The Effects of Dynamic Isolated Lumbar Extensor Training on Lumbar Multifidus Functional Cross-Sectional Area and Functional Status of Patients With Chronic Nonspecific Low Back Pain

Martin J. Willemink, MD,* Hendrik W. van Es, MD, PhD,* Pieter H. Helmhout, PhD,‡ Arjen L. Diederik, MD,* Johannes C. Kelder, MD,† and Johannes P. M. van Heesewijk, MD, PhD*

Study Design. A prospective single-arm trial.

Objective. To investigate whether dynamic isolated resistance training of global lumbar extensor muscles leads to changes in lumbar multifidus (LM) morphology in terms of cross-sectional muscle, and, if so, whether these changes are associated with observed changes in self-experienced functional status of chronic nonspecific low back pain (CNSLBP).

Summary of Background Data. LM morphology is associated with the recurrence of CNSLBP.

Methods. Sixteen male patients underwent a dynamic isolated resistance-training program for the lower back muscles of approximately 10 sessions in 12 weeks. In the next 12 weeks, frequency of training was tailored to the patients’ need. Participants underwent lumbar magnetic resonance imaging at baseline (T0), after 12 weeks (T12), and after 24 weeks (T24). Functional cross-sectional area was obtained by analyzing the magnetic resonance images. Functional status was assessed using the patient-specific functional scale, Roland-Morris disability questionnaire, and global perceived effect scale.

Results. Roland-Morris disability questionnaire and patient-specific functional scale scores showed significant and clinically relevant improvements between baseline and T12 with 44% and 39%, respectively. Between T12 and T24, these scores did not change significantly. Seven participants (44%) reported clinically relevant improvements in global perceived effect at T12. At T24, 1 more participant reported a relevant global perceived effect improvement, whereas 2 participants (13%) reported worsening of their condition. The magnetic resonance imaging analysis showed minor nonsignificant changes in functional cross-sectional area.

Conclusion. Our study shows that 10 weeks of dynamic isolated training of the lumbar extensors, once a week, leads to clinically relevant improvements in functional status of men with CNSLBP, without accompanying improvements in functional cross-sectional area of LM. These findings suggest that improvement in LM morphology is not a critical success factor in restoring functional status of patients with CNSLBP, at least in the short term (6 mo).

Key words: low back pain, lumbar multifidus, dynamic isolated resistance training, lumbar extensor muscles, MRI, functional cross-sectional area, functional status.

D espite the high incidence of chronic nonspecific low back pain (CNSLBP), its cause is poorly understood.1 A fair number of studies have indicated that the lumbar multifidus (LM) muscle plays an important role in CNSLBP.2 Fat infiltration of the atrophied LM is associated with spinal instability, hence playing a role in the recurrence of CNSLBP.1,3–7 Freeman et al8 describe an inhibitory feedback mechanism for localized LM atrophy, beginning with pain in the spine (possibly arising from facet joints or intervertebral discs), leading to reflex inhibition of the LM, ultimately leading to atrophy and fatty replacement of the muscle.

Literature supports the effectiveness of active reconditioning exercise in the treatment of CNSLBP, by reducing pain and improving function in activities of daily living.9,10 However, a number of questions regarding the method of its application and exact prescription still remain to be answered.10 Several previous investigations suggest that segmental muscle stabilization training, that is, directed at teaching patients to activate their LM and transversus abdominis muscles, is superior to “superficial” strengthening of global muscle groups (rectus abdominis, obliquus abdominis, erector spinae), both in improving LM morphology (e.g., muscle cross-sectional area, area of fatty infiltration [AFI]) and in restoring low back pain.1,4

Dynamic isolated lumbar extensor training, a progressive resistance training concept used in low back pain management,
is primarily aimed at controlled activation and strengthening of the global back muscles, respectively, by fixation of the pelvis and upper legs. Choi et al suggest, however, that isolated lumbar extensor training also has beneficial effects on the deeper “inner core” trunk muscles, including the LM. Spinal stability is enhanced by different back muscles, with the LM muscles accounting for more than two-thirds of the stiffness of the spine when in the neutral zone. Imaging studies of the LM have demonstrated pathological changes associated with CNSLBP. Therefore, it is especially interesting to analyze potential changes in LM morphology associated with specific back exercises.

