ORIGINAL ARTICLE

P. H. Helmhout
C. C. Harts
J. B. Staal
M. J. J. M. Candel
R. A. de Bie

Comparison of a high-intensity and a low-intensity lumbar extensor training program as minimal intervention treatment in low back pain: a randomized trial

P. H. Helmhout (✉) · C. C. Harts
Department of Training Medicine
and Training Physiology,
Occupational Health and Safety Service
Royal Netherlands Army,
LGen Knoopkazerne, Mineurslaan 500,
3521 AG Utrecht, The Netherlands
Tel.: +31-30-2366605,
Fax: +31-30-2366412,
e-mail: tgtf@army.dnet.mindef.nl

J. B. Staal · R. A. de Bie
Department of Epidemiology,
Maastricht University,
Maastricht, The Netherlands

M. J. J. M. Candel
Department of Methodology and Statistics,
Maastricht University,
Maastricht, The Netherlands

Abstract In a randomized, observer-blinded trial, the effectiveness of 3-month high-intensity training (HIT) of the isolated lumbar extensors was compared to low-intensity training (LIT). Eighty-one workers with non-specific low back pain longer than 12 weeks were randomly assigned to either of the two training programs.

Training sessions were performed on a modified training device that isolated the lower back extensors. Total intervention time was limited to 5–10 min (one or two training sessions) per week. Training effects were assessed in terms of changes in self-rated degree of back complaints, functional disability, and general, physical and mental health. Secondary outcomes in this study were muscle strength and fear of moving the back (kinesiophobia). Outcomes were evaluated at 1, 2, 3, 6, and 9 months after randomization.

The results showed that the two treatment programs led to comparable improvements in all outcome measures, except for mean isometric strength at 1, 2, 3, 6, and 9 months and kinesiophobia score at 2 and 9 months of follow-up. The high-intensity training group showed a higher strength gain (24 to 48 Nm) but a smaller decline in kinesiophobia (2.5 and 3.4 points, respectively), compared to the low-intensity training group.

It can be concluded that high-intensity training of the isolated back extensors was not superior to a non-progressive, low-intensity variant in restoring back function in nonspecific (chronic) low back pain. In further research, emphasis should be put on identifying subgroups of patients that will have the highest success rate with either of these training approaches.

Keywords Low back pain · Randomized trial · Lumbar extension strength · Training device · Functional disability

Introduction

Physical fitness is an important precondition for military readiness. Contrary to civilian jobs, many military functions require above-average levels of physical strength and endurance to cope with daily task-related physical demands such as repetitive heavy lifting and carrying, marching or driving.

Overall, musculoskeletal injuries are by far the largest group of occupational complaints in the Royal Nether-lands Army (RNLA). Low back pain (LBP) is, together with psychological overload, the most frequent diagnosed disorder during office hours of Dutch military company doctors, taking nearly 15% of their weekly office hours. These figures are expected to rise in the near future: the RNLA population is slowly ageing due to a lower recruitment rate of younger soldiers and a proposed age shift in functional dismissal, from 55 to 58 years. The risk of chronic and recurrent back pain is known to be positively correlated to age [14]. Therefore, efforts to optimize the treatment of back disorders in our working population are

highly valuable, both from a social and occupational perspective.

Recent systematic reviews indicate that exercise therapy is a successful approach for the restoration of chronic and recurrent low back pain, at least in the short term [16, 23]. However, higher quality studies generally show a lack of treatment specificity of different exercise modalities, such as aerobic exercises, strength and endurance reconditioning or mobilizing exercises [2, 10, 12]. Moreover, controversy remains regarding the impact of a training stimulus – in terms of intensity, duration and frequency – on the reduction of LBP. Different explanations for this lack of specificity are given in the literature, such as nonspecific, more centrally induced training effects, e.g., a shift in pain perception [12], or a large heterogeneity in the chosen study populations [6]. If, indeed, no specific dimension or type of exercise therapy is superior to another in producing optimal therapeutic outcomes, other aspects gain importance, e.g., cost, facilities, personnel time and capacity, treatment affinity, expectation and the patient's compliance.

From this perspective, back strength and endurance training in chronic LBP (CLBP) patients, with the use of training devices, is an interesting concept for military (and comparable) populations. Firstly, RNLA personnel are, from their initial military education, used to participate in physical exercise programs, including progressive resistance training on equipment. Second, the RNLA is well-equipped with an extensive line of modern fitness devices, including state-of-the-art lower back machines, on all major military locations throughout the country.

In recent years, we have gained experience with high-intensity, isolated training of the lumbar extensors in military personnel with low back pain, using special training devices. We used a sports medicine approach, partly according to established exercise protocols [5, 18], in which the following principles were emphasized:

– Isolation of the lumbar extensors through fixation of the pelvis and thighs
– Training in the individual's full range of motion
– Avoiding 'sticking points' – i.e., a point in the range of motion in which a relatively high resistance is experienced [3] – by tuning the load curve of the weight stack to the individual's strength curve.

The results we observed in our population were satisfying to sometimes excellent in terms of pain relief and functional restoration. This was when giving a training stimulus of not more than 5–10 min (one or two training sessions) per week.

The main objective of this study was to perform a randomized controlled trial on the relative efficacy of a high-intensity, progressive resistance training program (HIT) of the isolated lumbar extensors. We used specific training devices for a group of workers with nonspecific LBP. We compared the progressive training group with a group of workers who participated in a low-intensity, non-progressive training program (LIT) of the same extent. Total intervention time of both "minimal intervention programs" was limited to 14 sessions of 5–10 min over a period of 12 weeks.

A multi-center study is currently underway to compare this minimal intervention strategy with the usual treatment of LBP in the RNLA. In this paper, we describe the results of the pilot-study, which evaluates the clinical importance of the treatment's progressive, muscle-strengthening component. The outcome of both the HIT and LIT intervention programs is assessed in terms of changes in self-rated degree of back complaints, functional disability, and general, physical and mental health. Secondary outcomes in this study are muscle strength and fear of moving the back.

Materials and methods

Study design

We performed an observer-blinded randomized clinical trial. Measurements of outcomes were taken before randomization (baseline), at 1, 2 and 3 months after randomization (treatment period) and at 6 and 9 months after randomization (follow-up period). The study took place at the training medicine department of the RNLA in Utrecht, The Netherlands.

