Standardized evaluation of long-term results after anterior lumbar interbody fusion

Abstract A total of 113 patients, excluding those with tumor, spondylitis, and idiopathic scoliosis, underwent anterior lumbar interbody fusion (ALIF) with autologous iliac crest graft between 1984 and 1991 at our department. The proportion of these who were failed back patients was higher than that reported in the literature. Evaluation of functional outcome was feasible in 80 patients, utilizing Oswestry and Marburg scores, which were closely intercorrelated. The overall results yielded an improvement in the Oswestry score of 35.7 percentage points. A subset of 52 patients who were evaluated twice, showed the same results at an average of 6.6 years as they did at 2.3 years following surgery. Functional results showed a weak correlation with postoperative height loss of the intervertebral space. Influencing factors for the functional result were: postoperative compensation claim, age, and obesity. Of the professional people involved, 19.4% did not return to any occupation. Patients satisfied with the result had significantly greater functional improvement. Younger patients with additional dorsal distraction prior to ALIF for reduction of severe spondylolisthesis fared better than patients with ALIF alone. The rate of complications was low and did not contribute to the postoperative functional result. On the basis of these results further prospective studies have been designed and are currently underway.

Key words Anterior lumbar fusion · Follow-up studies · Radiology · Low back pain · Outcome

Introduction

Anterior lumbar interbody fusion (ALIF) in its present form has gone through a long phase of evolution. As early as 1910, Kausch [21], based on the experience of Müller [29], planned a transperitoneal ALIF in a tumor case using a prepared femoral allograft. He abandoned his plan after exposure of the vertebrae, because the patient was seemingly doomed by the extent of the tumor. Capener, in 1932 [6], was the first to describe results after transperitoneal ALIF, though Burns’ 1933 publication [5] is usually quoted as a milestone for spinal fusion. In 1944 Iwahara [20] introduced the retroperitoneal approach for ALIF. Harmon 1960 [16], among others, propagated ALIF for degenerative instability. We prefer a modified O’Brien technique [30] with bicortical autologous bone grafts.

Indications

At our department, since 1984 ALIF has been the method of choice for spinal fusion in cases of proven or suspected instability above L5-S1 (retroperitoneal approach) and cases of spondylolysis with spondylolisthesis at L5-S1 with more than Meyerding grade 1 slip (transperitoneal approach) with prior posterior distraction instrumentation. Posterior lumbar interbody fusion is not used at our department for various reasons. Since 1993, with the availability of a more sophisticated posterior instrumentation...
system (DDS, Plus, Switzerland), the use of ALIF has decreased and is now restricted to the above-mentioned spondyloysis cases and cases above L5 with isolated pathology of the anterior column. Relative contraindications were: low-grade, single-level instability at L5-S1 (uninstrumented posterolateral fusion advocated), previous intraabdominal or retroperitoneal operations, known vascular abnormalities (e.g., aneurysm), extreme osteopenia, and purely posterior encroachment of the spinal canal. Absolute contraindications were the same as those pertaining to other invasive spinal fusion procedures.

Little data is available about ALIF patients beyond a 5-year follow-up period. The aim of the study presented is to answer the following questions for ALIF patients, excluding those with tumor, idiopathic scoliosis, and spondylitis:

- What are the overall functional results?
- What are the (subgroup) results at about 2 years and again over 6 years following surgery?
- Do the radiological results correlate with the functional results?
- Are there predictors for the outcome of the patients?
- Are the formerly used evaluation tools adequate compared with other well-established procedures?
- Can ALIF alone be recommended for the above mentioned group of patients?

In preparation of a controlled prospective study the answers to these questions, based on a retrospective study, are mandatory.

**Materials and methods**

From June 1984 to December 1991 at our Department of Orthopedic Surgery 113 patients, on the basis of the above-mentioned indications, underwent ALIF. Conservative treatment had to prove ineffective before surgery was considered. Instability was documented via functional radiographs [23, 31] and/or discography [1, 22]. Sequestered disc herniations, if suspected, were excluded via CT scan. Facet blocks were often necessary to confirm the planned extent of the spinal fusion.

All patients underwent a standardized clinical assessment before surgery. If necessary, a neurologic examination was also conducted. The operations were performed under general anesthesia by four experienced surgeons, one of whom (P.G.) operated on the criteria of Stauffer and Coventry [34] and has been in use in our department since 1986. A subgroup of 52 patients who sent back their questionnaires had once previously participated in a Marburg score follow-up, conducted on average 2.3 years following surgery.

The Oswestry score is described elsewhere [9]. It consists of ten equally weighted items, whose scores are totalled. Disability is graded into five categories: 0–20% of the maximum possible score = minimal, 21–40% = moderate, 41–60% = severe, 61–80% = crippled, 81–100% = bed bound or exaggerating patients. No distinction is made between leg and back pain. A German version was developed in cooperation with native speakers. The percentage change in the Oswestry score “AMS” was calculated, in a similar way to AOS.

The Marburg score consists of eight items with a total of 12 questions and a maximum total of 48 points, indicating absence of disability. No English version has been developed so far. It allows for discrimination to be made between back and leg pain. Items involving compensation claims and back-to-work behavior are included. The final scores, which indicate the functional state of patients, can be broken down into three categories: 0–25% of the maximum score = poor, 26–75% = fair, and 76–100% = good. The percentage change in the Marburg score “AMS” was calculated, in a similar way to ∆OS.

