We have developed criteria to determine the appropriate indications for lumbar laminectomy, using the standard procedure developed at the RAND corporation and the University of California at Los Angeles (RAND-UCLA). A panel of five surgeons and four physicians individually assessed 1000 hypothetical cases of sciatica, back pain only, symptoms of spinal stenosis, spondylolisthesis, miscellaneous indications or the need for repeat laminectomy. For the first round each member of the panel used a scale ranging from 1 (extremely inappropriate) to 9 (extremely appropriate). After discussion and condensation of the results into three categories laminectomy was considered appropriate in 11% of the 1000 theoretical scenarios, equivocal in 26% and inappropriate in 63%. There was some variation between the six categories of malalignment, but full agreement in 64% of the hypothetical cases.

We applied these criteria retrospectively to the records of 196 patients who had had surgical treatment for herniated discs in one Swiss university hospital. We found that 48% of the operations were for appropriate indications, 29% for equivocal reasons and that 23% were inappropriate.

The RAND-UCLA method is a feasible, useful and coherent approach to the study of the indications for laminectomy and related procedures, providing a number of important insights. Our conclusions now require validation by carefully designed prospective clinical trials, such as those which are used for new medical techniques.

Disorders of the lumbar spine are common and represent a large proportion of health care: in the UK about 13 000 operations are carried out annually for low-back disorders and back pain causes loss of about 150 million days of work each year. In the USA there are over 250 000 operations for back pain. Despite many varied treatments, disability due to this cause appears to be increasing.

The quality of care is a major concern and a challenge for clinicians, researchers and policy makers. ‘Evidence-based medicine’ can only be useful when there is proper evidence, and many treatments used in daily practice are not supported by careful studies, let alone by rigorous controlled clinical trials. This applies to the indications for lumbar laminectomy, which are usually based, at least partly, on expert opinion. A procedure is considered to be appropriate when the medical benefit in terms of quality of life or life-span, reduction in pain and improved function are thought to exceed by a sufficiently wide margin, the medical risks of mortality, morbidity, and anxiety caused by the procedure.

Surgery is only one option in the broad spectrum of treatments for disorders of the lower back. The technical competence of the operation is difficult to assess, but the appropriateness of the indication can be evaluated. Most failures are attributed to the poor selection of patients, and this helps to explain the differences in reported outcome ranging as it does from 58% to 93% of ‘good’ results. There is no widely accepted consensus on the detailed indications for surgery for low-back pain, although there are many papers on patient selection and on criteria for prognostic factors, some of which provide decision algorithms. The scientific basis for such recommendations is weak. One of the rare randomised, prospective trials evaluating laminectomy against conservative treatment for lumbar disc hernia, designed by Weber, showed a significant postoperative benefit at one...
The assessment of appropriate indications for laminectomy

We convened a multispecialty expert panel in October 1995 in Switzerland with the aim of determining the appropriate indications for lumbar laminectomy and related procedures. The members were chosen by the principal specialty societies involved in the care of such patients; it included two orthopaedic surgeons and three neurosurgeons with one general practitioner, one neurologist, one rheumatologist and one general internal medicine physician. There were therefore five ‘doers’ of laminectomy and four who were ‘referrers’ for surgery.

We used the standard procedure of the RAND-UCLA method of assessing the appropriate indications. Each member of the panel was given a summarised review of the literature of 36 pages, covering 107 relevant publications on the effectiveness, complications and outcome of spinal surgery for disc hernia. These articles had been identified by a Medline search from 1966 to August 1994, using the medical subject headings lumbar vertebrae, lumbosacral region and intervertebral disc cross-referenced by terms which included surgery, laminectomy, discectomy and outcome. The panel members were asked to study the review and use the information together with their own clinical judgement to assess the appropriateness of laminectomy.

A list of all potential surgical indications for lumbar laminectomy was then prepared and discussed individually with the panel members in order to refine it and eliminate misinterpretation. The six major groups of indications were agreed to be sciatica, back pain only, symptoms of spinal stenosis, spondylolisthesis, miscellaneous and repeat laminectomy.

