

The effect of timing of rehabilitation on physical performance after lumbar spinal fusion: a randomized clinical study

Lisa G. Oestergaard · Claus V. Nielsen ·
Cody E. Bünger · Karen Svidt · Finn B. Christensen

Received: 21 May 2012/Revised: 4 December 2012/Accepted: 5 February 2013/Published online: 6 April 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract

Background The number of patients undergoing lumbar spinal fusion procedures (LSF) has risen in recent years, but only few studies have examined different rehabilitation strategies for this patient group.

Purpose To evaluate the impact of initiating rehabilitation either 6 or 12 weeks after a LSF based on the patients' physical performance using the 6-min walking test (6MWT) and the Åstrand Fitness test (AF-test) as measurement.

Methods A multicentre RCT including 82 patients with degenerative disc diseases undergoing LSF randomly assigned to initiate rehabilitation either 6 or 12 weeks after surgery. Both groups received the same group-based rehabilitation. The main outcome measures were the 6MWT and the AF-test, secondarily questionnaire-based measures. Follow-up at baseline as well as at 3, 6 and 12 months after surgery.

Results Comparing the two groups no statistically significant difference was found in walking distance or fitness over time. In both groups, the patients achieved an overall increase in walking distance ($p < 0.01$), but no improvement in fitness. The 6MWT showed significant correlation (-0.37 to -0.59) with the questionnaire-based outcome measures ($p < 0.01$). The AF-test did not correlate to either the 6MWT or any of the questionnaire-based outcome measures.

Conclusion No difference was found in the effect of initiating rehabilitation either 6 or 12 weeks after LSF on the patients' physical performance in terms of fitness and walking distance. The 6MWT showed fair to moderate correlation to the questionnaire-based outcome measures. The AF-test showed no significant independent value, and we question its use in LSF patients.

Keywords Spinal fusion · Rehabilitation · Physical performance · Occupational therapy · Physiotherapy

L. G. Oestergaard (✉)
Department of Occupational Therapy and Physiotherapy,
Aarhus University Hospital, Noerrebrogade 44,
8000 Aarhus C, Denmark
e-mail: lisaoest@rm.dk

L. G. Oestergaard · C. E. Bünger · F. B. Christensen
Department of Orthopaedic Surgery,
Aarhus University Hospital, Noerrebrogade 44,
8000 Aarhus C, Denmark

L. G. Oestergaard · K. Svidt
Orthopaedic Department, Region Hospital of Silkeborg,
8600 Silkeborg, Denmark

C. V. Nielsen
Section of Social Medicine and Rehabilitation,
Department of Public Health, Aarhus University,
8000 Aarhus C, Denmark

Introduction

Recent years have seen a significant rise in the number of lumbar spinal fusion (LSF) procedures, but only few studies have reviewed different rehabilitation strategies after LSF [1–5]. A study published in 2003 compared the effect of three different postoperative rehabilitation programs initiated 3 months after LSF. They found that group-based intervention deploying a bio-psycho-social approach outperformed video instruction and intensive physiotherapy [2, 6]. A study published in 2010 documented the superiority of a psychomotor therapy intervention focusing on cognition, behaviour and motor relearning activated 3–12 weeks after LSF [1]. In 2010, a concept study

evaluated preoperative training, pre-surgical information, patient-controlled epidural analgesia and intensive post-operative mobilisation after LSF and total disc replacement. They found improved functional mobility, shortened hospitalisation and reduced costs after this intervention protocol [4, 7].

Reviewing the literature, we found that no studies have examined the optimal time for initiating a rehabilitation program after LSF. Recent practise is that patients visit their surgeon 12 weeks after a LSF and then begin active rehabilitation. However, three studies have documented a positive effect of active rehabilitation initiated already 4–6 weeks after herniated disc surgery or microdiscectomy [8–10]. They reported improvements in patients' ability to cope with pain and perform activities of daily living (ADL) and in their general disability level 1-year post-surgery [8–10].

In a previous publication, we concluded that an early start of rehabilitation (6 vs. 12 weeks) after lumbar spinal fusion resulted in inferior functional outcomes according to the Oswestery Disability Index and Dallas Pain Questionnaire at 1-year follow-up. With the use of these well-recognised and validated patient-based questionnaires, the improvements in the 12 week-group was up to four times better than that of the 6 week-group. This indicates that the start-up time of rehabilitation is an important contributing factor for the overall outcome [11].

