Journal of Back and Musculoskeletal Rehabilitation 36 (2023) 799–813
DOI 10.3233/BMR-220220
IOS Press

Review Article

Effectiveness of non-surgical treatment combined with supervised exercise for lumbar spinal stenosis: A systematic review and meta-analysis

Ryunosuke Urata[a,1], Tatsuya Igawa[a,b,c,1,*], Shomaru Ito[d] and Akifumi Suzuki[e]
[a]*Department of Rehabilitation, International University of Health and Welfare Mita Hospital, Tokyo, Japan*
[b]*Department of Physical Therapy, School of Health Science, International University of Health and Welfare, Tochigi, Japan*
[c]*Department of Rehabilitation, International University of Health and Welfare Hospital, Tochigi, Japan*
[d]*Department of Rehabilitation, International University of Health and Welfare Narita Hospital, Chiba, Japan*
[e]*Toyota Tsusho All Life Corporation, Tokyo, Japan*

Received 12 July 2022
Accepted 13 December 2022

Abstract.
BACKGROUND: The benefits of combining supervised exercise in the non-surgical treatment of lumbar spinal stenosis (LSS) is unclear.
OBJECTIVE: To compare the effectiveness of non-surgical treatments with and without supervised exercise for pain intensity, symptom severity, functional impairment/disability, walking distance, and quality of life (QOL) in LSS patients.
METHODS: Randomized controlled trials (RCTs) evaluating combinations of supervised exercises were searched using four electronic databases up to August 13, 2020. Meta-analysis was conducted for immediate and long-term results.
RESULTS: Three studies were identified, including 244 participants. Immediate-term results showed that leg pain intensity (mean distance [MD]: -0.94, 95% confidence intervals [95% CI]: -1.60 to -0.29, $p < 0.01$) and symptom severity (MD: -0.29, 95% CI: -0.50 to -0.08, $p < 0.01$) were lower in the study group than in the control group, and walking distance (MD: 415.83, 95% CI: 298.15 to 533.50, $p < 0.001$) and QOL were higher in the study group. Long-term results showed that functional disability/impairment (MD: -0.27, 95% CI: -0.49 to -0.04, $p < 0.05$) was lower in the study group than in the control group, and walking distance and QOL were higher in the study group.
CONCLUSION: The number of studies on this topic was small and limited. Combinations of non-surgical treatment and supervised exercise may not provide significant benefits.

Keywords: Spinal stenosis, exercise therapy, systematic review as topic, meta-analysis as topic

[1]Tatsuya Igawa and Ryunosuke Urata contributed equally to this work.

[*]Corresponding author: Tatsuya Igawa, Department of Physical Therapy, School of Health Science, International University of Health and Welfare 2600-1, Kitakanemaru, Ohtawara, Tochigi, 323-8501, Japan. E-mail: igatatsu.7@gmail.com.

ISSN 1053-8127/$35.00

1. Introduction

Approximately 20% of adults over the age of 65 years suffer from lumbar spinal stenosis (LSS). Patients with LSS who have narrowed spinal canals or intervertebral foramina have lower limb pain and numbness, neurological claudication, and severely restricted physical activity [1]. The number of people with LSS tends to increase with age [2], and globally, it may increase further in the future, particularly in a rapidly aging community [3]. There is some evidence that exercise therapy for LSS is effective in improving pain and disability [4–8]. Although exercise therapy is generally prescribed before surgery [9], effective means of provision have not been established.

Supervised exercise is effective for conditions such as intermittent claudication, cancer-related fatigue, and hip osteoarthritis [10–12]. Supervision can properly improve adherence and exercise intensity due to the encouragement from and confidence of healthcare providers [11]. In recent reviews, supervised exercise has been suggested to be effective even in the non-surgical treatment of LSS [7,8]. The LSS clinical guidelines published by the Danish Health Authority in 2019 state that "supervised exercise is one of the recommended treatments" [7]. A previous review by Jacobi et al. [8] shows low-to-moderate evidence that a combination of supervised exercise and manual therapy was superior to exercise performed by patients independently in improving short-term walking capacity, pain intensity (leg and back pain), and symptom severity. Therefore, supervised exercise may be an effective approach for LSS patients.

Treatment selection for LSS is usually based on a combination of clinical evidence and individual patient characteristics and preferences [9]. In some cases, the two treatments may be combined. Long-term evidence for complex non-surgical treatments, including supervised exercise, is very important for clinicians and patients during treatment decisions. Although previous reviews have focused on supervised exercise versus other forms of exercise, the combination of supervised exercise with other non-surgical treatments has not been fully considered, and the synergistic effect of this combination is not clear.