Study population

The source population ($n=37,000$) consisted of military and civilian employees of the RNLA. In the second half of 1998, all military (company) doctors and physiotherapists of the RNLA were requested to recruit participants with CLBP for our study. In addition, advertisements in military union media called for volunteers. Written informed consent was to be obtained from all participants. Further, each volunteer was given to an intake test, which consisted of history taking by a sports physiotherapist, physical examination by a sports physician and pre-assessment of the isometric back strength by a human-movement scientist.
Inclusion criteria were:

– Male employees of the RNLA
– More than 12 weeks of continuous or recurrent (at least 3 times a week) episodes of back pain
– Pain localized between posterior iliac crests and angulus inferior scapulae
– Availability to visit our department 1 to 2 times a week during 12 consecutive weeks
– Willingness to abandon other treatment interventions for the lower back during the intervention period

Potential participants were excluded if they had undergone spinal surgery in the last 2 years, if they had severe back pain that was a hindrance to performing maximal isometric strength efforts, or if they had specific LBP, defined as herniated disc, ankylosing spondylitis, spondylolisthesis or other relevant neurologic diseases.

Assignment to the treatments

All included participants were randomly assigned to a high-intensity training group (HIT) or a low-intensity training group (LIT). Randomization was done by means of a computer-generated table

of random numbers and a restricted randomization procedure with a block size of 4 (provided by a researcher not involved in the project). Stratification was applied for baseline scores on the Roland Disability Questionnaire [22], with a cut-off level of 10 points, on the suspicion that initial level of back complaints influences the individual response to the exercise program.

Treatment content

The HIT program consisted of a 12-week, progressive resistance training of the isolated lumbar extensor muscle groups. The program included 10 training sessions (respectively 2 days/week in weeks 1–2 and 1 day/week in weeks 3–12) and four isometric strength tests (in weeks 1, 4, 8 and 12). The initial training load was set at approximately 35% of the maximal isometric back extension strength of the participant, measured in the first week. The goal of every training session was to perform 15 to 20 repetitions (weeks 1–2) or 10 to 15 repetitions (weeks 3–12) on the lower back machine, equivalent to approx. 50% and 70% of the one-repetition maximum (1 RM), respectively. If the participant was able to perform a higher number of repetitions, a 2.5 kg weight was added in the next training session. On the other hand, if the participant was unable to perform the minimal number of repetitions, the subsequent training load was lowered by 2.5 kg. This training protocol is partly based on existing protocols [5, 18], and partly on our own experiences. A comprehensive training protocol can be obtained from the authors.

In the LIT program a non-progressive, low-intensity resistance protocol was used. Throughout the complete training period, the training load was set at no higher than 20% of the maximal isometric strength, as measured in the first week. The goal of every training session was to perform 15 (1st and 2nd week after each test) or 20 repetitions (3rd and 4th week after each test) on the lower back machine. The lowest training load on the machine was a fixed-weight of 2.5 kg. In our opinion, resistance training at this level would not generate a significant physiological strength training stimulus.

In both exercise programs, training sessions were carried out on a standard lower back machine, used in most military sports gyms (Lower Back ROM, Technogym, Italy). For the purpose of this study, we modified the machine by applying a knee-lock system and a thigh-restraining belt. With these adjustments, hips and thighs could both be immobilized, thus allowing the participant only to move the isolated lower back. In addition, the original oval-shaped "cam" (i.e., the weight stack transfer), which applied an increasing load on the back towards extension, was replaced by a circular cam. In a pre-study we found that a circular shape better reflects the optimal load curve of healthy subjects (A. Breet, internal communication, 1998). See Fig. 1.

All training sessions in both programs were conducted by a trained physiotherapist (C.C.H.), who paid special attention to the execution of the training in terms of pace and movement. The flexion and extension of the lower back had to be executed in the participant's full range of motion. Movements had to be slow and controlled: moving in 2 s from maximal flexion to maximal extension (lifting the weight), and returning from maximal extension to maximal flexion in 4 s (lowering the weight). During this movement, emphasis was put on the hollowing and flattening of the lumbar lordosis. Every training session was preceded by a 5-min all-body warm-up on an arm/leg ergometer (Schwinn Airdyne Pro, Balans, Nieuwegein, The Netherlands). The weight load used and the number of repetitions during each training session were recorded.

Outcome measures

Self-assessed outcome measures

At the end of the treatment period and during follow-up, participants were asked to report the degree to which their back complaints had changed since the start of the program, expressed in percentages of increase or decrease. In addition, various aspects of back complaints were evaluated before and after treatment and during follow-up, by means of the following set of validated questionnaires:

– Roland Disability Questionnaire (RDQ): a 24-item scale, ranging from 0 to 24 points, reflecting functional disability due to LBP [22]. A high score reflects a high disability rate

Fig. 1 Modified lower back test and training machine, with subject on the left in flexed position and on the right in extended position
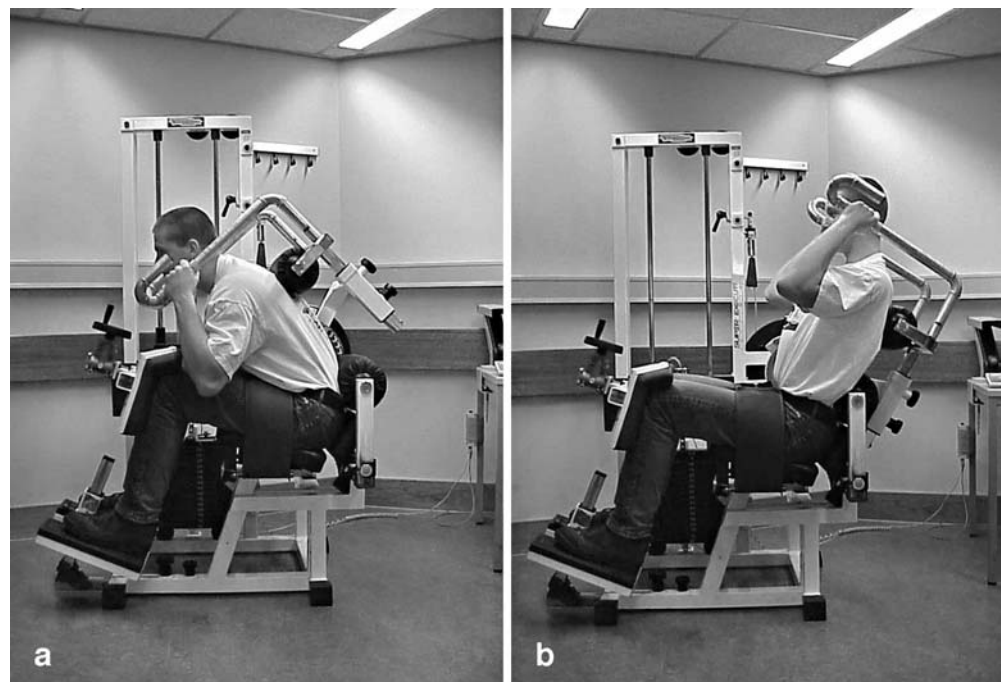
- Oswestry Questionnaire (OSW): a 10-item scale, ranging from 0 to 100%, to obtain a score for functional disability caused by LBP [4]. A high score reflects a high rate of pain-indicated limitations. This scale is considered to be complementary to the RDQ, as it is more sensitive to severe disability [1]
- Tampa Scale for Kinesiophobia (TSK): a 17-item scale, ranging from 4 to 68 points, measuring the extent to which a chronic back patient fears physical damage due to movement [7, 25]. A high score reflects a high degree of fear of movement
- MOS 36-item Short Form Health Survey (SF-36): ranging from 0 to 100%, to score self-experienced health related to quality of life [30]. Items are grouped into eight domains from which an overall summary score, a physical component score and a mental component score can be derived. A high score on each of the scales reflects a high level of self-experienced health [23]