In the literature on LSF, we only identified a single study evaluating the patients' physical performance by the use of physical tests [4]. Nielsen et al. used the two physical tests, timed-up-and-go and sit-to-stand, to evaluate the effect of their intervention on the patients' physical performance. However, the results of these tests did not correlate with the findings in their questionnaire-based outcome measures [4]. In the clinical setting, patients walking ability and fitness is considered to be both an indicator for surgery and effect of rehabilitation. Patients' walking ability and fitness have been measured with different methods. The most frequently used current tests are: the 5-min walking test, the 15-m walking test, the 6-min walking test (6MWT), the 1-min stair-climbing, and the Åstrand Fitness test (AF-test) [12–14]. The 6MWT is often used in clinical settings as a single measure of physical functional status [13, 15]. The AF-test has been used to measure aerobic fitness in non-surgical low-back pain patients, and is considered an easy test to perform in the clinical setting [16, 17].

Aim

The dual aim of the present paper was to evaluate the impact of initiating rehabilitation either 6 or 12 weeks after

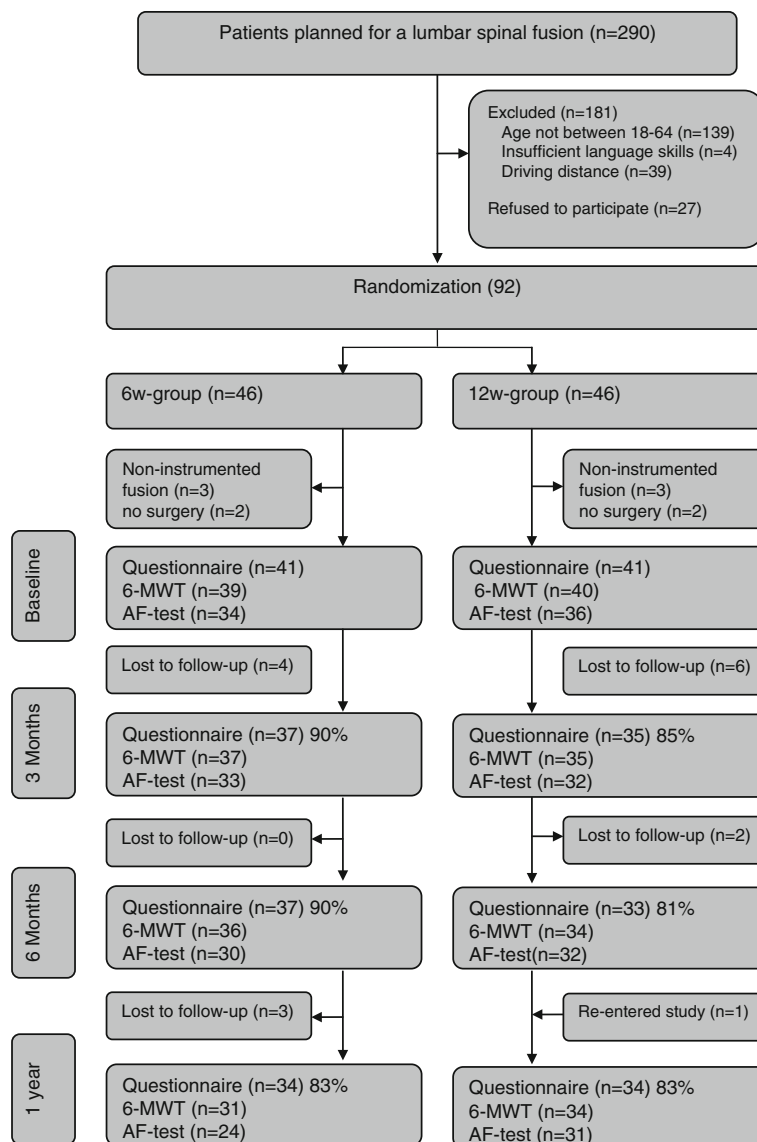
a LSF on the patients' physical performance. Secondary, to evaluate the value of the 6MWT and the AF-test in performance testing LSF patients.

Materials and methods

This randomized controlled trial (RCT) recruited patients in a one-to-one fashion for rehabilitation start either 6 (6w-group) or 12 weeks after surgery (12w-group). All patients were planned for an instrumented LSF due to degenerative disc disease or spondylolisthesis grade I or II. Exclusion criteria were age below 18 years or above 64 years, more than 100 km driving distance to the hospital and inability to speak and understand Danish.