We hypothesized that combining supervised exercise with the non-surgical treatment of LSS would result in superior immediate and long-term clinical outcomes. This is because the synergistic effects of supervised exercise and other treatments may benefit patients with LSS. Thus, the purpose of this study was to conduct a systematic review comparing non-surgical treatments with and without supervised exercise to clarify the immediate and long-term effectiveness of the combination.

2. Materials and methods

This systematic review was reported according to the PRISMA 2020 statement [14]. The revised flowchart in the PRISMA 2020 statement distinguishes between original and updated systematic reviews [13]. As our review is not an update of a previous review, we chose the search strategy of the original systematic review. The protocol of this systematic review was prospectively registered with PROSPERO (CRD42020199232) [14]. The study included literature that had already been published; thus, Institutional Review Board approval was waived.

2.1. Eligibility criteria

The eligibility criteria were as follows: (i) studies that included patients with LSS who were above 50 years of age and were diagnosed by medical history, physical and neurological examinations, and diagnostic imaging; (ii) studies that compared the combination of supervised exercise and non-surgical treatment (e.g., medications, injections, exercise instruction, and education) with the non-surgical treatment without supervised exercise; and (iii) Randomized controlled trials (RCTs) that were published in English after 1990 until August 13, 2020. A previous study defined supervised exercise as treatment sessions that occur at least twice a week and last for at least 6 weeks [15]. We included treatments that met at least 80% of the required frequency for supervised exercise (a total of at least 10 frequencies) in our review. Non-RCTs, cohorts, observational studies, case-control studies, case reports, and studies related to surgery were excluded.

2.2. Search strategy and study selection process

Records were identified using multiple electronic search databases, including MEDLINE via PubMed, Cochrane Central Register of Controlled Trials, Physiotherapy Evidence Database, and CINAHL (the search was completed on August 13, 2020). The search formula included words such as "spinal stenosis," "exercise," and "randomized controlled trial" (Table 1 and Appendix). The identified records were stored in End-

Table 1
Search strategy in MEDLINE via PubMed

MEDLINE via PubMed
(("Spinal stenosis"[MeSH Terms] OR "spinal stenosis"[ALL Fields] OR "spinal stenoses"[ALL Fields]) OR (("lateral recess"[ALL Fields] OR "foraminal"[ALL Fields]) AND ("pathologic constriction"[ALL Fields] OR "constriction, pathologic"[MeSH Terms] OR "stenosis"[ALL Fields])) OR (("lumbar"[ALL Fields] OR "lumbo"[ALL Fields] OR "lateral"[ALL Fields] OR "central"[ALL Fields] OR "foraminal"[ALL Fields]) AND ("spinal stenosis"[ALL Fields] OR "spinal stenoses"[ALL Fields] OR "canal stenosis"[ALL Fields] OR "canal stenoses"[ALL Fields] OR "stenosis"[ALL Fields] OR "stenoses"[ALL Fields]))) AND ("transcutaneous electric nerve stimulation"[MeSH Terms] OR "pain management"[MeSH Terms] OR "rehabilitation"[MeSH Terms] OR "pain management"[ALL Fields] OR "rehabilitation"[ALL Fields] OR "physical therap*"[ALL Fields] OR "physiotherap*"[ALL Fields] OR "exercise"[MeSH Terms] OR "exercis*"[ALL Fields] OR "Physical Fitness"[ALL Fields] OR "Physical Conditioning"[All Fields] OR "Physical Activit*"[All Fields] OR ("Physical"[All Fields] AND ("Conditioning"[All Fields] OR "activit*"[All Fields])) OR "Training"[ALL Fields] OR "Gymnastic*"[ALL] OR "Walking"[ALL Fields] OR "Ambulation"[ALL Fields] OR "health education"[MeSH Terms] OR "Health education"[ALL Fields] OR ("Health"[ALL Fields] AND "education"[ALL Fields]) OR "education"[ALL FIelds]) AND "English"[Language] AND ("randomized controlled trial"[Publication type] OR "randomized controlled trial"[Title/Abstract] OR "randomized controlled trials"[Title/Abstract])

note Ver. X9 (USACO, Tokyo, Japan), and duplicate papers were deleted.

Two independent researchers performed a two-step screening process to rigorously assess the eligibility of the study. For the primary screening, the titles and abstracts were evaluated, and for the secondary screening, the full texts were read in detail. Disagreements were first discussed by two researchers, and if they were not resolved, a third researcher was invited to participate in the discussion.