Physical outcome measures

Individual back-extension strength progression was evaluated before and after treatment and during follow-up, using repeated isometric measurements. The same modified lower back machine was used in both the measurements and training exercises. To ensure valid and reliable measurements, we made several additional modifications: arm supports were attached, so that the position of the arms could be standardized in all tests. A lengthwise adjustable cable system could be attached through a notch system, after disconnection of the original weight stack-cable system. This allowed locking the arm of the machine at several angles, forwards and backwards. The force on the new system – the translation of the force on the machine arm (backward pushing or forward hanging) – was conveyed through a force sensor in the cable (99% accuracy up to 5 kN) to a personal computer, via a transducer (Digimax RS 232, Interface Mechatronic, Germany). Forces were expressed in torques (Nm), multiplying by the radius of the circular cam (0.12 m). A homemade software program graphically showed the measured force signals.

Participants underwent two measurements: a passive test and an active test. A passive test was conducted to correct for the non-muscular torque (nonMT). The nonMT is a combination of factors that, together with the torque of the lumbar extensors, provide the gross force signal. It consists of the upper body weight, dead weight of the machine arm and the stored energy of compressed and stressed tissue in the flexed or extended torso. Correction for these factors is necessary to determine the true strength of the back extensors, the so-called net muscular torque (netMT). NonMT was measured in five positions of the machine arm: 55° (full flexion), 40°, 25°, 10° and –5° (full extension) relative to the vertical. During the passive test, participants were asked to sit as relaxed as possible in each position. In the three positions in front of the center of gravity of the torso (55° to 25°), the arm supports could be used to hang on the fixed machine arm. If the participant was unable to reach one of the exterior positions (55° or –5°) because of pain or immobility, the measurement in this position was omitted. Every measurement took 6 s, of which the last 2 s were converted into a mean value of nonMT (in Nm).

An active test was carried out immediately after the passive test, to evaluate the maximal strength of the lumbar extensors in the same five positions. The machine arm was fixed at each angle and participants had to exert maximal backward force into the arm pad. The gross force measured was converted into a torque: the functional torque (FT). The netMT could then be calculated by subtracting the nonMT from the FT at every angle. A well-defined protocol was used to conduct the active test: a non-maximal trial in each position, followed by a maximal effort for 6 s (2 s to smoothly build up force, 2 s to maintain maximal force and 2 s to relax slowly). Force values between the 2nd and 4th s were converted into one mean value of netMT (in Nm).

A previous test-retest reliability study showed that the reproducibility of the netMT was lower between an initial and second test, than between a second and third test. Based on these results, we decided to do a pre-test at the intake.

All passive and active tests were conducted by two trained test leaders, each testing a different but equally distributed group of participants from both treatment groups. Participants were verbally encouraged in every test, but no visual feedback of test performance was given. At intake, the positioning of the participant into the lower back machine was assessed based on his anthropometrics. Subjects were instructed not to exercise within 24 h before testing. A complete, written test protocol can be obtained from the authors.

Follow-up period

Participants who completed the intervention program received questionnaires and strength measurements at 6 and 9 months after randomization. Additionally, participants were asked to indicate whether they had received any treatment(s) for their LBP during this follow-up period.

Blinding

The present study is an observer-blinded trial. The training sessions and data collection (both from questionnaires and from strength measurements) were conducted by different people. The two test leaders were not aware of the allocation of individual participants to one of the two groups. Conversely, the therapist who conducted the training sessions was not informed about participant scores on the strength tests; only relevant information about the measurements was given (i.e., the initial maximal isometric strength for the determination of the initial training load and the positioning into the machine).

The two intervention programs were presented to the participants as potentially equally effective for the lower back while targeting different aspects: strength in the HIT group versus mobility in the LIT group. Training and test sessions were planned at separate times during the week, to prevent participants from one group from exchanging treatment experiences with the other.

Analysis

All data were entered and analyzed with SPSS 11.0.1 (SPSS, 2001) or MlwiN [19]. Checks for missing and incorrect values and for normality were conducted. Scores of self-assessed degree of back complaints were broken down into "improved" (20% or more improvement) versus "not improved" (less than 20% improvement). We chose a cut-off point of 20%, which in our belief reflects a clinically relevant change in the health status of a participant.

Analyses were done using all randomized participants who provided any post-baseline data. Demographic and clinical characteristics, as well as baseline outcome measures, were summarized by descriptive statistics. Multiple Linear Regression statistics were used to assess baseline-adjusted, between-group differences in all but one of the outcome measures at 1, 2, 3, 6, and 9 months after randomization. Chi-square tests were used otherwise. According to the criteria suggested by Pocock et al [17], baseline covariates (characteristics and other outcomes) were entered into the regression model if related to the outcome variable ($r > 0.5$).

Mean strength development in participants who had a complete 9-month follow-up were displayed for both treatment groups. Multilevel analysis (MlwiN) was used to analyze possible trend differences between HIT and LIT in mean isometric strength across time, simultaneously correcting for baseline differences in mean strength. All tests were judged at a two-sided significance level of $\alpha = 0.05$.

