Introduction – Literature Overview

Surgical treatment of a herniated lumbar disc, as described by Mixter and Barr (1) dates back to 1934. Within a decade, the technique became widespread, and though a variety of modified surgical techniques have been developed over the years, lumbar discectomy as originally described, remains the most universal technique.

A considerable body of literature reports on short-term outcome results following lumbar discectomy, which are generally satisfactory. Studies on long-term outcome report on varying success rates. The first long-term studies were encouraging: in 1949, Spurling (2) reported on excellent outcome in the majority of the patients after a follow-up period of 10 years. Guardjian (3) confirmed a success rate of 76% in 1961. Barr and Mixter (4) shared the early enthusiasm.

Naylor (5) published on excellent and good results in respectively 41% and 18%. He also reported on imperfect operative results: 17% of the patients had persistent back pain and 16% had persistent neurological symptoms (numbness, paresthesiae, cramps, sciatic pain). In a study by Salenus (6), 36% of the patients stated that the pain was the same as pre-operatively after a 6 year period, and 8% stated that the pain had become worse. Different types of surgery were however included in this study.

In a Belgian study by Van Loon (7), only 8% of the patients reported residual or aggravation of sciatic pain. Low back pain (LBP) however, persisted moderately in 34% and heavily in 30%. They concluded to an overall good result in 85% of the patients.

Frymoyer (8) stated that in 37% surgery was considered a failure. Weber (9) reported on satisfactory results in 64% at 1 to 2 years of follow-up.

Lewis (10) published a prospective study: in his series, 62% had complete relief of back pain and leg pain at a minimum of 5 years follow-up, and 28% had partial relief. The results presented by Dvorak (11) were more pessimistic: at 4 to 17 years follow-up, 70% still complained of LBP, of which 23% of constant heavy pain.

Waddell (12) found the outcome of different types of back surgery to be dependent upon the way of outcome measurement: patient’s self-assessment versus physician’s assessment versus ability to return to work. Abramovitz (13) reported on a prospective, multiple centres study with 840 discectomy patients. At 1-year follow-up, physician’s assessed the result to be good in 77%; the patient in 73%. Davis (14) confirmed good to excellent long term results in 89% after a mean period of 10 years.

The results of a self-report questionnaire, published by Loupasis (15) were again more nuanced: late results were in 64% satisfactory and unsatisfactory in 34%. Thirty-eight percent still complained of significant back or leg pain. Yorimitsu (16) stated that residual LBP was found in 74.6% of the patients. Mariconda (17) published on patient-reported disability in terms of Oswestry score at a mean of 25 years after operation, which was minimal or absent in 75.6%. In a recent study, Bakhash (18) stated that surgery provided immediate pain relief in 79.4%, but that the long-term outcome was not satisfactory. Pain recurred in 12.82% of cases after 1 year, and in 35.89% during the first 5 years.

Since the late nineties, the need to measure the effectiveness of surgery in relieving symptoms and improving quality of life in a standardized way has been stressed. Only a few of the aforementioned studies have however used validated and/or back-specific questionnaires (15, 17, 19). The current
study not only does so, but also addresses chronic LBP from a biopsychosocial perspective.

Patients and methods

SUBJECTS

All patients (n = 282) who underwent a lumbar discectomy at the Ghent University Hospital, department of Neurosurgery, between September 1st 1996 and August 31st 2000 were contacted by mail (n = 278; 4 patients had migrated). Age at the time of surgery ranged from 19 to 83 years (mean 44.1 y). The patient sample was divided into 3 age groups (19-40 y, 41-60 y, and 61-83 y) to examine the influence of age on outcome. Patient characteristics are listed in table 1. Time since surgery at the moment of evaluation ranged from 42 months to 83 months (mean 63 months). To examine the influence of time since surgery on outcome, the patient group was divided into 5 groups, according to the year patients were operated in (1996 to 2000).