The objective of this prospective study was to investigate whether dynamic isolated resistance training of global LM extensor muscles in CNSLBP patients leads to changes in LM morphology in terms of muscle cross-sectional area and AFI, and, if so, whether these changes are associated with observed changes in self-experienced functional status of CNSLBP.

MATERIALS AND METHODS

Study Design
At the start of the study (T₀), participants underwent baseline magnetic resonance imaging (MRI). After 12 weeks of dynamic training of the lower back (T₁₂), participants underwent a second MRI scan. Another 12 weeks later (T₂₄), a third MRI scan was done; in the meantime, training continued at a customized level. To quantify the morphology of the LM, muscle cross-sectional area and AFI were measured in the axial plane at 3 disc levels: L3–L4, L4–L5, and L5–S1. Concurrently with each MRI, patients filled in validated questionnaires concerning functional limitations caused by low back pain. Results of the scans and questionnaires were analyzed together.

Participants
Twenty consecutive patients with a clinical presentation of CNSLBP who presented during regular consulting hours to a physiotherapy practice were examined for eligibility in the study. All participants had nonspecific low back pain lasting for at least 12 weeks. In our hospital, women in their fertile age have to undergo a pregnancy test before MRI. To prevent this additional testing, only male patients were included. Other exclusion criteria were history of surgical lumbar intervention, age less than 30 years, myopathy, muscular dystrophy, spinal deformity, idiopathic scoliosis, vertebral fractures, congenital malformations, severe hernia in need of a surgical intervention, spondylolysis, osteoporosis, vertebral metastasis, and ankylosing spondylitis. Because participants followed a progressive resistance training program, additional exclusion criteria included angina pectoris, untreated hypertension, untreated diabetes mellitus, epilepsy, active rheumatic disease, severe neurological degenerative diseases like multiple sclerosis, and severe coagulation disorders. Besides, patients who attended a fitness training program for the low back muscles at the time of recruitment were excluded. MRI contraindications included claustrophobia, metal splinter in the eye, and medical or biostimulation implants. All participants gave informed consent. Approval for this study was obtained by the local ethical committee.

Intervention
Participants underwent progressive resistance training of the isolated lumbar extensors for approximately 24 weeks. The training program was carried out on the Lower Back Revival System (Figure 1, OriGENE Concepts BV, Delft, the Netherlands). This back training device prevents the lower extremities and hips to move during the exercise by locking the knees and thighs, thereby allowing only the individual's upper body to move. Participants were instructed to move in a relatively slow and controlled manner through full range lumbar motion (in approximately 2 seconds from flexion to extension, and in approximately 3 seconds back to flexion), thereby activating both global and deeper trunk muscles. Another principle of isolated lumbar extensor training is the relatively low number of training sessions prescribed for optimal results. The program included approximately 10 training sessions (once a wk) during the first 12 weeks. Thereafter, training continued at a frequency that was tailored to the patients’ convenience. This regime follows current practice-based training doctrines used in the Lower Back Revival System program. Previous studies on the concept have reported significant improvements in strength after training at a frequency of 2 days or even 1 day a week, comparable with improvements of 3 times a week.
The goal of each training session was to perform 4 sets of up to 10 repetitions, in which the patients had to move slowly and controlled from maximal flexion to maximal extension, and back to maximal flexion. The initial training load depended on the individual's strength and was selected by the physiotherapist. During the first session, the training load was set at a moderately low level and increased to a level at which the participant was able to perform 4 sets of each 10 repetitions in a comfortable way. The load for every next training session was increased by 2.5 kilograms if the participant was able to perform 4 sets in a comfortable way. The physiotherapist noted weight load and number of repetitions during each training session. All sessions were conducted by the same provider, whether a physiotherapist or a sports physician. During the study, participants were not allowed to attend other back muscle therapies. No medicine restrictions were imposed.

**Outcomes**

**Questionnaires**

To analyze low back specific functional status at each time point, patients completed 2 validated questionnaires: the Dutch 24-item version of the Roland-Morris disability questionnaire (RMDQ), and the patient-specific functional scale (PFS) described by Beurskens *et al.* Moreover, the global perceived effect (GPE) was determined. RMDQ scores range from 0 (no disability) to 24 (severe disability). PFS ranges from 0 (i.e., no disability in performing the 3 most disabling activities in daily living) to 300 (i.e., unable to perform these activities at all). A week before the planned MRI, patients received a letter intended to give them time to reflect on their most important limiting activities. GPE was determined by a 7-point Likert scale question concerning the patient’s global improvement or worsening (1 equals complete recovery, 7 equals much worse).