Results

Patient flow and treatment compliance

Our recruitment efforts yielded 88 employees who visited our department in Utrecht. A study population of 81 participants with low back pain was selected, based on our inclusion and exclusion criteria. After signing informed consent, participants were randomly assigned to the HIT group (n=41) or to the LIT group (n=40).

Fig. 2 shows a flow chart of participants and withdrawals. The HIT group showed a higher treatment compliance than the LIT group: 29 participants (71%) in the HIT group, versus 19 participants (48%) in the LIT group, attended all 14 training sessions. Ten participants (24%), versus 13 participants (33%), missed one session; two participants (5%) versus eight participants (20%) missed two, three or four sessions.

In both groups, participants reported no co-interventions during the 12-week treatment period. Four participants (12%) in the HIT group and eight participants (28%) in the LIT group reported no back training and/or treatment during the follow-up period (months 3–9). Twenty-four participants (73%) in the HIT group and 14 participants (48%) in the LIT group participated in different kinds of exercise training, including on our specific device, and five participants (15%) in the HIT group and seven participants (24%) in the LIT group sought professional help for their backs in this period.

Baseline characteristics and outcomes

Table 1 shows demographic properties of the population as well as baseline values for the outcome measures. Despite our randomization strategy, the HIT and LIT groups showed small differences on several baseline values. The low-compliance group (missing two to four training sessions) showed considerably worse scores on the baseline RDQ, Oswestry and SF-36 scores, than the medium and good compliers.

Table 2 shows the post-treatment minus pre-treatment scores of the outcomes at 1, 2, 3, 6, and 9 months after randomization. No significant between-group differences

Fig. 2 Flow chart of participants and withdrawals (months of follow-up counted after randomization)
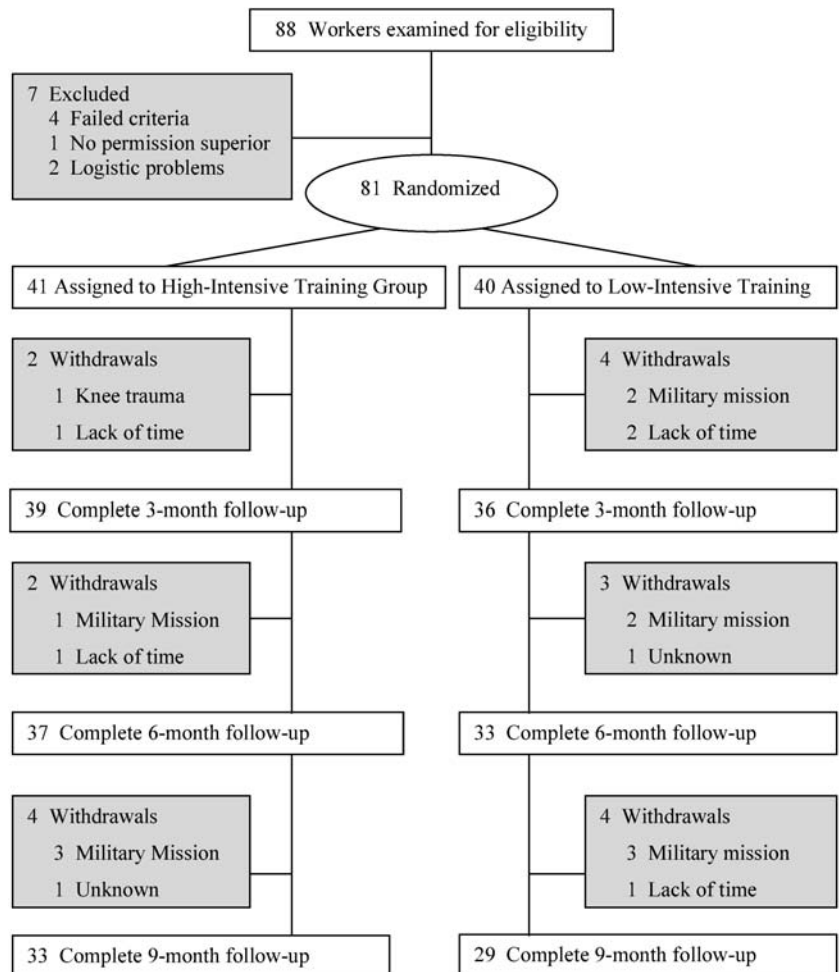
Table 1 Demographic and clinical characteristics (*MMH* manual materials handling) and baseline outcome measures (*RDQ* Roland Disability Questionnaire, *OSW* Oswestry Questionnaire, *TSK* Tampa Scale for Kinesiophobia, *SF-36* Short-Form-36 Health Survey) are presented for trial participants in both treatment groups, participants with medium and low compliance, and the withdrawals at intake. Values (from all outcome measures, mean values and SD) are presented in number of participants, unless otherwise stated