QUESTIONNAIRES

The 7-page questionnaire included a general information form (medical history), the Macnab Classification System (20), the Quebec Back Pain Disability Scale (21) (QBPDS) and the West Haven Yale Multidimensional Pain Inventory (22), Part I (MPI-I). The Macnab outcome measure is a frequently used single rating scale with 4 response options: excellent, good, fair and poor (appendix 1). QBPDS (21) is a 20-item self-administered instrument. The psychometric properties of the Dutch version are similar to those of the English one (i.e. high levels of reliability and validity) (23). It is designed to measure the level of functional disability in persons with back pain and focuses exclusively on 6 domains of physical function: bed rest, sitting-standing, walking, moving, bending and moving heavy objects. Twenty questions are scored on a 6-point scale: from 0 (“not difficult at all”) to 5 (“unable to do so”). The total score ranges from 0 to 100, with lower scores indicating better function. The MPI (22) measures several of the pain-relevant aspects on the psychosocial dimension and on pain behaviour. Part I measures pain-relevant psychosocial-aspects, and is made up of 5 scales: pain severity (3 items), interference with the daily life due to pain (11 items), perceived life control (4 items), affective distress (3 items) and social support (3 items). Items are scored on a 7-point scale (0 to 6). The total scale score is the sum of the item scores. Pain severity scores range thus from 0 to 18 (higher scores indicating more pain); interference scores from 0 to 66 (higher scores corresponding with more interference); perceived life control scores from 0 to 24 (higher scores indicating more control); affective distress scores from 0 to 18 (higher scores corresponding with more distress); social support scores from 0 to 18 (higher scores indicating more support). The Dutch language version of the MPI has been shown to produce reliable, valid information for diagnostic purposes and for therapy-outcome studies (24).

Current treatment and readiness to participate in an experimental exercise program were questioned. Patients without complaints (Macnab score 1), were allowed to skip the QBPDS and the MPI-I. Patients who had not returned the questionnaire within a month were sent a second (identical) questionnaire. If not answered, patients were contacted by telephone.

Responders were divided into 4 groups according to lumbar surgery history.

Table 1

Patient characteristics: age at the time of surgery and level of operation

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>%</th>
<th>mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>161</td>
<td>57.9</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>117</td>
<td>42.1</td>
<td></td>
</tr>
<tr>
<td>Age at the time of surgery (y)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19-40</td>
<td>112</td>
<td>40.3</td>
<td></td>
</tr>
<tr>
<td>41-60</td>
<td>134</td>
<td>48.2</td>
<td></td>
</tr>
<tr>
<td>61-83</td>
<td>32</td>
<td>11.5</td>
<td></td>
</tr>
<tr>
<td>Level of operation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L3-L4</td>
<td>15</td>
<td>5.0</td>
<td></td>
</tr>
<tr>
<td>L4-L5</td>
<td>128</td>
<td>43.2</td>
<td></td>
</tr>
<tr>
<td>L5-S1</td>
<td>142</td>
<td>47.8</td>
<td></td>
</tr>
<tr>
<td>2 levels</td>
<td>8</td>
<td>2.7</td>
<td></td>
</tr>
<tr>
<td>level unknown</td>
<td>4</td>
<td>1.3</td>
<td></td>
</tr>
</tbody>
</table>
Statistical Analysis

Parametric tests were performed since there was a Gaussian distribution of the data. ANOVA and Independent t-tests were used to examine differences between groups. Statistical significance was set at p = 0.05.

The relationship between the non-continuous Macnab scale and the continuous QBPDS was examined by means of a Univariate General Linear Model with fixed factor ‘Macnab’ and dependent factor ‘QBPDS’. The Partial Eta Squared value was determined. All responders were included to this aim. The same was done for the dependent factors ‘pain severity’, ‘interference’, ‘life control’, ‘affective distress’ and ‘social support’.

Correlations between the QBPDS and the sub-scales of the MPI-I were investigated using Pearson’s Correlations Coefficient; all responders were included to this aim. Significance for correlations was set at the 0.01 level (2-tailed); high Pearson correlates was defined as 0.65 < r < 0.85; moderate Pearson correlates as 0.4 < r < 0.65; weak Pearson correlates as 0.2 < r < 0.4.

All data were processed using the Statistical Package for Social Science for Windows (SPSS 16.0).

Results

Response Rate

The response rate was 85.61% (135 men (57.7%), 99 women (42.3%)). One hundred and fifty six patients returned the questionnaire within 30 days. Sixty-three patients responded the second mailing. Phone calls were made to 58 patients, of which another 19 answered the questionnaire. Twelve patients could not be contacted (1 had died).

Responders

Of the 234 responders, 41 had undergone prior lumbar surgery, and 13 patients had a spinal fixation or spinal prosthesis later on. These patients were excluded from further analysis.

Twenty-five patients had a revision or second lumbar discectomy in the follow-up period. Responder subgroups are listed in table 2. Patient characteristics of the included patients are listed in table 3.

Macnab score – QBPDS - MPI-I

Results of the self-rated Macnab score are reported in table 4. Excellent (score 1) and Good (score 2) results were reported in 58.99%. Mean Scores of the QBPDS and the MPI-I are listed in table 5.