**Magnetic Resonance Imaging**

Imaging techniques such as ultrasound computed tomography (CT) and MRI allow analysis of LM morphology. The total cross-sectional area (TCSA) of muscles, including the LM, can be divided in functional cross-sectional area (FCSA), and AFI. The FCSA is the cross-sectional area of the muscle isolated from fat. Because significant atrophy of lumbar paraspinal muscles can occur without reduction of TCSA, LM morphology should be quantified by analyzing FCSA and AFI together.

The resolution of ultrasound is low, which leads to inferior tissue discrimination. CT and MRI result in higher contrast images and therefore allow good tissue discrimination. In a reliability study, Hu *et al.* found that CT and MRI are acceptable for measuring the FCSA and fatty infiltration of atrophied lumbar paraspinal muscles, with slightly favorable results for MRI. In the present study, MRI was preferred to CT because of its reliability results and the absence of additional radiation exposure.

The MRI system used in this study comprised a 1.5 Tesla MR unit (Philips Medical Systems, Best, the Netherlands). For transmission and reception, a spiral surface coil was used. Images were obtained using a balanced fast-field echo sequence, with matrix size 512 × 512, field of view...
of 5 patients underwent one extra MRI, 4 weeks before the baseline MRI (T−4). During these 4 weeks, patients underwent no training at all. Both the extra MR images at T−4 and the baseline MR images at T0 of these 5 patients were analyzed twice by 1 observer and once by the other. Measurements were repeated after 2 weeks by the first observer.

Intraobserver reliability was analyzed using the ICC with a 2-way random model (ICC 2,1); interobserver reliability with a 2-way mixed model (ICC 3,1). TCSA, FCSA, and AFI median (± standard deviation [SD]) ICC values were 0.94 (±0.12), 0.94 (±0.17), and 0.96 (±0.08), respectively, for intraobserver reliability, and 0.82 (±0.16), 0.80 (±0.30), and 0.92 (±0.04), respectively, for interobserver reliability.

Main Statistical Analysis
Mainly for reasons of costs and feasibility, the number of participants examined for eligibility in this study was limited to 20. A one-within multivariate and univariate analysis of variance for repeated measures was performed, respectively, to analyze temporal changes (0, 12, 24 wk) in outcomes. If necessary, post hoc testing of within-subjects contrasts was performed to determine statistically significant pairwise comparisons, using the Bonferroni corrections between levels of evaluation. Normality of the outcome variables at different levels of evaluation was tested with the Kolmogorov-Smirnov test. The
The results of this study are displayed in Table 1. RMDQ and PFS scores significantly improved between baseline and T12 by 43.9% and 38.5%, respectively. Between T12 and T24 these scores declined by 2.9% and 4.1%, respectively. Seven participants (43.8%) had GPE scores of 1 or 2 at T12 and no participant scored higher than 4. At T24, 50% had GPE scores of 1 or 2 (N=8) and 12.5% (N=2) scored higher than 4.

Between baseline and T12, patients underwent, on average, 9.8 (±0.93) training sessions compared with 4.69 (±2.5) sessions between T12 and T24. The average weight lifted increased from 19.1 (±3.2) kg at baseline to 36.4 (±4.1) kg at T12, with a minor increase of 1% thereafter to 36.9 (±5.1) kg at T24. For both RMDQ and PFS, tests of within-subjects contrasts were significant between T0 and T12, and not significant between T12 and T24. MR image analysis showed no significant changes in TCSA, AFI, and FCSA during both follow-up periods. Results are graphically shown in Figure 5.