The factors considered for each of the clinically specific scenarios were neurological abnormalities, radiological abnormalities, previous conservative treatment, disability, duration of symptoms, nature of symptoms, pending insurance litigation, and the results of other imaging studies. The definitions of all the terms used were agreed and are available on request from the authors. Our intention was to cover virtually all imaginable clinical patterns of degenerative diseases for which lumbar laminectomy might be considered.

The panel then evaluated 1000 sets of clinical indications. Most included sciatica (n = 720), spondylolisthesis (n = 156), or spinal stenosis (n = 48). Three examples, concerning acute, subacute and chronic sciatica, are shown in Table 1 to indicate how different levels of severity were combined to provide descriptions of clinically specific situations.

Each of the specific situations was then rated separately by each panel member on a nine-point scale as to the appropriateness of laminectomy (1, extremely inappropriate; 5, equivocal or uncertain; 9, extremely appropriate). At a second stage the panel met to discuss the results with particular attention to disagreements on the first round. This multidisciplinary discussion was chaired by a trained moderator and the panel members then reassigned their ratings for the appropriateness of each indication.

These second-round ratings were then used to determine the consensus on indications for laminectomy. To facilitate comparisons, we condensed the nine-point scale into three levels: inappropriate (1 to 3), equivocal (4 to 6) and appropriate (7 to 9) recording the median value for the whole panel together with a measure of agreement between members. Agreement was present when, after discarding one extreme high and one extreme low rating, the remaining seven ratings all fell within any single three-point region. Disagreement was defined as occurring when at least two panellists rated an indication in the range of 1 to 3 and at least two others rated it in the range of 7 to 9. Intermediate situations, where there was

Table I. Three examples of hypothetical cases of disc hernia considered by the panel, related to the variables used in the evaluation matrix

<table>
<thead>
<tr>
<th>Variable</th>
<th>Clinical description</th>
<th>Subacute sciatica</th>
<th>Acute sciatica</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Duration</td>
<td>Less than 6 months</td>
<td>Less than 6 weeks</td>
<td>More than 6 months</td>
</tr>
<tr>
<td>2. Neurology</td>
<td>With non-progressive major muscular weakness</td>
<td>Minor neurological abnormalities and a positive root-tension sign</td>
<td>With (stable) minor neurological abnormalities, persistent asymmetry of ankle reflexes</td>
</tr>
<tr>
<td>3. Radiology</td>
<td>Herniated disc on imaging</td>
<td>Herniated disc on imaging</td>
<td>A herniated disc with a free fragment is seen on imaging</td>
</tr>
<tr>
<td>4. Previous conservative treatment</td>
<td>Already treated with two non-operative treatment regimes</td>
<td>No non-operative treatments</td>
<td>Having had one non-operative treatment regime</td>
</tr>
<tr>
<td>5. Disability</td>
<td>Is moderately disabled</td>
<td>Is moderately disabled</td>
<td>Is moderately disabled</td>
</tr>
<tr>
<td>6. Insurance claim or litigation</td>
<td>No insurance claim pending for this problem</td>
<td>No insurance claim pending for this problem</td>
<td>An insurance claim is pending for this problem</td>
</tr>
</tbody>
</table>
neither agreement nor disagreement, were qualified as uncertain.

We used chi-squared tests for comparison of distributions and paired \( t \)-tests for between-group differences in the average median rating for identical indications. For scenarios which included a primary diagnosis of sciatica, we used multiple logistic regression techniques to determine the relative and absolute importance of each of the variables. This multifactorial approach was needed to formulate the hypothetical individual indications for each primary diagnostic category considered by the panel.

The criteria which had been developed were then retrospectively applied to the records of 196 consecutive patients who had been treated by low-back surgery in one university hospital.

### Results

Of the 1000 theoretical clinical scenarios considered by the panel only 106 (11\%) were rated as appropriate for laminectomy (Table II), while 26\% had equivocal indications and 63\% were inappropriate. The panel agreed on one of the three categories in 64\% of the hypothetical cases, and the results for each clinical category are shown in Table II.

For all of the 36 clinical scenarios which described low-back pain only, the panel had 100\% agreement that laminectomy was inappropriate. The four ‘miscellaneous cases’ also produced 100\% agreement: laminectomy was appropriate for the two examples of cauda equina syndrome and the two showing progressive weakness, but only when there was known to be a herniated disc.