Comparison of outcome according to the Macnab score

Significant differences were found for the QBPDS (p < 0.001), pain severity (p < 0.001), interference (p < 0.001), control (p < 0.001), and distress (p = 0.009). The support score (p = 0.369) was not significantly different between the groups.

Patients with Macnab score 2 had significant better scores than patients with Macnab score 3 on the QBPDS (p < 0.001), pain severity (p < 0.001), interference (p = 0.002), distress (p = 0.010) and control (p < 0.001) scales. Similar differences were found for patients with Macnab score 2 versus Macnab score 4 (QBPDS p < 0.001; pain severity p < 0.001; interference p = 0.002; distress p = 0.006; control p < 0.001), and for patients with Macnab score 3 versus Macnab score 4 (QBPDS p < 0.001; pain severity p < 0.001; interference p = 0.04; distress p = 0.041; control p = 0.016).

Mean scores of the QBPDS, pain severity, interference, support, distress and control subscales are given in table 6.

Influence of age on outcome

Between the age groups 19-40 y, 41-60 y and 61-83 y, a significant difference in outcome score was
found for the QBPDS score only (p = 0.028). Mean QBPDS score was lower in patients aged between 19-40 y (mean score 24.42) compared with patients aged between 41-60 y (mean score 33.43) (p = 0.008). No significant difference in QBPDS score was found between the 41-60 y group compared with the 61-83 y group (p = 0.664), and between the 19-40 y group compared with the 61-83 y group (mean score 30.91; (p = 0.270). The mean pain severity (p = 0.293), interference (p = 0.755), life control (p = 0.248), affective distress (p = 0.465) and social support scores (p = 0.774) were not significantly different between the age groups.

### INFLUENCE OF TIME SINCE SURGERY

Between the 5 patient groups, no significant difference was found for any of the outcome scores (QBPDS 0.461; pain severity 0.515; interference 0.521; control 0.338; affective distress 0.097; support 0.699).
LONG TERM OUTCOME IN LUMBAR DISCECTOMY

Table 6
Mean scores and standard deviations and of the QBPDS, pain severity, interference, support, distress and control subscales of the MPI-I for the different Macnab scores

<table>
<thead>
<tr>
<th>Macnab</th>
<th>QBPDS</th>
<th>Pain severity</th>
<th>Interference</th>
<th>Control</th>
<th>Affective Distress</th>
<th>Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>(&lt; 0.001)</td>
<td>(&lt; 0.001)</td>
<td>(&lt; 0.001)</td>
<td>(&lt; 0.001)</td>
<td>0.009</td>
<td>0.369</td>
</tr>
<tr>
<td>3</td>
<td>40.74 ± 17.14</td>
<td>8.1 ± 3.23</td>
<td>29.97 ± 12.58</td>
<td>15.52 ± 4.31</td>
<td>8.54 ± 2.76</td>
<td>12.76 ± 4.67</td>
</tr>
<tr>
<td>4</td>
<td>59.07 ± 13.69</td>
<td>13.00 ± 3.06</td>
<td>42.20 ± 4.66</td>
<td>13.90 ± 4.20</td>
<td>9.15 ± 2.83</td>
<td>12.83 ± 4.50</td>
</tr>
<tr>
<td>Partial Eta²</td>
<td>.59</td>
<td>.48</td>
<td>.19</td>
<td>.15</td>
<td>.08</td>
<td>.02</td>
</tr>
</tbody>
</table>

Table 7
Pearson correlations between outcome scores

<table>
<thead>
<tr>
<th>Outcome scores</th>
<th>r</th>
<th>p</th>
<th>r²</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Pearson correlate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QBPDS – Interference</td>
<td>0.82</td>
<td>&lt; 0.001</td>
<td>0.67</td>
</tr>
<tr>
<td>QBPDS – Pain severity</td>
<td>0.81</td>
<td></td>
<td>0.66</td>
</tr>
<tr>
<td>Pain severity – Interference</td>
<td>0.76</td>
<td></td>
<td>0.58</td>
</tr>
<tr>
<td>Moderate Pearson correlate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain severity – Affective distress</td>
<td>0.44</td>
<td>&lt; 0.001</td>
<td>0.19</td>
</tr>
<tr>
<td>Pain severity – Control</td>
<td>-0.46</td>
<td></td>
<td>0.21</td>
</tr>
<tr>
<td>Control – QBPDS</td>
<td>-0.43</td>
<td></td>
<td>0.18</td>
</tr>
<tr>
<td>Weak Pearson correlate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interference – Affective distress</td>
<td>0.39</td>
<td>&lt; 0.001</td>
<td>0.15</td>
</tr>
<tr>
<td>Affective distress – Control</td>
<td>0.36</td>
<td></td>
<td>0.13</td>
</tr>
<tr>
<td>Affective distress – QBPDS</td>
<td>0.10</td>
<td></td>
<td>0.01</td>
</tr>
<tr>
<td>Interference – Control</td>
<td>-0.32</td>
<td>≥ 0.001</td>
<td>0.10</td>
</tr>
</tbody>
</table>