Patients were included from January 2008 to January 2010 prior to surgery at three local spine centres. They received written and verbal information about the conditions of their participation and signed written informed consent. They were randomly assigned to the 6w-group or the 12w-group by use of sealed envelopes. The patients were block-randomized at the spine centre from which they were allocated. Those who included the patients did not participate in the treatment. The conditions of the study allowed no blinding of the therapists, surgeons, or the patients. As shown in Fig. 1, 290 patients underwent instrumented LSF of whom 181 patients did not meet the inclusion criteria. Among the 109 patients who met the criteria, 27 declined the invitation to participate (44 % male, mean age 47, 28 % were receiving disability pension), which left the study with 92 participants (49 % male, mean age 53, body mass index (BMI) 28, 28 % receiving disability pension).

The 12w-group received the standard treatment in which the patients met the surgeon 12 weeks after surgery to discuss their post-surgery condition based on new X-ray images of their lumbar spine. Rehabilitation commenced after this meeting. The 6w-group had a similar meeting scheduled 6 weeks after the surgery, and their rehabilitation started immediately after this meeting. Training was setup at the rehabilitation units and undertaken in groups of 3–6 persons. The two groups followed the same rehabilitation program, inspired by the Christensen study, which comprised four 2-h sessions [2]. Each session started with a 20-min exchange of experiences of pain and physical incapacity, problems and solutions in performing ADL, expressions of doubts that had arisen during the rehabilitation and straight tips and psychological support from kindred spirits. An exercise bike was used for warm-up. A physiotherapist instructed the patients in home exercises focusing on active stability training of the truncus and the large muscle groups. The patients were instructed in six active exercises and five stretching exercises. As the

Fig. 1 Patient flow

patients progressed, further four exercises with exercise band was introduced to the patients. The patients were advised to perform approximately 2×10 repetitions per exercise once a day. An occupational therapist joined one of the sessions to instruct the patients in proper ergonomics and working posture in relation to the ADL problems experienced by the patients and in relation to their return to work. For each participating patient, the following baseline characteristics were registered: gender, age, working status, diagnosis, type of surgery, quality of life, and previous LSF.

The physical parameters were the 6MWT and the AF-test. The 6MWT measures the distance that a patient can quickly walk on a flat, hard surface in 6 min [15]. The AF-test is a submaximal cycle ergometer aerobic fitness test calculating the patient's maximum oxygen uptake ($\text{ml O}_2/\text{min/kg}$) [12]. Both tests were performed based on the

standardised test protocols, and the physiotherapists were trained to conduct the tests in conformity with the protocols. We performed a pilot test on eight patients before study inclusion. Four patients were tested prior to surgery and four patients 3 months after having a LSF. Prior and after each test, the patients were asked to rate their current pain, measured on a numerical rating scale with "no pain" at the "0" end and "worst pain possible" at the "10" end of the scale [18, 19]. The physical tests were performed on other days than those at which the patients participated in the rehabilitation program. All patients first performed the 6MWT, then had a 15-min break, and then performed the AF-test. The patients also answered the following questionnaires: the Oswestery Disability Index (ODI) [20, 21], the Dallas Pain Questionnaire (DPQ) [22] and the Low Back Pain Rating Scale (LBPRS) [23]. All parameters were measured at baseline and 3, 6 and 12 months after surgery.

The power calculation performed for this study was based on the clinical evaluation of the primary outcome measure, ODI, previously reported [11]. Based on earlier studies, the standard deviation of the ODI was set to 20 points [1, 24]. A 15-point difference in this category was considered clinically significant. To fulfil these criteria, the study would need a total of 76 patients (38 in each group).

Statistical analysis

STATA 11.0 was used for statistical evaluation. Data analysis was performed according to the intention-to-treat principle and the significance level was set at 0.05. The one-way analysis of variance (ANOVA) was used with the group variable as the independent variable and the physical test as the continuous variable. The mean and 95 % confidence intervals (95 % CI) were reported for both the physical parameters. Spearman's rho was used to determine the level of correlation between the 6MWT and the AF-test. Similarly, Spearman's rho was used to determine the level of correlation between the two physical tests and ODI, DPQ and LPBRS. The expected normal walking distance matched by age, gender and weight was calculated according to Enright's formulae [25].