The study was conducted by two researchers (M.P., M.B.) other than the responsible surgeons. The data of each patient was taken from the patient files, and consisted of: diagnosis, weight and height at time of operation, Broca index, previous operations, technical details such as number and level of segments involved, graft type, name of the surgeon, postoperative stay, bracing, and complications.

The functional postoperative follow-up consisted of a self-evaluation questionnaire. The questionnaire was divided into a retrospective part (assessment of the preoperative status) and a present time part, which were clearly separated from each other. They were mailed out to the patients in combination with a short clinical evaluation form, and patients were asked to appear at our clinic or at their orthopedist for completion of the form and new radiographs. In cases of non-return, a second letter was mailed out after 12 weeks and an attempt was made to contact by phone.

The questionnaire consisted of two parts: the Oswestry (Disability Index) score and the Marburg score. The latter is based on the criteria of Stauffer and Coventry [34] and has been in use in our department since 1986. A subgroup of 52 patients who sent back their questionnaires had once previously participated in a Marburg score follow-up, conducted on average 2.3 years following surgery.

The Oswestry score is described elsewhere [9]. It consists of ten equally weighted items, whose scores are totalled. Disability is graded into five categories: 0–20% of the maximum possible score = minimal, 21–40% = moderate, 41–60% = severe, 61–80% = crippled, 81–100% = bed bound or exaggerating patients. No distinction is made between leg and back pain. A German version was developed in cooperation with native speakers. The percentage change in the Oswestry score “AMS” was calculated, in a similar way to ∆OS.

For both scores the maximum value had occasionally to be adjusted if a question had been left out by the patient.

Several other items were screened on a yes/no basis without being scored: improvement in quality of life, worsening after initial improvement of pain, retrospective agreement with the decision for operative treatment, and satisfaction with the result. The duration of preoperative pain was also determined.

The evaluation of radiological results utilized a newly developed method for calculating the normalized height differences on lateral radiographs between the preoperative and immediately postoperative segments, and between the immediately postoperative and follow-up segments, respectively.

First, the four projected corners of the topmost vertebra to be involved in the fusion were defined on the preoperative lateral radiograph. They were connected by diagonals. At the intersection of the diagonals, a bisector line was drawn to measure the projected AP diameter, B, of the vertebra. Lines were then drawn connecting the lower two corner points with the upper corner points of the lowest vertebra to be involved in the fusion (Fig. 1). Two lines, a1 and a2, resulted and their lengths were averaged, giving:

$$a_1 + a_2 \over 2 = A$$

The study was conducted by two researchers (M.P., M.B.) other than the responsible surgeons. The data of each patient was taken from the patient files, and consisted of: diagnosis, weight and height at time of operation, Broca index, previous operations, technical details such as number and level of segments involved, graft type, name of the surgeon, postoperative stay, bracing, and complications.

The functional postoperative follow-up consisted of a self-evaluation questionnaire. The questionnaire was divided into a retrospective part (assessment of the preoperative status) and a present time part, which were clearly separated from each other. They were mailed out to the patients in combination with a short clinical evaluation form, and patients were asked to appear at our clinic or at their orthopedist for completion of the form and new radiographs. In cases of non-return, a second letter was mailed out after 12 weeks and an attempt was made to contact by phone.

The questionnaire consisted of two parts: the Oswestry (Disability Index) score and the Marburg score. The latter is based on the criteria of Stauffer and Coventry [34] and has been in use in our department since 1986. A subgroup of 52 patients who sent back their questionnaires had once previously participated in a Marburg score follow-up, conducted on average 2.3 years following surgery.

The Oswestry score is described elsewhere [9]. It consists of ten equally weighted items, whose scores are totalled. Disability is graded into five categories: 0–20% of the maximum possible score = minimal, 21–40% = moderate, 41–60% = severe, 61–80% = crippled, 81–100% = bed bound or exaggerating patients. No distinction is made between leg and back pain. A German version was developed in cooperation with native speakers. The percentage change in the Oswestry score “AMS” was calculated, in a similar way to ∆OS.

For both scores the maximum value had occasionally to be adjusted if a question had been left out by the patient.

Several other items were screened on a yes/no basis without being scored: improvement in quality of life, worsening after initial improvement of pain, retrospective agreement with the decision for operative treatment, and satisfaction with the result. The duration of preoperative pain was also determined.

The evaluation of radiological results utilized a newly developed method for calculating the normalized height differences on lateral radiographs between the preoperative and immediately postoperative segments, and between the immediately postoperative and follow-up segments, respectively.

First, the four projected corners of the topmost vertebra to be involved in the fusion were defined on the preoperative lateral radiograph. They were connected by diagonals. At the intersection of the diagonals, a bisector line was drawn to measure the projected AP diameter, B, of the vertebra. Lines were then drawn connecting the lower two corner points with the upper corner points of the lowest vertebra to be involved in the fusion (Fig. 1). Two lines, a1 and a2, resulted and their lengths were averaged, giving:

$$a_1 + a_2 \over 2 = A$$