded by Astra Zeneca, Biotronik, Biosensors, St Jude Medical and The Medicines Company. The other authors report no conflict of interest.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** No additional data are available.

**Open Access** This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

#### REFERENCES

1. Konstantinou K, Dunn KM. Sciatica: review of epidemiological studies and prevalence estimates. *Spine (Phila Pa 1976)* 2008;33:2464–72.
2. Stafford MA, Peng P, Hill DA. Sciatica: a review of history, epidemiology, pathogenesis, and the role of epidural steroid injection in management. *Br J Anaesth* 2007;99:461–73.
3. Koes BW, van Tulder MW, Peul WC. Diagnosis and treatment of sciatica. *BMJ* 2007;334:1313–17.
4. Pinto RZ, Maher CG, Ferreira ML, et al. Drugs for relief of pain in patients with sciatica: systematic review and meta-analysis. *BMJ* 2012;344:e497.
5. Gibson JN, Waddell G. Surgical interventions for lumbar disc prolapse. *Cochrane Database Syst Rev* 2007(2):CD001350.
6. Taylor VM, Deyo RA, Cherkin DC, et al. Low back pain hospitalization. Recent United States trends and regional variations. *Spine (Phila Pa 1976)* 1994;19:1207–12. discussion 13.

7. Weinstein JN, Tosteson TD, Lurie JD, *et al.* Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT): a randomized trial. *JAMA* 2006;296:2441–50.
8. Boden SD, Davis DO, Dina TS, *et al.* Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects. A prospective investigation. *J Bone Joint Surg Am* 1990;72:403–8.
9. el Barzouhi A, Vleggeert-Lankamp CL, Lycklama à Nijeholt GJ, *et al.* Magnetic resonance imaging in follow-up assessment of sciatica. *N Engl J Med* 2013;368:999–1007.
10. Jensen MC, Kelly AP, Brant-Zawadzki MN. MRI of degenerative disease of the lumbar spine. *Magn Reson Q* 1994;10:173–90.
11. Deyo RA, Cherkin DC, Weinstein J, *et al.* Involving patients in clinical decisions: impact of an interactive video program on use of back surgery. *Med Care* 2000;38:959–69.
12. Atlas SJ, Deyo RA, Keller RB, *et al.* The Maine Lumbar Spine Study, Part II. 1-year outcomes of surgical and nonsurgical management of sciatica. *Spine (Phila Pa 1976)* 1996;21:1777–86.
13. Patrick DL, Deyo RA, Atlas SJ, *et al.* Assessing health-related quality of life in patients with sciatica. *Spine (Phila Pa 1976)* 1995;20:1899–908. discussion 909.
14. Weinstein JN, Lurie JD, Tosteson TD, *et al.* Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT) observational cohort. *JAMA* 2006;296:2451–9.
15. Weber H. Lumbar disc herniation. A controlled, prospective study with ten years of observation. *Spine (Phila Pa 1976)* 1983;8:131–40.
16. Buttermann GR. The effect of spinal steroid injections for degenerative disc disease. *Spine J* 2004;4:495–505.
17. Osterman H, Seitsalo S, Karppinen J, *et al.* Effectiveness of microdiscectomy for lumbar disc herniation: a randomized controlled trial with 2 years of follow-up. *Spine (Phila Pa 1976)* 2006;31:2409–14.
18. Peul WC, van den Hout WB, Brand R, *et al.* Prolonged conservative care versus early surgery in patients with sciatica caused by lumbar disc herniation: two year results of a randomised controlled trial. *BMJ* 2008;336:1355–8.
19. Peul WC, van Houwelingen HC, van den Hout WB, *et al.* Surgery versus prolonged conservative treatment for sciatica. *N Engl J Med* 2007;356:2245–56.
20. Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ* 1996;312:1215–8.
21. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med* 2000;342:1887–92.
22. McKee M, Britton A, Black N, *et al.* Methods in health services research. Interpreting the evidence: choosing between randomised and non-randomised studies. *BMJ* 1999;319:312–15.
23. Pocock SJ, Elbourne DR. Randomized trials or observational tribulations? *N Engl J Med* 2000;342:1907–9.
24. Delamarter RM. Microdiscectomy and microsurgical laminotomies. In: Frymoyer JW, ed. *The adult spine: principles and practice*. Philadelphia: Lippincott-Raven Publishers, 1997:1961–88.
25. Spengler DM. Lumbar discectomy. Results with limited disc excision and selective foraminotomy. *Spine (Phila Pa 1976)* 1982;7:604–7.
26. Roberts ST, Willick SE, Rho ME, *et al.* Efficacy of lumbosacral transforaminal epidural steroid injections: a systematic review. *PM R* 2009;1:657–68.
27. Simopoulos TT, Kraemer J, Nagda JV, *et al.* Response to pulsed and continuous radiofrequency lesioning of the dorsal root ganglion and segmental nerves in patients with chronic lumbar radicular pain. *Pain Physician* 2008;11:137–44.
28. Schochat T, Rehberg W, von Kempis J, *et al.* [The North American Spine Society Lumbar Spine Outcome Assessment Instrument: translation and psychometric analysis of the German version in rehabilitation patients with chronic back pain]. *Z Rheumatol* 2000;59:303–13.
29. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473–83.
30. Pose B, Sangha O, Peters A, *et al.* [Validation of the North American Spine Society Instrument for assessment of health status in patients with chronic backache]. *Z Orthop Ihre Grenzgeb* 1999;137:437–41.
31. Bullinger M. German translation and psychometric testing of the SF-36 Health Survey: preliminary results from the IQOLA Project. International Quality of Life Assessment. *Soc Sci Med* 1995;41:1359–66.
32. Schaaeren S, Bischoff-Ferrari HA, Knupp M, *et al.* A computer touch-screen version of the North American Spine Society outcome assessment instrument for the lumbar spine. *J Bone Joint Surg Br* 2005;87:201–4.
33. da Costa BR, Gahl B, Jüni P. Tools & techniques—statistics: propensity score techniques. *EuroIntervention* 2014;10:761–7.
34. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000;11:550–60.
35. Atlas SJ, Keller RB, Wu YA, *et al.* Long-term outcomes of surgical and nonsurgical management of sciatica secondary to a lumbar disc herniation: 10 year results from the maine lumbar spine study. *Spine* 2005;30:927–35.