The study was approved by the Danish Data Protection Agency (J.nr. 2007-41-1605) and Ethical Committee (J.nr. 2007-0264).

Results

The patients in the two groups were comparable at baseline in terms of age, fusion level, surgical procedures, decompression and outcome measures (Table 1). By random the 6w-group included 53 % male patients, whereas the 12w-group included 42 % male patients. Three patients in the 6w-group and five patients in the 12w-group had a re-fusion within the first year after surgery. Two patients from each group had their instrumentation removed approximately 1 year after surgery.

At baseline, two patients from the 6w-group and one patient from the 12w-group could not complete the 6MWT due to pain. An equal number of patients were unable to complete the 6MWT at the 1-year follow-up. The follow-up rate of patients completing the 6MWT was 88 % at the 3-month follow-up, 84 % at the 6-month follow-up, and 78 % at the 1-year follow-up (Fig. 1).

At baseline, nine patients in the 6w-group and five patients in the 12w-group could not complete the AF-test because of pain (11 patients), heart disease (2 patients) or failure of achieving steady state (1 patient). At the 3-month follow-up, a total of 12 patients could not complete the AF-test because of pain (7 patients), heart disease (4 patients)

Table 1 Baseline characteristics of the study population

| | 6w-group (n = 41) | 12w-group (n = 41) |
|--|----------------------|-----------------------|
| Age \pm SD | 52.0 (SD 8.5) | 51.3 (SD 9.9) |
| Gender (% male) | 53 % (21) | 42 % (17) |
| Body mass index (kg/ m ² \pm SD) | 27.8 (SD 5.1) | 28.4 (SD 4.4) |
| Occupational status | | |
| Workforce (sick listed) | 64 % (26) | 61 % (25) |
| Social security | 7 % (3) | 7 % (3) |
| Early retirement | 5 % (2) | 12 % (5) |
| Disability pension | 24 % (10) | 20 % (8) |
| Diagnosis | | |
| Isthmic spondylolisthesis grade I–II | 12 % (5) | 15 % (6) |
| Degenerative disc disease | 88 % (36) | 85 % (35) |
| Surgical procedures | | |
| PLF | 76 % (31) | 85 % (35) |
| TLIF | 24 % (11) | 15 % (6) |
| Decompression | 85 % (35) | 73 % (30) |
| Earlier spondylolysis | 19 % (8) | 21 % (9) |
| Fusion levels (median/25.p/ 75.p) | 2 (1;2) | 2 (1;2) |
| Quality of life (EQ-5D) | 0.66 (0.39;0.72) | 0.63 (0.44;0.71) |

PLF posterior lumbar fusion, TLIF transformational lumbar interbody fusion

or failure of achieving steady state (1 patient). The follow-up rate of patients completing the AF-test was 85 % at baseline, 79 % at the 3-month follow-up, 76 % at 6-month follow-up, and 67 % at 1-year follow-up (Fig. 1).

Figure 2 shows that the mean walking distance at baseline was 446 m (411;480) in the 6w-group and 432 m (404;461) in the 12w-group. A comparison of the two groups showed no statistically significant difference in walking distance over time. Nor did we find any statistically significant overall difference between the two groups. In both the 6w-group and the 12w-group, we found a statistically significant positive change over time ($p < 0.001$). At the 1-year follow-up, the 6w-group had increased their mean walking distance by 57 m (25;89) ($p < 0.001$) and the 12w-group had increased their mean walking distance by 62 m (31;93) ($p < 0.001$). From Fig. 2 it may be seen that both the 6w-group and the 12w-group approached the calculated normal area at the 6-month and the one-year follow-up. Even so, we observed a statistically significant difference between either group and the calculated the normal area at all follow-up times ($p < 0.001$).