Parameters	HIT (N=41)	LIT (N=40)	With-drawals (N=6)	Medium compliance (n=29)[a]	Low compliance (n=8)[b]
Demographic characteristics					
Mean age in years (SD)	41 (10)	40 (9)	34 (13)	41 (11)	38 (12)
Range of age in years	21–55	22–55	21–50	22–55	22–55
Physically active (%)[c]	31 (76)	22 (55)	4 (67)	16 (55)	4 (50)
Military:civil employees	36:5	36:4	5:1	24:5	6:2
Recruited:referred patients	36:5	37:3	3:3	24:5	5:3
Clinical characteristics					
Time since first LBP episode					
3–6 months	0	3	–	–	–
6–12 months	5	7	–	–	–
1–5 years	16	14	–	–	–
≥5 years	19	16	–	–	–
Radiation[d]	14	15	–	–	–
Earlier therapies					
Physical therapy	13	16	–	–	–
Physiotherapy	15	24	–	–	–
Manual therapy	11	8	–	–	–
Orthopedic specialist	9	6	–	–	–
No therapy	5	2	–	–	–
Combination of earlier therapies					
1 therapy	21	12	–	–	–
2 therapies	11	19	–	–	–
≥3 therapies	2	2	–	–	–
Change of work due to LBP	4	4	–	–	–
Work absenteeism due to LBP	11	14	–	–	–
Frequency of work absenteeism					
<1 week in last year	4	3	–	–	–
1–3 weeks in last year	29	17	–	–	–
≥3 weeks in last year	8	20	–	–	–
Type of work					
Sedentary	24	25	–	–	–
Physical without MMH	6	8	–	–	–
Physical with MMH	5	4	–	–	–
Outcome measures					
RDQ score, 0–24 points	7.1 (4.8)	7.9 (4.8)	7.2 (6.0)	7.6 (4.3)	12.1 (1.7)
OSW score, 0–100%	13.9 (8.0)	17.4 (13.2)	14.7 (13.2)	18.6 (13.8)	25.3 (4.5)
TSK score, 17–68 points	37.0 (6.1)	40.0 (7.0)	37.0 (8.6)	37.8 (7.5)	35.4 (9.1)
SF-36 score, 0–100%					
Overall summary score	71.2 (14.2)	66.9 (15.2)	66.5 (22.4)	68.3 (14.8)	57.7 (12.6)
Physical component score	67.9 (18.1)	63.6 (20.8)	60.9 (22.9)	62.8 (21.7)	49.4 (16.6)
Mental component score	83.9 (18.2)	79.4 (19.8)	77.7 (27.5)	82.0 (18.4)	71.7 (17.8)
Strength (Nm)[e]					
Mean of five angles	233 (71)	203 (64)	208 (52)	228 (75)	212 (86)
55° flexion	254 (81)	229 (78)	275 (83)	249 (87)	229 (112)
40° flexion	255 (77)	214 (72)	229 (50)	234 (81)	204 (104)
25° flexion	245 (72)	206 (72)	227 (66)	237 (201)	196 (100)
10° flexion	212 (77)	176 (70)	184 (46)	201 (83)	176 (77)
5° extension	201 (76)	163 (80)	132 (96)	200 (83)	166 (82)

[a]Participants who missed one training session
[b]Participants who missed two (N=6), three (N=1) or four (N=1) training sessions
[c]Defined as 30 min or more physically active per day (walking/cycling to work, physical work activities, sports, physically active hobbies)
[d]Defined as pain radiation or tingling in the legs
[e]Isometric net muscular torque

Table 2 Between-group differences in the outcome measures at 1, 2, 3, 6, and 9 months after randomization (*HIT* high-intensity training group, *LIT* low-intensity training group, *CI* confidence interval)

Outcome measures	Mean values (SD)		Between-group differences[c] (95% CI)	P value
	HIT[a]	LIT[b]		
Improved (%) [d]				
3 months	64.1	50.0	14.1	0.22[e]
6 months	78.4	70.0	8.4	0.41[e]
9 months	84.8	72.4	12.4	0.23[e]
Improved (%)[f]				
3 months	39.4 (32.1)	27.5 (30.6)	11.9 (–1.4 to 25.3)	0.08[e]
6 months	51.1 (38.7)	42.1 (32.9)	9.0 (–8.1 to 22.8)	0.35[e]
9 months	63.0 (31.0)	49.8 (33.2)	13.2 (–3.1 to 29.5)	0.11[e]
RDQ (0–24 points)				
1 month	4.7 (4.1)	6.0 (4.7)	–0.6 (–2.0 to 0.9)	0.42
2 months	4.1 (3.5)	5.7 (4.6)	–0.9 (–2.3 to 0.5)	0.19
3 months	4.2 (4.0)	5.5 (4.4)	–0.6 (–2.3 to 1.1)	0.47
6 months	4.5 (4.5)	4.3 (4.2)	0.7 (–1.4 to 2.9)	0.50
9 months	3.2 (4.2)	3.2 (3.3)	0.6 (–1.2 to 2.4)	0.52
OSW (%)				
1 month	13.2 (8.6)	15.1 (13.0)	0.5 (–2.9 to 3.9)	0.77
2 months	11.5 (8.7)	13.2 (12.2)	0.7 (–2.3 to 3.8)	0.63
3 months	10.8 (7.6)	14.0 (12.5)	–0.5 (–3.7 to 2.8)	0.55
6 months	10.3 (9.5)	13.6 (13.4)	–0.4 (–4.8 to 4.0)	0.87
9 months	9.0 (8.5)	11.6 (14.0)	0.5 (–3.0 to 4.0)	0.77
TSK (4–68 points)				
1 month	36.5 (6.9)	38.8 (7.0)	0.2 (–2.0 to 2.3)	0.88
2 months	36.5 (8.0)	36.9 (6.9)	2.5 (0.3 to 4.8)	0.03
3 months	35.0 (8.3)	37.1 (7.3)	0.6 (–2.1 to 3.2)	0.24
6 months	35.0 (8.9)	36.5 (6.9)	1.2 (–1.9 to 4.3)	0.44
9 months	34.9 (7.8)	33.8 (6.4)	3.4 (0.5 to 6.4)	0.03
SF-36 overall (%)				
1 month	72.3 (13.9)	71.8 (14.1)	–2.6 (–6.9 to 1.7)	0.23
2 months	74.6 (14.1)	71.5 (16.0)	–0.1 (–4.5 to 4.3)	0.96
3 months	76.4 (14.6)	73.3 (15.9)	0.5 (–4.3 to 5 4.)	0.83
6 months	76.9 (14.5)	75.0 (15.7)	0.8 (–5.6 to 7.2)	0.81
9 months	78.0 (14.4)	78.7 (14.3)	–3.1 (–9.3 to 3.0)	0.31
SF-36 physical (%)				
1 month	69.3 (18.3)	70.8 (18.1)	–4.1 (–11.2 to 3.1)	0.26
2 months	73.6 (17.1)	69.8 (22.1)	1.5 (–4.9 to 8.0)	0.64
3 months	76.6 (17.9)	73.5 (21.5)	1.3 (–6.1 to 8.7)	0.73
6 months	76.8 (20.0)	74.1 (25.1)	1.9 (–9.0 to 12.7)	0.73
9 months	82.6 (16.0)	80.6 (19.5)	0.2 (–8.4 to 8.8)	0.96
SF-36 mental (%)				
1 month	84.2 (18.3)	82.7 (18.6)	–2.0 (–7.3 to 3.3)	0.46
2 months	86.3 (15.9)	81.8 (20.1)	1.5 (–4.8 to 7.7)	0.64
3 months	86.5 (17.5)	82.5 (17.9)	1.9 (–4.5 to 8.3)	0.56
6 months	87.1 (16.4)	84.9 (19.4)	1.3 (–5.7 to 8.4)	0.71
9 months	84.4 (19.7)	85.6 (18.3)	–3.8 (–11.2 to 3.6)	0.31
Strength (Nm)[g]				
1 month	268 (88)	202 (66)	40 (14 to 66)	0.00
2 months	274 (71)	210 (85)	58 (22 to 94)	0.00
3 months	265 (73)	217 (76)	31 (12 to 50)	0.00
6 months	242 (75)	221 (66)	24 (1 to 47)	0.04
9 months	270 (68)	217 (74)	29 (2 to 55)	0.03