2.3. Data collection process

Two independent researchers extracted data on study characteristics, participants, interventions for study and comparator groups, outcomes, and methodology. We used pain intensity (numerical rating scale [NRS]), symptom severity (symptom severity domain of the Zurich claudication questionnaire and the Swiss spinal stenosis questionnaire [ZCQS]), and functional impairment/disability (Oswestry disability index [ODI], and physical function domain of the Zurich claudication questionnaire and the Swiss spinal stenosis [ZCQF]) as the primary outcomes of this review. Secondary outcomes were self-reported questionnaires on health-related quality of life (QOL), walking distance or speed, and adverse events. Each outcome was categorized as immediate (less than 1 month after treatment) or long-term ($>$ 1 year after randomization) outcome.

2.4. Risk of bias assessment

The Cochrane risk of bias tool version 2.0 was used to assess the risk of bias in each study [16]. Two independent researchers evaluated the randomization process, deviations from intended interventions, missing outcome data, outcome measurement, and selection of the reported results, and the overall risk of bias was rated on a 3-point scale ("Low," "Some concerns," and "High"). Disagreements were first discussed by two researchers, and if not resolved, a third researcher was invited to participate in the discussion.

2.5. Data synthesis and analysis

Effective exercise modalities for LSS have not been established [9]. Supervised exercise programs were different across the studies; therefore, we performed a meta-analysis with a random-effects model. Mean difference (MD) or standardized MD (SMD) and 95% confidence interval (CI) were calculated for the prespecified outcomes. The significance levels were set at 5%. For statistical heterogeneity, the chi-square test and the I^2 value were calculated, and the significance level of the chi-square test was set at 10%. I^2 values were interpreted as follows: 0%, nonsignificant; 30%–60%, moderate heterogeneity; 50%–90%, substantial heterogeneity; and 75%–100%, significant heterogeneity [17]. If heterogeneity was detected, subgroup analyses based on the number of exercises, duration of intervention, or frequency of intervention were performed. Publication bias was visually evaluated using a funnel plot. Statistical analysis was performed using Review Manager Ver. 5.4 software (Cochrane Collaboration, Software Update, Oxford, UK).

The certainty of evidence in each outcome was assessed according to the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) methodology framework [18]. Three researchers scrutinized the following items: risk of bias, inconsistency, indirectness, imprecision, and publication bias, and they then rated the certainty of the total evidence on a scale ("High," "Moderate," "Low," and "Very low"). Downgrading criteria were defined as follows:

Risk of bias: Downgrade was considered if "Some concern" or "High" was included in the overall risk of bias integrated.

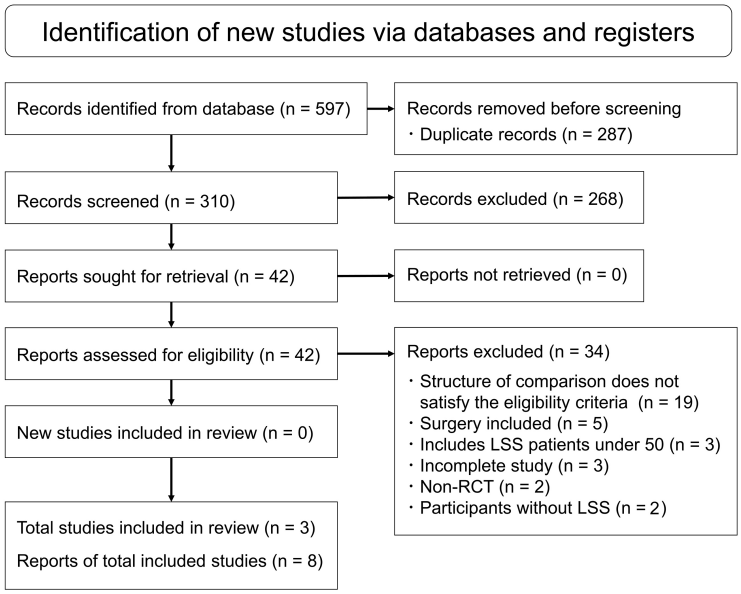


Fig. 1. The PRISMA flowchart of the study selection process.

Inconsistency: Downgrade was considered when the similarity of point estimates, overlapping confidence intervals, χ^2 test, and I^2 values were evaluated, and there was heterogeneity for which no cause could be identified.

Indirectness: Downgrade was considered if there were substantial differences between what was considered in the systematic review and the population, interventions, outcomes, and controls measured in the integrated study.