Figure 3 shows that the maximum oxygen uptake at baseline (ml O₂/min/kg) was 25.1(22.6;27.7) in the 6w-group and 26.5 (25.2;33.4) in the 12w-group. A comparison of the two groups revealed no statistically significant

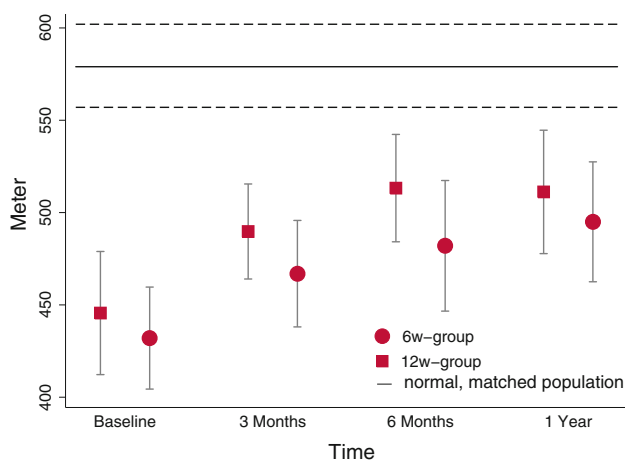


Fig. 2 Six minute walking test (6MWT). Dots showing mean meter walked with 95 % CI for the 6w-group and the 12w-group. Vertical line indicating the mean expected walking distance, with 95 % CI in a normal age- and sex-matched population

difference in maximum oxygen uptake over time. Similarly, no overall statistically significant differences between the groups were found. Nor did we ascertain any statistically significant difference from baseline to the 1-year follow-up within either of the two groups.

In the absence of any difference in physical parameters between the two intervention groups, we decided to analyse the merged data. Surprisingly, we found no correlation between the 6MWT and the AF-test. Analysis of the correlation between the 6MWT and the questionnaire findings [11] yielded the following results: ODI $r = -.37$ ($p < 0.01$), DPQ—daily activities $r = -.43$ ($p < 0.001$), DPQ—work/leisure $r = -.55$ ($p < 0.001$), DPQ—anxiety/depression $r = -.57$ ($p < 0.001$), DPQ—social concerns $r = -.59$ ($p < 0.001$), LBPRS back pain $r = -.41$ ($p < 0.01$), and LBPRS leg pain $r = -.45$ ($p < 0.001$). In

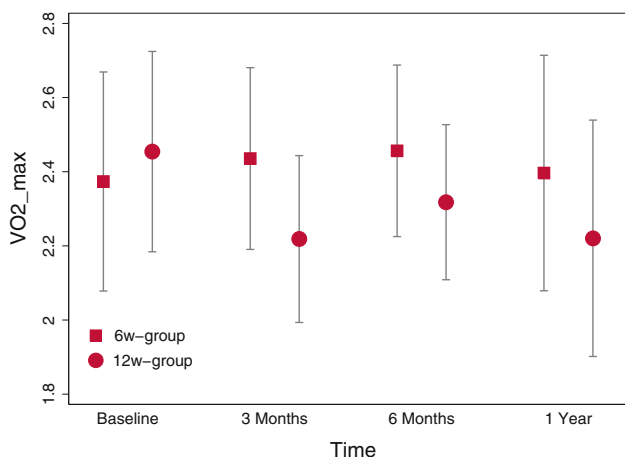


Fig. 3 Åstrand Fitness test (AF-test). Dots showing mean maximum oxygen uptake (ml O₂/min/kg) with 95 % CI for the 6w-group and the 12w-group

a similar comparison no significant correlation, whatsoever, was found between the AF-test and the ODI, the DPQ, the LBPRS back pain, and the LBPRS leg pain.

Both at baseline and at the 3-month follow-up, the patients experienced more pain after performing the 6MWT compared to prior to the test ($p < 0.0001$). However, at 6-month and 1-year follow-up, the 6MWT had no impact on pain. The AF-test did not increase the pain level at any time point.

Discussion

The main finding of the present study was that no clinical benefit in terms of longer walking distance or better fitness was accrued from early versus late initiation of rehabilitation following LSF. The 6MWT and the AF-test showed no difference over time (baseline—1-year follow-up), when comparing the 6w-group and the 12w-group. We have previously documented that the 12w-group experienced significantly larger improvements than the 6w-group in terms of ADL performance, functional mobility and back-pain [11]. However, we were unable to verify these results using the physical parameters, the 6MWT and the AF-test. A possible explanation is that the timing of rehabilitation mainly influences the patients' ability to cope with pain and perform daily activities, but that it does not affect the patients' walking ability or fitness.

We did not register the individual duration of the surgery. Based on the fact that there were no differences in fusion technique or site of fusion between the two groups we have no reason to expect the duration of surgery to differ between the two groups. Our power estimation was calculated on the basis of the questionnaire-based ODI and not the two physical parameters. We therefore, cannot rule out that our study may not have been sufficiently powered to demonstrate differences in the physical parameters. However, the absence of any tendency towards differences between the groups makes it unlikely that a more powerful design would have affected the overall conclusions.