RELATIONSHIP BETWEEN THE OUTCOME MEASURES

The Macnab scale related most to the QBPDS (Partial Eta² = 0.59) and to the pain severity subscale (Partial Eta² = 0.48). Partial Eta² values for the Macnab as fixed factor are presented in table 6. Pearson correlates between the continuous outcome scales are presented in table 7.

COMPARISON OF OUTCOME IN SINGLE DISCECTOMY VERSUS MULTIPLE DISCECTOMY

The QBPDS score (p = 0.002) was significantly higher in patients with more than one lumbar discectomy (mean score 41.98) compared with patients with a single lumbar discectomy (mean score 27.24). Pain severity score (p = 0.01) was significantly higher in multiple discectomy (8.89) compared with single lumbar discectomy (6.27). The mean scores of the Interference subscale (p = 0.079), of the Support subscale (p = 0.669), of the Distress subscale (p = 0.566) and of the Control subscale (p = 0.286) were not significantly different. Mean scores and standard deviations for single and multiple discectomy are listed in table 5.

CURRENT TREATMENT

Eighty-seven (48.3%) took pain medication on a regular basis. Forty-nine persons (27.2%) had taken pain killers on the day of response. Reported pain medication is listed in table 8. Thirty-three patients (18.33%) receive current treatment for back pain. Physiotherapy (41.12 valid %) and osteopathy (11.83 valid %) were most frequently reported.

EXPERIMENTAL EXERCISE THERAPY

113 patients (68.9%) were ready to participate in an experimental exercise study.

Discussion

RESPONSE RATE/RESPONDERS/QUESTIONNAIRES

A difference in completion of the different outcome measures was seen. A low compliance was seen for the interference subscale (41% incomplete), whereas the Macnab and the QBPDS were generally completely filled in. The presence of unanswered
items has been reported to be inherent to self-administered, mailed questionnaires (25).

POST-OPERATIVE OUTCOME (MACNAB SCORE, QBPD S, MPI-I)

In the current study, 12.36% of the patients were still completely relieved of their back and/or sciatic pain at long term follow-up. Eighty-three of the 156 painful patients (53.21%) had Macnab score 2: occasional back or leg pain (biological factor) of sufficient severity to interfere with the patient’s ability to do his normal work (social factor) or his capacity to enjoy himself in his leisure hours (psychosocial factor). The co-existence of psychosocial problems is clearly documented by the scores of the MPI-I subscales.

The results are similar to those reported by Loupasis (15), who found satisfactory results in 64% using a self-administered back specific questionnaire.

COMPARISON OF OUTCOME ACCORDING TO THE MACNAB SCORE

Patients with a higher Macnab score experience more psychosocial problems: more interference in daily life, less control and more distress. Getting more insight in the complexity and the extent of the pain problem, might help patients to understand and manage their back pain. The outcome results in a biopsychosocial approach acknowledge that not only the LBP has to be managed, but also the psychological consequences in terms of distress and control.

INFLUENCE OF AGE ON OUTCOME

As all patients operated on over a 4-year period were included, the age of the patients ranged from 19 y to 83 y. Between the 3 age groups, a significant difference in outcome was found for the QBPD S score only. Patients under 41 y reported fewer problems with back specific activities than patients aged 41-60 y. The relevance of this difference is unclear. A relation with better tissue recovery or better physical fitness is unlikely, since no significant difference was found between patients aged 19-40 y and 61-83 y.

INFLUENCE OF TIME SINCE SURGERY

Success rates of lumbar discectomy have been reported to diminish over time: the longer the follow-up period, the higher the pain and disability rates (18). In the current study, all patients were evaluated at the same moment. As a consequence, time since surgery varied between 43 months and 83 months. Though some bias due to the difference in time of the evaluation was therefore expected, time since surgery seemed to cause no substantial bias”.

RELATIONSHIP BETWEEN THE OUTCOME MEASURES

Fifty-nine percent of variance in the QBPD S score and 48% of variance in the pain severity score can be account for by the Macnab score. There seems little relationship between the Macnab scale and the other subscale of the MPI-I. The short Macnab outcome scale, which is often used in evaluation of surgical procedures, seems a fair screening instrument for back related pain and function. Since multiple dimensions are combined in one score, it is impossible to evaluate pain and function separately. As there are only 4 response options, subtle shifts in improvement or aggravation are not detected.