Imprecision: Downgrade was considered if the sample size was small and there was a wide range of effect estimates. The sample size for complement accuracy was based on an optimal information size (OIS) of 400 people. In addition, the downgrade was considered if the lower limit of the 95% CI for the estimated effect was below the minimal clinically important difference (MCID). The outcomes of MCID application were NRS leg pain, back pain, ZCQS, and ZCQF, with thresholds of 1.5, 1.25, 0.36, and 0.1, respectively [19].

Publication bias: Downgrade was considered when publication bias was suspected.

3. Results

3.1. Study selection and study characteristics

A total of 597 records were identified from the electronic databases. Eventually, three studies were included in this systematic review [20–22] (Fig. 1). Detailed characteristics of each study are presented in Table 2. The three included studies published in 2018 and 2019 were performed in Canada, the United States, and Japan, respectively. A total of 244 participants (110 men and 134 women), with a high proportion of females, met the eligibility criteria for this study. The mean age of participants in all the studies was 70.2 years. In all the studies, the participants were classified into two groups, non-surgical treatment with and without supervised exercise. In all three studies, the supervised exercise for the study group included manual therapy, stretching, muscle endurance and stabilization exercises, cycling, and weight-supported treadmill training. These were provided individually by a physical therapist or chiropractor, and the frequency of treatment met the eligibility criteria (a total of 10–12 frequencies). In two studies, a voluntary training program, including daily walking and home exercise, was provided to both the study and control groups [20,22]. In one study, epidural steroid injections (ESI) and patient education were provided to both groups [21]. The included studies were homogeneous; therefore, the meta-analysis included all the studies.

3.2. Risk of bias

The overall risk of bias in all the studies was a concern (Fig. 2). They were rated as high risk in outcome

Table 2
Characteristics of the included studies

Study	Country	Sample size			Age mean (SD)			Sex male: female		Intervention		Analyzable outcomes
		All	SG	CG	All	SG	CG	SG	CG	SG	CG	
Ammendolia et al., 2018	Canada	104	51	53	70.6	69.4 (7.7)	71.7 (9.5)	18: 33	27: 26	**Comprehensive group** 1. Structured comprehensive conservative training program (twice a week): self-management strategies using a cognitive-behavioral approach; a standardized set of 18 exercises; manual therapy. 2. Self-directed training program (every day): same as the control group. Duration of intervention: 6 weeks	**Self-directed group** 1. Self-directed training program (every day): instructional videos, workbooks, and pedometers on how to perform the required exercises and self-management strategies; 15- to a 30-minute training session with an experienced independent licensed chiropractor (only once). Duration of intervention: 6 weeks	ZCQS, ZCQF, ODI, and NRS (leg and back pain), SF-36 Follow-up: 8 weeks and 3, 6 and 12 months
Hammerich et al., 2019	United States	54	23	31	67.2 (9.7)	66.3 (1.9)	67.8 (1.8)	11: 12	15: 16	**ESI + PT group** 1. Evidence-based multimodal PT program (once a week): manual physical therapy; mobility; aerobic exercise; muscle endurance and stabilization exercises. 2. One to 3 ESIs are performed according to standardized algorithms. 3. Education by the Back Book. Duration of intervention: 10 weeks	**ESI group** 1. One to 3 ESIs are performed according to standardized algorithms. 2. Education by the Back Book. Duration of intervention: 10 weeks	ODI, NPRS, and SF-36 Follow-up: 10 weeks and 6 and 12 months
Minetama et al., 2019	Japan	86	43	43	72.7	72.3 (6.9)	73.2 (8.2)	20: 23	19: 24	**PT group** 1. Supervised PT sessions (twice a week): manual therapy; individually tailored stretching and strengthening exercises; cycling and body weight-supported treadmill walking. 2. Daily walk 3. HE program (every day) Duration of intervention: 6 weeks	**HE group** 1. Daily walk 2. HE program (every day): three 30-second bouts of both single and double knee-to-chest exercises; ten 6-second bouts of trunk raise and bridging in the supine position; a four-point kneeling exercise. Duration of intervention: 6 weeks	ZCQS, ZCQF, SPWT, NRS (leg and back pain), and SF-36 Follow-up: 6 weeks

Abbreviation. SG, study group; CG, control group; SD, standard deviation; ESI, epidural steroid injection; PT, physical therapy; HE, home exercise; ZCQS, symptom severity domain of the Zurich claudication questionnaire; ZCQF, physical function domain of the Zurich claudication questionnaire; ODI, Oswestry Disability Index; NRS or NPRS, numerical (pain) rating scale; SPWT, self-paced walking test; SF-36, MOS 36-Item Short-Form Health Survey.