We analysed the walking distance by gender because the literature suggests that women in general walk shorter distances than men [25]. At baseline, the women's mean walking distance was 413 m (382;444) compared with the men's 469 m (441;498) ($p < 0.01$). Women improved (79 m, 48;111) more than men (37 m, 7;66) ($p = 0.0476$) from baseline to the 1-year follow-up. We cannot rule out that the slightly skewed distribution of men in the two groups could have underestimated the effect in the 6w-group. However, we do not believe that the extra four men in the 6w-group could have influenced the overall findings of the study.

A limitation of the study relates to the 17 % drop-out rate at the 1-year follow-up. More than half of the drop-

outs were due to a re-LSF that was not associated with the rehabilitation. The remaining drop-outs included three non-responders in the 6w-group and four in the 12w-group. It has not been possible to gather information on their status or motives for non-response, but the largely identical distribution of patients lost to follow-up in the two groups makes it unlikely that drop-out has affected our overall conclusions.

Both the 6w-group and the 12w-group achieved an overall increase in 6MWT, most of which happened within the first 6 months after surgery (Fig. 2). The walking distance in general did not improve from 6 to 12 months post-surgery. We do not know if our patient population had already reached their maximum walking distance at 6 months, at which time their maximum walking distance remained significantly shorter than that of the general matched population. However, one could speculate that continued physical rehabilitation in the time span from 6 months to 1 year after surgery would support the patients in achieving a maximum walking distance comparable to that of the general population. Analysing the correlation between the 6MWT and the questionnaire-based outcome measures (ODI, DPQ, LBPRS), we found a statistically significant correlation between -0.37 and -0.59 , which is considered to be a fair to moderate correlation [26].

Analysing the AF-test, we surprisingly found no correlation with either the 6MWT or any of the questionnaire-based outcome measures. A number of patients had problems coping with the AF-test, which influenced their compliance, and this might have influenced the results. The AF-test was used as an outcome measure in a RCT from 2011 that examined an intervention to assist return to work among non-surgical low-back pain patients [16]. In the 2011 study, they found that the AF-test reflected the clinical differences found by the use of questionnaire-based outcome measures [16]. However, our patient group seems to differ from their patient group in terms of poorer fitness at baseline. We therefore, raise the question whether or not patients undergoing LSF can cope with the AF-test and if the information gathered with the AF-test provides us with valid information on patients' physical performance.

In future studies, it might be relevant to evaluate other physical parameters as an adjunct to the information gathered from the 6MWT.

Conclusion

In the conclusion, we found that walking distance and fitness were not influenced by early initiation of post-surgery rehabilitation. The physical parameters did not support an early initiation of rehabilitation, nor could it explain the positive outcome of late initiation of rehabilitation reflected

in questionnaire-based outcome measures. The overall outcome of this RCT seems to be in favour of postponing the initiation of home-based rehabilitation to 12 weeks after a LSF. The 6MWT reflected an overall improvement in both groups that seemed to occur already 6 months after surgery. The 6MWT showed significant correlation with the more generally used questionnaire-based outcome measures. The AF-test showed no significant independent value, and we question its use in LSF patients.

Acknowledgments For financial support Lundbeck Foundation (UCSF), the Danish Rheumatism Associations (Gigtforeningen), the Health Insurance Foundation (Helsefonden), the Danish Strategic Research Council, the Central Denmark Region and the Occupational Therapy and Physiotherapy Department, University Hospital of Aarhus. Thanks to patients, surgeons, physiotherapists, and occupational therapists participating in this study.

Conflict of interest None.