### Value of multidisciplinary discussion.

Comparison between the ‘first-round’ and the ‘second-round’ ratings showed a slight shift of the median score away from appropriateness, from 13\% to 11\% of the hypothetical cases.

The impact of the interaction of panel members with other specialists was confirmed by an improvement from ‘first-round’ agreement in 39\% of the 1000 scenarios to 64\% after discussion of divergent ratings (Fig. 1), with ultimate unresolved disagreement between the three values in only 6\% of the hypothetical cases.

### Differences within the panel.

Within the panel there were some divergent views between the five surgeons, the ‘doers’ and the four physician ‘referrers’. The ‘doers’ found 13\% in the appropriate range for discectomy (7 to 9), as against 7\% so rated by the referring physicians \((p < 0.001)\) (Fig. 2).

---

Table II. The final opinion of the panel on the indications for laminectomy (appropriate, equivocal or inappropriate) for the six major clinical categories among clinical hypothetical clinical cases by number and percentage

<table>
<thead>
<tr>
<th>Indications</th>
<th>Number</th>
<th>Appropriate</th>
<th>Equivocal</th>
<th>Inappropriate</th>
<th>Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sciatica</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute</td>
<td>180</td>
<td>22 (12)</td>
<td>51 (28)</td>
<td>107 (59)</td>
<td>114 (63)</td>
</tr>
<tr>
<td>Subacute</td>
<td>270</td>
<td>39 (14)</td>
<td>64 (24)</td>
<td>167 (62)</td>
<td>170 (63)</td>
</tr>
<tr>
<td>Chronic</td>
<td>270</td>
<td>16 (6)</td>
<td>94 (35)</td>
<td>160 (59)</td>
<td>177 (66)</td>
</tr>
<tr>
<td>Back pain only</td>
<td>36</td>
<td>-</td>
<td>-</td>
<td>36 (100)</td>
<td>36 (100)</td>
</tr>
<tr>
<td>Symptoms of spinal stenosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Central</td>
<td>12</td>
<td>3 (25)</td>
<td>1 (8)</td>
<td>8 (67)</td>
<td>10 (83)</td>
</tr>
<tr>
<td>Lateral</td>
<td>36</td>
<td>5 (14)</td>
<td>7 (19)</td>
<td>24 (67)</td>
<td>29 (81)</td>
</tr>
<tr>
<td>Spondylolisthesis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Without fusion</td>
<td>78</td>
<td>4 (5)</td>
<td>14 (18)</td>
<td>60 (77)</td>
<td>52 (67)</td>
</tr>
<tr>
<td>With fusion</td>
<td>78</td>
<td>12 (15)</td>
<td>22 (28)</td>
<td>44 (56)</td>
<td>29 (37)</td>
</tr>
<tr>
<td>Miscellaneous*</td>
<td>4</td>
<td>2 (50)</td>
<td>2 (50)</td>
<td>4 (100)</td>
<td></td>
</tr>
<tr>
<td>Repeat laminectomy</td>
<td>36</td>
<td>3 (8)</td>
<td>8 (22)</td>
<td>25 (69)</td>
<td>21 (58)</td>
</tr>
<tr>
<td>Total</td>
<td>1000</td>
<td>106 (11)</td>
<td>261 (26)</td>
<td>633 (63)</td>
<td>642 (64)</td>
</tr>
</tbody>
</table>

* Two cases of cauda equina syndrome and progressive motor weakness, one showing good imaging confirmation and one with no imaging evidence of a disc hernia (see text)
Evaluation of the process. The panel members were asked to complete a numerical questionnaire (1, lowest; 5, highest) at the sessions. The mean score for both the validity and the utility of the method was 3.9, the discussion had a mean of 4.8, and the panel members’ satisfaction was rated at 4.7.

Logistic regression study. For the two primary diagnostic categories of sciatica and spondylolisthesis, we estimated the probability that laminectomy was considered inappropriate as against equivocal or appropriate in relation to the dichotomous logistical functions of the six variables within each primary category. We found that all six variables included in the models were necessary in the final model for sciatica (goodness-of-fit p = 0.7). This confirmed the value of our multifactorial approach to defining the appropriate nature of indications.