**DISCUSSION**

This study shows that 24 weeks of dynamic resistance training of the isolated lumbar extensors led to improved functional status in a group of male patients with CNSLBP without significant changes in LM morphology.
TABLE 1. Changes in Functional Status and Multifidus Cross-Sectional Areas of Patients (N = 17) at Baseline (T₀), After 12 Weeks (T₁₂), and After 24 Weeks (T₂₄). Mean Values ± Standard Deviation are Presented

<table>
<thead>
<tr>
<th></th>
<th>T₀</th>
<th>T₁₂</th>
<th>T₂₄</th>
<th>T(dfₐ, dfₑ)*</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMDQ score (0–24)</td>
<td>7.35 ± 4.30</td>
<td>4.12 ± 3.02</td>
<td>4.24 ± 5.91</td>
<td>4.22 (2.0, 32.0)</td>
<td>0.024</td>
</tr>
<tr>
<td>PFS score (0–300)</td>
<td>179.94 ± 52.75</td>
<td>110.68 ± 72.40</td>
<td>115.24 ± 81.82</td>
<td>15.97 (2.0, 32.0)</td>
<td>0.000</td>
</tr>
<tr>
<td>L3–L4 TCSA (cm²)</td>
<td>7.04 ± 1.30</td>
<td>7.11 ± 1.36</td>
<td>6.83 ± 1.19</td>
<td>1.04 (1.16, 18.59)†</td>
<td>0.333</td>
</tr>
<tr>
<td>FCSA (cm²)</td>
<td>5.96 ± 1.19</td>
<td>6.08 ± 1.29</td>
<td>5.83 ± 1.04</td>
<td>1.10 (2.0, 32.0)</td>
<td>0.345</td>
</tr>
<tr>
<td>AFI (cm²)</td>
<td>1.08 ± 0.55</td>
<td>1.02 ± 0.52</td>
<td>1.01 ± 0.61</td>
<td>0.65 (1.35, 21.56)†</td>
<td>0.475</td>
</tr>
<tr>
<td>L4–L5 TCSA (cm²)</td>
<td>9.48 ± 1.70</td>
<td>9.61 ± 1.87</td>
<td>9.45 ± 1.79</td>
<td>1.60 (2.0, 32.0)</td>
<td>0.218</td>
</tr>
<tr>
<td>FCSA (cm²)</td>
<td>7.68 ± 1.29</td>
<td>7.82 ± 1.42</td>
<td>7.69 ± 1.37</td>
<td>1.32 (2.0, 32.0)</td>
<td>0.280</td>
</tr>
<tr>
<td>AFI (cm²)</td>
<td>1.80 ± 0.81</td>
<td>1.79 ± 1.08</td>
<td>1.76 ± 1.01</td>
<td>0.08 (1.17, 18.79)†</td>
<td>0.820</td>
</tr>
<tr>
<td>L5–S1 TCSA (cm²)</td>
<td>10.54 ± 1.34</td>
<td>10.59 ± 1.23</td>
<td>10.51 ± 1.45</td>
<td>0.28 (2.0, 32.0)</td>
<td>0.760</td>
</tr>
<tr>
<td>FCSA (cm²)</td>
<td>8.26 ± 1.32</td>
<td>8.44 ± 1.37</td>
<td>8.35 ± 1.55</td>
<td>1.49 (2.0, 32.0)</td>
<td>0.241</td>
</tr>
<tr>
<td>AFI (cm²)</td>
<td>2.27 ± 0.99</td>
<td>2.14 ± 0.98</td>
<td>2.17 ± 1.00</td>
<td>1.98 (2.0, 32.0)</td>
<td>0.155</td>
</tr>
<tr>
<td>GPE score (%)‡, §</td>
<td>1 or 2</td>
<td>7 (43.8%)</td>
<td>8 (50.0%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3–5</td>
<td>9 (56.3%)</td>
<td>6 (37.6%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5–7</td>
<td>0 (0%)</td>
<td>2 (12.5%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Numbers of training‡</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lifting weight (kg)‡</td>
<td>19.14 ± 3.19</td>
<td>36.41 ± 4.08</td>
<td>36.88 ± 5.12</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

AFI indicates area of fatty infiltration; FCSA, functional cross-sectional area; GPE, global perceived effect; PFS, patient-specific functional scale; RMDQ, Roland-Morris disability questionnaire; TCSA, total cross-sectional area.

*F represents the ratio of the repeated measures analysis of variance, with dfₐ representing the degrees of freedom of the model, and dfₑ representing the degrees of freedom for the residuals of the model.

†Greenhouse-Geisser corrected df values (assumption of sphericity not met).

‡ N = 16 instead of 17.