References

- Abbott AD, Tyni-Lenne R, Hedlund R (2010) Early rehabilitation targeting cognition, behavior, and motor function after lumbar fusion: a randomized controlled trial. *Spine (Phila Pa 1976)* 35(8):848–857
- Christensen F, Laurberg I, Bunger C (2003) Importance of the back-caffe concept to rehabilitation after lumbar spinal fusion: a randomized clinical study with a 2-year follow-up. *Spine* 28(23):2561–2569
- Fairbank J, Frost H, Wilson-MacDonald J et al (2005) Randomised controlled trial to compare surgical stabilisation of the lumbar spine with an intensive rehabilitation programme for patients with chronic low back pain: the MRC spine stabilisation trial. *BMJ* 330(7502):1233
- Nielsen PR, Jorgensen LD, Dahl B, Pedersen T, Tonnesen H (2010) Prehabilitation and early rehabilitation after spinal surgery: randomized clinical trial. *Clin Rehabil* 24(2):137–148
- Oestergaard LG, Maribo T, Bunger CE, Christensen FB (2012) The Canadian Occupational Performance Measure's semi-structured interview: its applicability to lumbar spinal fusion patients. A prospective randomized clinical study. *Eur Spine J* 21(1):115–121
- Sogaard R, Bunger C, Laurberg I, Christensen F (2008) Cost-effectiveness evaluation of an RCT in rehabilitation after lumbar spinal fusion: a low-cost, behavioural approach is cost-effective over individual exercise therapy. *Eur Spine J* 17(2):262–271
- Nielsen P, Andreasen J, Asmussen M, Tonnesen H (2008) Costs and quality of life for prehabilitation and early rehabilitation after surgery of the lumbar spine. *BMC Health Serv Res* 8:209
- Danielsen JM, Johnsen R, Kibsgaard SK, Hellevik E (2000) Early aggressive exercise for postoperative rehabilitation after discectomy. *Spine (Phila Pa 1976)* 25(8):1015–1020
- Dolan P, Greenfield K, Nelson RJ, Nelson IW (2000) Can exercise therapy improve the outcome of microdiscectomy? *Spine (Phila Pa 1976)* 25(12):1523–1532
- Kjellby-Wendt G, Styf J, Carlsson SG (2001) Early active rehabilitation after surgery for lumbar disc herniation: a prospective, randomized study of psychometric assessment in 50 patients. *Acta Orthop Scand* 72(5):518–524

11. Oestergaard LG, Nielsen CV, Bungler CE et al (2012) The effect of early initiation of rehabilitation after lumbar Spinal fusion: a randomized clinical study. *Spine (Phila Pa 1976)* 37(21):1803–1809
12. Anderson D (2004) The Åstrand-Ryhming test/method under the magnifying glass. A review of research articles
13. Enright PL (2003) The six-minute walk test. *Respir Care* 48(8):783–785
14. Smeets RJ, Hijdra HJ, Kester AD, Hitters MW, Knotterus JA (2006) The usability of six physical performance tasks in a rehabilitation population with chronic low back pain. *Clin Rehabil* 20(11):989–997
15. Carpo R (2002) ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med* 166(1):111–117
16. Jensen LD, Maribo T, Schiottz-Christensen B et al (2012) Counselling low-back-pain patients in secondary healthcare: a randomised trial addressing experienced workplace barriers and physical activity. *Occup Environ Med* 69(1):21–28
17. Smeets RJ, Soest MV (2009) The usability of a modified Astrand bicycle test to assess the aerobic capacity in patients with musculoskeletal pain and healthy controls. *Disabil Rehabil* 31(24):1988–1995
18. Mannion AF, Balague F, Pellise F, Cedraschi C (2007) Pain measurement in patients with low back pain. *Nat Clin Pract Rheumatol* 3(11):610–618
19. Haefeli M, Elfering A (2006) Pain assessment. *Eur Spine J* 15(Suppl 1):S17–S24
20. Fairbank J, Couper J, Davies J, O'Brien J (1980) The Oswestry low back pain disability questionnaire. *Physiotherapy* 66(8):271–273
21. Fairbank J, Pynsent P (2000) The Oswestry Disability Index. *Spine* 25(22):2940–2952
22. Lawlis GF, Cuencas R, Selby D, McCoy CE (1989) The development of the Dallas Pain Questionnaire. An assessment of the impact of spinal pain on behavior. *Spine* 14(5):511–516
23. Manniche C, Asmussen K, Lauritsen B et al (1994) Low Back Pain Rating scale: validation of a tool for assessment of low back pain. *Pain* 57(3):317–326
24. Brox JI, Reikeras O, Nygaard O et al (2006) Lumbar instrumented fusion compared with cognitive intervention and exercises in patients with chronic back pain after previous surgery for disc herniation: a prospective randomized controlled study. *Pain* 122(1–2):145–155
25. Enright PL, Sherrill DL (1998) Reference equations for the six-minute walk in healthy adults. *Am J Respir Crit Care Med* 158(5 Pt 1):1384–1387
26. Landis JR, Koch GG (1977) An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics* 33(2):363–374