Retrospective testing. We applied the newly developed criteria for appropriate surgery in a retrospective manner to the records of 196 patients who had had surgery for lumbar disc herniation at a University Hospital in Switzerland. The indications for surgical treatment were appropriate in 48%, equivocal in 29% and inappropriate in 23% of these patients.

Discussion

In the era of ‘evidence-based medicine’, we must ask ‘What do we do when the evidence is insufficient?’ It is untenable simply to stop providing a certain treatment, but surgeons, patients and third-party payers do need standards by which they can evaluate the appropriateness of care.

To help towards defining decisions for treatment for back surgery, we used the RAND-UCLA method,10,31,32 which is gaining wide acceptance. This uses clinical, radiological and functional factors in a theoretical model which covers all possible clinical scenarios. We found a low rate of appropriate indications (11%) for laminectomy within all potential clinical situations in which the operation may be considered. This rather restrictive attitude to surgery reflects the general trend of recent papers;33,34 the only indications should be clear-cut clinical signs and symptoms which are in agreement with neuroradiological findings.

The absence of a trial of conservative treatment was considered to be a contraindication to operation, except for patients with signs of compression of the cauda equina or a rapidly progressive paresis due to a disc hernia confirmed by imaging. Persistent disability after several non-operative treatments, with major weakness, a positive nerve-root tension sign and confirmatory imaging, was an appropriate indication for surgery. Between the few unequivocal indications in either direction, there is a large grey zone with unknown or uncertain benefit-to-risk ratios for operation. In such cases, the criteria established by the panel can help therapeutic decisions by providing unbiased and detailed estimates of appropriateness. The scenarios which left the panel undecided are especially appropriate for carefully designed clinical trials in the future. The final agreement rate of 64% was higher than those for previous reports of the RAND-UCLA method,31 but there was a highly significant statistical difference of opinion between surgeon ‘doers’ and ‘referring physicians’ (p < 0.001). This probably reflects the fact that different specialists tend to favour their own approach to treatment, but could also be interpreted as showing a biased selection of patients by different specialists. The failures of back surgery usually return to non-operative specialists, while patients with excellent recoveries are not seen again. This emphasises the importance of a multidisciplinary approach to the problem, to encompass the largest possible potential patient population.

The value of the stages of the RAND process is shown by the improved intrapanel agreement after the multidisciplinary discussion (Fig. 2). When each panel member evaluated the appropriateness of laminectomy independently, without the feedback of other experts, there was agreement in under 40% of hypothetical cases. Exchange of opinions between experts from different specialties, with discussion of discordant ratings, increased agreement to almost two-thirds. Interdisciplinary discussion is a key element of the RAND method, as is the non-directive involvement of a trained moderator and non-dominant participation by the panel members. In our panel, these elements combined to produce very low disagreement (6%) for the hypothetical cases. This result is all the more
impressive because the method does not force consensus. The panel members give their final opinion privately and anonymously, and consensus is achieved naturally by the clinical specificity of the method and by the development and use of clear definitions.

The logistic regression models for sciatica and spondylo-
olisthesis confirmed the need to consider the whole constella-
tion of factors, and not emphasise any single variable in any
determination of the appropriateness of laminectomy. These
statistical findings also helped to validate the process and
show that the models made sense. The panel members
confirmed the overall utility and validity of the criteria, and
were impressed by the methodology. The true value of the
method and results, however, can be proven only by their
prospective application in future controlled trials.

Our retrospective application of the criteria to 196
patients who had had surgery for herniations of the lumbar
disc indicated that only 48% had fully appropriate indica-
tions and that in 29% the choice was equivocal. The 23% of
patients considered to have had inappropriate operations
could be evidence of an unduly liberal use of laminectomy,
but our results contrast with those of a previously published
study using criteria developed in 1987 in the USA. When
we applied these 1987 criteria to the same 196 cases, we
found that surgical treatment was deemed appropriate in
only 26% of the cases, equivocal in 33% and inappropriate
in 41%. This considerable difference proves the need for
periodic and continuing updating of the criteria.