§ GPE score: 1, completely recovered; 2, much improved; 3, slightly improved; 4, no change; 5, slightly worsened; 6, much worsened; 7, worse than ever.

Between baseline and T₁₂, RMDQ scores significantly decreased with, on average, 3.2 points (44%). Scores on the PFS questionnaire significantly lowered with, on average, 69.3 points (39%). With literature presenting minimal clinically important difference values of 2.3 points using a 0 to 10 scale for PFS and corresponding to 23 points on a 0–100 scale, and 2 to 3 points for RMDQ with baseline values below 9, respectively, the differences seen between baseline and T₁₂ may be considered as clinically relevant. At T₂₄, 44% of the participants rated their complaints on the GPE scale as completely recovered or much improved over time, which are considered qualifications for clinically important improvements.

There were no significant changes in LM TCSA (0.5%–1.4%) and FCSA (1.8%–2.2%) during treatment. These findings are in line with a study by Danneels et al., who also reported no significant changes in FCSA (0%–1.9%) of the LM after 10 weeks of a combined stabilization and progressive resistance training regime for the trunk and lower leg muscles. Significant changes in FCSA (6.3%–7.3%) of the LM were only found in the intervention arm that added a static component to the training regime. Some studies did show significant increases in muscle size because of dynamic extension training, but TCSA values were assessed in compound muscle groups (erector spinae, multifidus, and psoas) and not in the LM alone, making direct comparison difficult.

Functional status scores and lifting weight increased predominantly during treatment between baseline and T₁₂. Thereafter, improvements stabilized, showing little or no increase or decrease, which suggests that this second training period, in which training frequency was customized to the participants’ convenience, merely served as a “maintenance” phase.

The small and insignificant increases in LM are unlikely to have caused the clinical improvements in self-assessed functional status seen in our study group of mostly chronically disabled patients. Several possible explanations can be given for these results. First, the dimensions of our treatment program, in terms of frequency (1 training session a week or less) and duration (approximately 30 min per session), were limited compared with previous studies on multifidus training in back pain patients. Training stimuli may not...
have specifically challenged the segmental multifidi muscles, or may simply have been too small and/or short to actually activate and change the morphology of the LM muscle. Because mean training weights of participants almost doubled in the first 12 weeks, it can be speculated that other training benefits emerge from our specific lumbar extensor training, such as joint mobilization, improved intra- or intermuscular coordination, strengthening of global movers, and/or decreased fear of movement. Different training frequencies and/or intensities may be indicated to further explore the impact of dynamic isolated lumbar extensor training on the segmental LM.

Second, despite good overall reliability results (FCSA intraobserver and interobserver reliability scores were excellent and good), the variance in the data concerning multifidus TCSA, FCSA, and AFI was quite considerable. Although the precision of this study may improve with a larger sample size, this would probably not have changed the main outcomes of this study, that is, clinically relevant improvements in functional status of the subjects and only marginal changes in LM morphology.

Third, participation in this study was voluntary, which may have attracted a group of patients with a more than average positive attitude toward the specific lumbar extensor exercises practiced. Because our study is noncontrolled, we cannot rule out that this potential recruitment bias may have led to the reported improvements in functional status. However, the fact that our study population had, on average, longstanding complaints (71% had back pain for ≥5 yr) emphasizes the clinical relevance of the observed improvements.

Because the current study assessed the short-term effects of one training modality, future research may address long-term effects using different training frequency/intensity groups and a control group to differentiate between training modalities for CNSLBP. A larger sample size could be used to evaluate differences between subgroups of participants with and without improvements in back complaints.

Conclusively, our study suggests that 10 weeks of dynamic isolated training of the lumbar extensors, once a week, leads to clinically relevant improvements in functional status of men with CNSLBP, without accompanying improvements in LM morphology. These findings suggest that improvement in LM morphology is not a critical success factor in restoring function of patients with CNSLBP, at least in the short term (6 mo).

Acknowledgments
The authors thank physiotherapist Mike Geurtsen for his commitment and Rene G. H. Willemink, PhD, for his contribution.

References

Key Points
- Specific training of global back muscles can lead to clinically relevant improvements in CNSLBP, without accompanying changes in LM FCSA.
- Improvement in LM morphology is not a critical success factor in restoring function of patients with chronic low back pain on the short term.
- However, future research concerning the long-term effects of different training frequencies and intensities is recommended.