[a]*N*=39 at 3 months, *N*=37 at 6 months and *N*=33 at 9 months of follow-up
[b]*N*=36 at 3 months, *N*=33 at 6 months and *N*=29 at 9 months of follow-up
[c]Multiple linear regression statistics are used to assess baseline-adjusted, between-group differences for all outcome measures except "improved"
[d]Percentage participants with a self-assessed percentage decrease (+) in back complaints of more than 20% compared to baseline (measured only after the training program)
[e]Chi-square test
[f]Mean self-assessed percentage decrease (+) or increase (–) in back complaints compared to baseline (measured only after the training program)
[g]Mean isometric net muscular torque of five angles

Fig. 3 Strength development (mean isometric strength of 5 angles) in participants from both intervention groups who had a complete 9-month follow-up
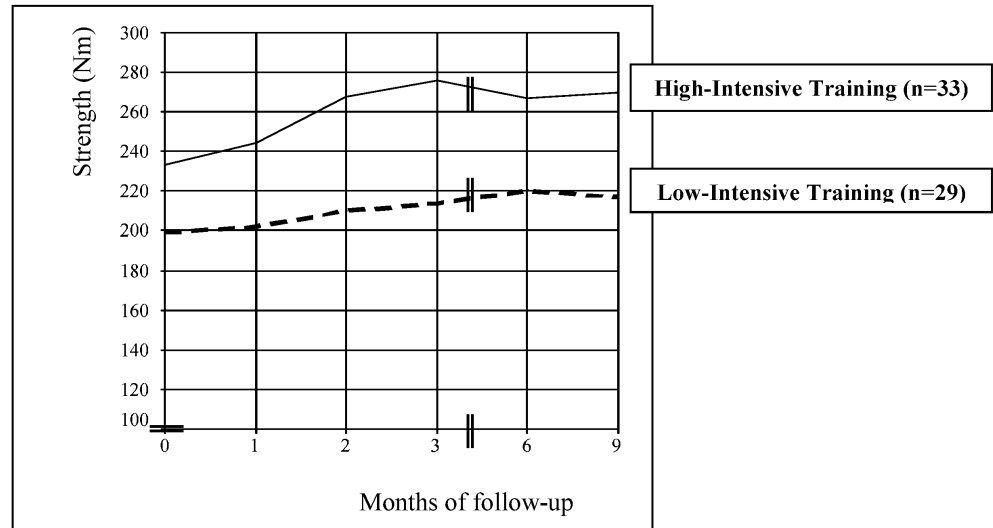

Table 3 Multilevel analysis for isometric back strength across time (*Time* linear trend, *Time²* quadratic trend, *Time³* cubic trend, *Time⁴* quartic trend)

Predictor	Parameter estimate	Standard error	T value	P value
Time[a]	136.985	13.365	10.25	<0.001
Time×baseline strength	0.853	0.049	17.408	<0.001
Time×group[b]	11.970	4.549	2.631	0.009
$Time^2$	−95.710	6.428	−14.890	<0.001
$Time^2$×baseline strength	−0.202	0.015	−13.467	<0.001
$Time^2$×group	−1.018	0.467	−2.180	0.032
$Time^3$	19.547	1.135	17.222	<0.001
$Time^3$×baseline strength	0.013	0.001	13	<0.001
$Time^4$	−1.168	0.065	−17.969	<0.001
Random effects				
Variance of random slope for time	242.736	62.217		
Variance of random slope for $time^2$	1.685	0.668		
Error variance	1272.918			

[a]Test moment at 1, 2, 3, 6, and 9 months, respectively
[b]Treatment group: HIT = 1, LIT = 0. Group effects were corrected for baseline differences in strength

in outcomes were found, except for mean isometric strength on all test moments (HIT scored 24–58 Nm higher than LIT) and for TSK score at 2 and 9 months (LIT scored 2.5 and 3.4 points, respectively, lower than HIT). Moreover, there was a tendency towards a significant difference in the percentage of self-assessed improvement between both groups at 3 months, the HIT group scoring 12% higher than the LIT group.

In Fig. 3, back strength measurements (mean isometric strength of five angles) for both groups are displayed for all test moments in the 9-month follow-up. The trend we see in this figure has been analyzed further, employing multilevel analysis.

Table 3 presents strength across time, corrected for the effects of baseline differences in strength between both groups.

The results show that the difference in strength development between HIT and LIT follows a curvilinear trend, which can be described by: $11.970*t–1.018*t^2$. In the first ±6 months (t=5.88), the group difference in strength in-creases until about 35 Nm, in favor of HIT, followed by a decrease. During the total study period (9 months), there is a positive effect of HIT versus LIT on strength.

Moreover, the effect of one unit of increase in the base-line value of strength on the strength at time t can be ex-pressed as: $0.853*t–0.202*t^2+0.013*t^3$. This means that, at every test moment after randomization (t=1, 2, 3, 6, and 9), the initial strength measurement has a positive effect on strength at that moment.

Discussion

Main findings

In a randomized, observer-blinded trial, the effectiveness of a 12-week high-intensity exercise program of the iso-lated lumbar extensors was compared to a low-intensity program. Our results show that the two treatment pro-grams led to comparable improvements in all outcome

measures, except for mean isometric strength at 1, 2, 3, 6, and 9 months, and TSK score at 2 and 9 months of follow-up. The high-intensity training group showed a higher strength gain (24–58 Nm) and a smaller decline in fear of movement (2.5–3.4 points less), compared to the low-intensity training group.

Generalizability of the study population

The population of this study represents a group of male employees with low back pain with, on average, long-standing, but moderate, functional disability that enables them to work with only occasional absenteeism.