The QBPD S correlates strongly with the pain severity subscale and the interference subscales of
the MPI-I. In the current study, QBPDS was preferred over other established back pain questionnaires (such as the Roland-Morris questionnaire or the Oswestry questionnaire) because it examines several activities which are not questioned in other back pain questionnaires, such as carrying, pulling or throwing an object, reaching or running. QBPDS includes no questions relating to pain intensity, personal care and social role activities. The strong correlation with pain severity and general interference of pain in daily life is therefore interesting. The QBPDS seems however less appropriate and therefore less indicated for examination of some other psychosocial factors such as subjective life control or affective distress. It is likely that next to back complaints, these factors are influenced by a person’s professional life, relation or familial stress.

PAIN MEDICATION – OTHER TREATMENT MODALITIES

A surprisingly high percentage of the painful patients take Non-Steroidal Anti-Inflammatory Drugs (NSAID) on a regular basis. The use of non-narcotic pain killers, in contrast, is low. Even more surprising, tricyclic anti-depressants and selective serotonin re-uptake inhibitors, first-line medications for chronic neuropathic pain (26), were not taken by any of the chronic pain patients.

NSAID have been proven effective in acute LBP (27). However, the benefit of NSAID for treating chronic LBP is unclear, and there is no evidence supporting the use of NSAID in sciatic pain (28). The rational for NSAID treatment in chronic LBP is probably based on the analgesic properties and the assumption that an inflammatory component is present in the chronic back pain mechanism. The risk of development of ulcers in regular users of NSAID is however well-known. Choosing pharmacologic therapy in function of risks and evidence-based benefits remains a challenge.

Pain medication, which can be regarded as therapy for the biological problems, is far more common than any other treatment modality. Sixty nine percent of the patients was however willing to participate in an experimental exercise program. This might reflect patient’s hope that exercise therapy will affect residual complaints, the wish to become more active under medical supervision, tiredness of medication intake or even – indirectly – the patient’s wish for a more psychosocial approach. Several patients sent an accompanying letter, expressing their gratitude for asking about residual problems. This might reflect a patient’s feeling that not enough attention is being paid to the post-operative follow-up.

COMPARISON OF OUTCOME IN SINGLE DISCECTOMY VERSUS MULTIPLE DISCECTOMY

It seems little surprising that patients who underwent more than one operation had a higher pain severity score and more problems with back-related functioning. Surgeons therefore should always carefully select operative indications, in special second operation or revision. Information about what to expect after the operation – in particular the presence of some back discomfort – is part of the surgeon’s responsibility.

LIMITATIONS

This study reflects the operative results of a single department in a university hospital. Though the operative technique itself is fairly universal, it is possible that the patient group seeking help in a university hospital is somewhat different in nature than a patient group in a peripheral hospital.

As this is a retrospective study, the results are in part based on the memory and the subjective view of the patients, and thus prone to some recall bias.

Some limitations are the consequence of the questionnaire used. In the presumption that pain free patients would not be motivated to complete a 7 page questionnaire, those patients were allowed to skip the QBPDS and the MPI-I. As a consequence, there are no mean QBPDS and MPI-I scores available for pain free discectomy patients. Normative data are to be established.

Conclusion

Long term after lumbar discectomy, outcome is fairly good. Patients with higher levels of pain experience more interference in daily life, have less control and are more distressed. Evaluation of outcome in a biopsychosocial perspective might help to gain insight in the complexity of the pain problem and managing it.

Acknowledgements

The authors wish to thank the Fund for Scientific Research-Flanders, Dr. Peersman for statistical advice, and all the patients who participated in the study.

REFERENCES


Katie Bouche,
Department of Physical Medicine and Rehabilitation,
Faculty of Medicine and Health Sciences,
Ghent University Hospital,
De Pintelaan 185,
9000 Ghent (Belgium).
E-mail: katie.bouche@uzgent.be
Appendix

Macnab scale:

1: Excellent: No pain. No restriction of activity.
2: Good: Occasional back or leg pain of sufficient severity to interfere with the patient’s ability to do his normal work or his capacity to enjoy himself in his leisure hours.
3: Fair: Improved functional capacity, but handicapped by intermittent pain of severity to curtail or modify work or leisure activities.
4: Poor: No improvement or insufficient improvement to enable increase in activities. Further operative intervention required.