Before our criteria for appropriate indications can be
widely propagated, there are several questions. The most
important is the need to test their validity in a prospective
clinical trial of the type used for any new technology in
medicine. Such a validation requires the design and per-
formance of rigorous studies to answer the vital questions:
‘Does the use of this decision-making tool lead to improved
health?’ or ‘does it lead to reduced cost without jeopardis-
ing health-care quality and outcome?’ It will not be easy to
provide clear evidence in an area of medicine in which
rigorous trials are rare, but it is necessary for such a break
with tradition.

A second important requirement is to examine the extent
of the need for local adaptations of the criteria. The Insti-
tute of Medicine in the USA recently suggested that such
local adaptations are often necessary to create guidelines
which are more relevant and more acceptable. They must,
however, be made with equal scientific rigour to ensure that
they have minimal or no negative effects on the well-being
of patients.

A third task is to find practical ways of making these
criteria widely available in a form that is easily accessible
and understood. The clinical specificity of the criteria which
we developed gives them a clear advantage over vague and
general guidelines, but also makes their practical imple-
mentation more difficult. It is not reasonable to expect
surgeons and physicians to wade through 1000 clinical
scenarios to find the one that most closely resembles a
specific case. It seems possible to integrate our criteria into
a user-friendly and easily accessible computer application
which could be available on the World Wide Web. This
could provide direct access to panel ratings. Other approa-
ches are possible, such as one based on a logistic predictive
model, but need further careful examination and validation.

At present there is no clear-cut published evidence of the
benefit-to-risk ratio of laminectomy for hernia of the lum-
bar disc and spinal stenosis. We found that the RAND-
UCLA method was a feasible and useful approach to
developing criteria for the appropriate indications.

The method has inherent internal validity, and the coher-
ence of the criteria was confirmed by a multivariate logistic
model, but it remains an entirely theoretical approach to
selection for back surgery. It needs prospective validation
in clinical settings, and we recommend the use of these
criteria in trials designed to assess their validity and
utility.

Financial support for this study was received from the Performance Fund
of the Cantonal Public Health Department, Vaud, Switzerland by grant no.
11 (1995). Value Health Sciences provided the literature review which we
used in the panel process.

The authors are particularly grateful to the panel members without
whom the project would not have been possible. They were Professor
Arnaldo Benini, Professor Walter Dick, Dr Johann Gaitzsch, Dr André
Kaelin, Dr Rudolf Kissling, Dr Henri Ott, Dr Jacob Roffler and Dr
Philippe Vaudens with the first author as the ninth member.

One or more of the authors have received or will receive benefits for
personal or professional use from a commercial party related directly or
indirectly to the subject of this article.

References

1. Clinical Standards Advisory Group. Epidemiology review: the
epidemiology and cost of low back pain. Annex to the CSAG report on

2. Andersson GBJ, Pope MH, Frymoyer JW, Snook S. Epidemiology
and cost. In: Pope MH, Andersson GBJ, Frymoyer JW, Chaffin DB,
eds. Occupational low back pain: assessment, treatment and preven-

3. Frymoyer JW, Cats-Baril WL. An overview of the incidences and

4. Herron LD, Turner J. Patient selection for lumbar laminectomy and
discectomy with a revised objective rating system. Clin Orthop 1985;
199:145-52.

5. Rish BL. A critique of the surgical management of lumbar disc disease

for herniation of a lumbar disc: additional experience with an objec-

7. Deyo RA. Practice variations, treatment fads, rising disability: do we

8. Stevens CD. The case for the Clinton plan for health care reform

9. Volinn E, Turczyn KM, Loeser JD. Epidemiology review: the epi-
demiology and cost of low back pain. Annex to the CSAG report on

assessment of the appropriateness of medical technologies. Int J

of coronary artery bypass graft surgery in New York State. JAMA
1993;269:753-60.

12. Phelps CE. The methodologic foundations of studies of the appro-

13. Davis RA. A long-term outcome analysis of 984 surgically treated

14. Eddy DM. A manual for assessing health practices and designing
practice policies: the explicit approach. American College of Physi-