Baseline scores on functional disability (RDQ and OSW) were comparable to those of other study populations with moderate severe low back pain [2, 8, 12]. However, scores on fear of movement (TSK) [25, 26] and self-experienced health (SF-36) [9, 29] were generally higher. Compared to a Dutch population of healthy men between 25 and 55 years of age [29], our CLBP population scored 8 points lower on the SF-36 Overall Summary Score, 16 points lower on the SF-36 Physical Component Score and only 1 point lower on the SF-36 Mental Component Score. Mc Horney et al [15] noted that a difference of 23 points in the domain of "physical functioning" (PF) reflects the impact of a complicated chronic medical condition, while a difference of 27 points in the domain of "mental health" (MH) is equivalent to the impact of serious depressive symptoms. Our population scored 11 points lower on PF and 2 points lower on MH, indicating that perceived health problems were not severe and mainly focused on physical aspects.

The fact that a strictly male population was used prevents this study from possible gender-related bias, e.g., different morphologic adaptations of strength training in men and women [6, 21].

Finally, participants in this study were partly recruited via advertisements (N=73), which may have attracted a group with good self-motivation to alleviate their back problems, compared to the participants (N=8) who were referred by a clinician. Nevertheless, recruitment bias is not to be expected, since both groups were randomized rather evenly into each of the interventions. A comparable distribution was found in the subgroups (withdrawals, medium compliers, low compliers). We therefore do not suspect a recruitment bias.

Considerations regarding the main findings

The main goal of the current study was to compare the effectiveness of a high-intensity and low-intensity back strengthening program, both of limited dimension. We hypothesized that the progressive, high-intensity character of the intervention was a key element in the restoration of back function.

Our results do not confirm this hypothesis: both training groups showed comparable improvements on all outcomes.

This is in agreement with the recent literature, in which training intensity is not seen as the crucial factor in restoring back function. In a systematic review, Van Tulder et al [24] conclude that there is strong evidence (level 1) that strengthening exercises are not more effective than other types of exercises. This conclusion is mainly based on four high-quality trials that made use of progressive trunk and leg exercises [6, 10, 11], sometimes in combination with other treatments [2]. Total intervention time in these studies varied from 15 to 90 min, from two to three times a week, from 4 to 54 weeks. Mannion et al [12] mention a "paradox in the concept of training specificity," in which the precise mode of therapy seems of limited importance for its effectiveness. More centrally induced training effects, like modifications in pain perception and analgesic effects of exercise, are suggested to be responsible for the improvements in back function. They conclude that physical restoration of moderate CLBP using specific types of exercise seems not to be justified [13].

Despite the lack of differences in outcomes, both groups did improve over time. Mainly for practical reasons and reasons of study power, we chose not to divide the total study population into three groups by including a waiting-list group. Consequently, we are not able to judge whether possible placebo, learning, novelty or Hawthorne effects might have led to an overestimation of the post-treatment improvements found. Since the majority of our participants already had longstanding back pain before entering the study, the post-treatment improvements found in this study cannot simply be explained by spontaneous recovery. The continuous character of these improvements – up to half a year after the end of the exercise program – could be explained from the relative high percentage of participants (84%) continuing some sort of back training or treatment during the follow-up period.

Analyses on strength development show that the strength-gaining effect of the HIT program exceeds the training period by approximately another 3 months. This is an interesting finding, considering the minimal intervention dimensions of the program. It might be explained by the high percentage of patients who continued to participate in exercise training in the follow-up period. Moreover, we can learn from the analysis that a higher initial back strength leads to a higher strength increase in time. From a physiological point of view, we know that strength training normally follows the law of diminishing returns, i.e., individuals with low initial strength have a higher strength development potential than individuals who are already strong at baseline [27]. We do not have a solid explanation for these findings, other than speculative ones, e.g., back patients with a stronger back – possibly with a more physically active background – have more faith in, less fear of, and higher expectations for our back strengthening approach.

An interesting finding is the higher degree of functional disability and lower self-experienced (physical) health at baseline among the low compliers, i.e., participants who missed two to four training sessions, compared to the medium and high compliers. It seems that the severity of complaints at least partly influences the motivation to participate in the treatment program.

Limitations of the study

The contrast between both interventions appeared to be less than expected. The training load for the LIT group was set at a maximum of 20% of the initial mean isometric strength of five angles. We observed that, for some participants in the LIT group with extreme low back strength at baseline, even a weight-stack without loads, together with the upper body weight in flexed position, exceeded the 20% intensity training stimulus. These factors might have diminished the contrast between both exercise programs. The LIT group showed a somewhat lower treatment compliance than the HIT group, which in its turn may have enlarged the intervention contrast between both groups.

The number of dropouts in our study was relatively high but evenly distributed among the two training groups, as well as the reasons for withdrawal. There was no indication that treatment dropouts had different characteristics and baseline outcomes than the remaining participants.

Comparison of treatment effects with references

By using effect sizes, calculated as [post-mean – pre-mean]/pre-standard deviation, the outcome responses of our study can be compared with those of similar intervention studies in comparable study populations. We only looked at the outcome responses that showed (near) significant post-treatment differences between the HIT and LIT group: percentage self-assessed improvement, mean strength, and fear of movement.

Manniche et al [11] reported 74% responders ("satisfied to very satisfied post-treatment") in a group that had followed a high-intensity back strengthening program, and 42% responders in a low-intensity group. These percentages are comparable to our results, in which "responder" is defined as 20% or more improvement in back function: 64% responders in the HIT group and 50% in the LIT group. Leggett et al [9] reported a higher responder rate of 80% after a multimodal exercise program including isolated back extensor strengthening, using a three-level appraisal score: better, same or worse.

With respect to isometric back-extension strength, effect sizes for our HIT group of 0.56–0.65 (flexion–exten-

sion) at 12 weeks, were comparable to those in a study by Risch et al [20] using back-isolating devices for patients with long-standing CLBP: 0.61–0.68 at 10 weeks. The effect sizes of our LIT group were somewhat higher (0.00–00.36) than those of their waiting-list group (0.13–0.00), which can be explained by a combination of training response and attention bias (Hawthorne effect). In a study by Leggett et al [9] using isolated back strengthening as part of a multimodal exercise intervention, higher effect sizes of 1.03 (better), 0.86 (same) and 0.65 (worse) were reported in patients from two outpatient CLBP treatment centers. The above-mentioned studies, as well as our study, used their training machines as measurement devices, which presumably overestimates the effect sizes found.

To the authors' knowledge, no intervention study comparable to ours has been done using the Tampa Scale for Kinesiophobia.

Overall, few studies have focused on the use of specific training devices as the only intervention treatment for low back pain, and no study has used a "minimal intervention strategy" comparable to ours. The magnitude of improvements in back function in the current study are in line with those reported in other studies using more extended exercise programs. This finding does not confirm a dose-response relationship between exercise and back complaints, suggested by Bronfort et al [2]. Vuori [28] mentions that physiological training principles, in terms of type and dose of exercise, cannot simply be applied when other than physiological conditioning factors of physical activity are involved, e.g., individual preferences and perceptions.

Conclusions

In concordance with the recent literature, we were unable to demonstrate that progressive, high-intensity training of the isolated back extensors was superior to a non-progressive, low-intensity variant in restoring back function. In terms of fear of movement, the gradual training approach even seems favorable. Nevertheless, other findings in this study indicate that some individuals with chronic LBP might benefit more from an aggressive approach: our high-intensity training group showed a trend towards a higher improvement rate immediately after treatment, as well as a higher treatment compliance and a higher willingness to participate in physical exercise in the longer term.

In further research, emphasis should be put on identifying subgroups of patients that will have the highest success rate with either of these different training approaches. Clearly, larger studies with more power are needed to do such an analysis.

References

1. Beurskens AJ, de Vet HC, Köke AJ, van der Heijden GJ, Knipschild PG (1995) Measuring the functional status of patients with low back pain. Assessment of the quality of four disease-specific questionnaires. Spine 20:1017–1028

2. Bronfort G, Goldsmith CH, Nelson CF, Boline PD, Anderson AV (1996) Trunk exercise combined with spinal manipulative or NSAID therapy for chronic low back pain: A randomized, observer-blinded clinical trial. J Manipulative Physiol Ther 19:570–582

3. Darden E (1990) The Nautilus Book. Contemporary Books, Chicago

4. Fairbank JC, Couper J, Davies JB, O'Brien JP (1980) The Oswestry low back pain disability questionnaire. Physiotherapy 66:271–273

5. Graves JE, Webb DC, Pollock ML, Matkozich J, Leggett SH, Carpenter DM, Foster DN, Cirulli J (1994) Pelvic stabilization during resistance training: Its effect on the development of lumbar extension strength. Arch Phys Med Rehabil 75:210–215

6. Hansen FR, Bendix T, Skov P, Jensen CV, Kristensen JH, Krohn L, Schioeler H (1993) Intensive, dynamic back-muscle exercises, conventional physiotherapy, or placebo-control treatment of low-back pain: A randomized, observer-blind trial. Spine 18:98–108.

7. Kori SH, Miller RP, Todd DD (1990) Kinisiophobia: A new view of chronic pain behavior. Pain Management Jan/Feb:35–43

8. Kuukkanen T, Mälkiä E (1996) Muscular performance after a 3 month progressive physical exercise program and 9 month follow-up in subjects with low back pain. A controlled study. Scand J Med Sci Sports 112–121

9. Leggett SH, Mooney V, Matheson LN, Nelson B, Dreisinger T, van Zytveld J, Vie L (1999) Restorative exercise for clinical low back pain: a prospective two-center study with 1-year follow-up. Spine 24:889–898

10. Ljunggren AE, Weber H, Kogstad O, Thom E, Kirkesola G (1997) Effect of exercise on sick-leave due to low back pain. Spine 22:1610–1617

11. Manniche C, Lundberg E, Christensen I, Bentzen L, Hesselsøe G (1991) Intensive dynamic back exercise for chronic low back pain: a clinical trial. Pain 47:53–63

12. Mannion AF, Müntener M, Taimela S, Dvorak J (1999) A randomised clinical trial of three active therapies for chronic low back pain. Spine 24:2435–2448

13. Mannion AF, Junge A, Taimela S, Müntener M, Lorenzo K, Dvorak J (2001) Active therapy for chronic low back pain: part 3. Factors influencing self-rated disability and its change following therapy. Spine 26:920–929

14. Marras WS (2000) Occupational low back disorder causation and control. Ergonomics 43:880–902

15. McHorney CA, Kosinski M, Ware JE Jr (1994) Comparisons of the costs and quality of norms for the SF-36 health survey collected by mail versus telephone interview: Results from a national survey. Med Care 32:551–567

16. Mior S (2001) Exercise in the treatment of chronic pain. Clin J Pain 17: S77–85

17. Pocock SJ, Assmann SE, Enos LE, Kasten L (2002) Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. Stat Med 21:2917–2930

18. Pollock ML, Leggett SH, Graves JE, Jones A, Fulton M, Cirulli J (1989) Effect of resistance training on lumbar extension strength. Am J Sports Med 17:624–629

19. Rasbash J, Browne W, Goldstein H, Yamg M, Plewis I, Draper D, Healy M, Woodhouse G (1999) A user's guide to MlwiN. Institute of Education, London

20. Risch SV, Norvell NK, Pollock ML, Risch ED, Langer H, Fulton M, Graves JE, Leggett SH (1993) Lumbar strengthening in chronic low back pain patients. Spine 18:232–238

21. Rissanen A, Kalimo H, Alaranta H (1995) Effect of intensive training on the isokinetic strength and structure of the lumbar muscles in patients with chronic low back pain. Spine 20:333–340

22. Roland M, Morris R (1983) A study of the natural history of back pain. I. Development of a reliable and sensitive measure of disability in low-back pain. Spine 8:141–144

23. Shmueli A (1998) The SF-36 profile and health-related quality of life: an interpretative analysis. Qual Life Res 7: 187–195

24. Tulder van M, Malmivaara A, Esmail R, Koes B (2000) Exercise therapy for low back pain: a systematic review within the framework of the Cochrane Collaboration Back Review Group. Spine 25:2784–2796

25. Vlaeyen JW, Kole-Snijders AM, Boeren RG, van Eek H (1995) Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. Pain 62:363–372

26. Vlaeyen JW, Kole-Snijders AM, Rotteveel AM, Ruesink R, Heuts PH (1995) The role of fear of movement/(re)injury in pain disability. J Occup Rehabil 5:235–252

27. Vrijens J (1984) Basis voor verantwoord trainen. PVLO, Gent, p 120

28. Vuori IM (2001) Dose-response of physical activity and low back pain, osteoarthritis, and osteoporosis. Med Sci Sports Exerc 6:S551–586

29. Ware J, Sherbourne C (1992) The MOS 36 item Short Form Health Survey (SF-36). Med Care 30:473–483

30. van der Zee K, Sanderman R (1993) Het meten van de algemene gezondheidstoestand met de RAND-36: Een handleiding. Noordelijk Centrum voor Gezondheidsvraagstukken, Groningen