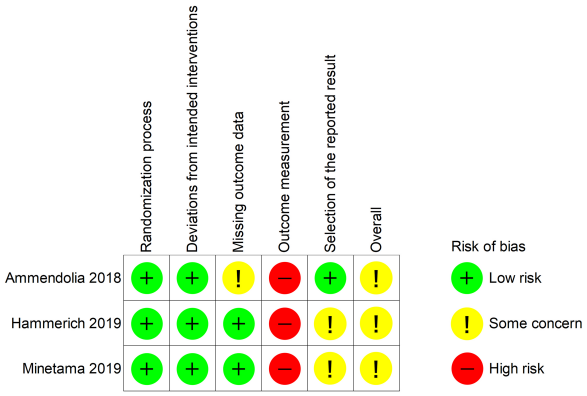


Fig. 2. Summary of risk of bias.

measurements because of inadequate blinding of therapists or participants.

3.3. Results of each predetermined outcome

3.3.1. Pain intensity

The three included studies reported pain intensity as assessed by the NRS; two of them reported leg pain intensity [20,22], and the three studies reported back pain intensity [20–22]. Synthesis of results was performed separately for leg pain and low back pain. Of the two studies that reported leg pain intensity, one reported only immediate results [22], and one reported both immediate and long-term results [20]. The meta-analysis performed to assess immediate effects showed that the leg pain intensity of the study group improved significantly compared to that of the control group (n = 188, MD [95% CI] = −0.94 [−1.60 to −0.29], p = 0.005) (Fig. 3). One study that combined supervised exercise with a self-directed training program [20] found no significant difference between the study and control groups in long-term results (n = 104, MD [95% CI] = −0.50 [−1.60 to 0.60], p = 0.37). All three studies that reported back pain intensity reported immediate results [20–22], and two reported long-term results [20,22]. Back pain intensity showed no significant difference between the study and control groups in immediate (n = 242, MD [95% CI] = −0.34 [−2.08 to 1.40], p = 0.70) and long-term results (n = 158, MD [95%CI] = 0.97 [−1.81 to 3.75], p = 0.49) (Fig. 3).

3.3.2. Symptom severity

Symptom severity, as assessed by the ZCQS, was reported in two studies [20,22]. The two reported immediate results [20,22], and one reported long-term results [20]. The meta-analysis performed to assess immediate results showed that the study group improved significantly compared to the control group (n = 188, MD [95% CI] = −0.29 [−0.50 to −0.08], p = 0.006) (Fig. 4). The study that combined supervised exercise with a self-directed training program [20] showed no significant difference between the study and control groups in the long-term results (n = 104, MD [95% CI] = −0.22 [−0.47 to 0.02], p = 0.08).

3.3.3. Functional impairment/disability

Of the studies that assessed functional impairment/disability, two used the ZCQF [20,22], and two used the ODI [20,21]. Synthesis of results was performed separately for ZCQF and ODI. The two studies that reported ZCQF reported immediate results [20,22], and one reported long-term results [20]. The meta-analysis performed to assess immediate effects showed no significant difference between the study and control groups (n = 188, MD [95% CI] = −0.21 [−0.58 to 0.16], p = 0.27) (Fig. 5). Regarding long-term results, the group that was prescribed both supervised exercise and the self-directed training program improved significantly compared to the group that was prescribed the self-directed training program alone (n = 104, MD [95% CI] = −0.27 [−0.49 to −0.04], p = 0.02) [20]. Regarding the ODI, immediate and long-term results were reported in two studies [20,21]. The meta-analysis showed no significant difference between the study and control groups (immediate: n = 158, SMD [95% CI] = −0.07 [−0.39 to 0.24], p = 0.65; long-term: n = 158, SMD [95% CI] = −0.05 [−0.47 to 0.36], p = 0.80) (Fig. 5).

3.3.4. Walking distance

Walking distance, as assessed by the self-paced walking test (SPWT), was reported in two studies [20,22]. Two of them reported immediate results [20,22], and one reported long-term results [20]. The meta-analysis performed to assess immediate results showed that the study group improved significantly compared to the control group (n = 188, MD [95% CI] = 415.83 [298.15 to 533.50], $p < 0.00001$) (Fig. 6). Regarding long-term results, the self-directed training program with supervised exercise group improved significantly compared to the self-directed training program alone group (n = 104, MD [95% CI] = 473.20 [203.90 to 742.40], $p < 0.0006$) [21].

3.3.5. Quality of life

The QOL, as assessed by the 36-Item Short-Form Health Survey (SF-36), was reported in three stud-

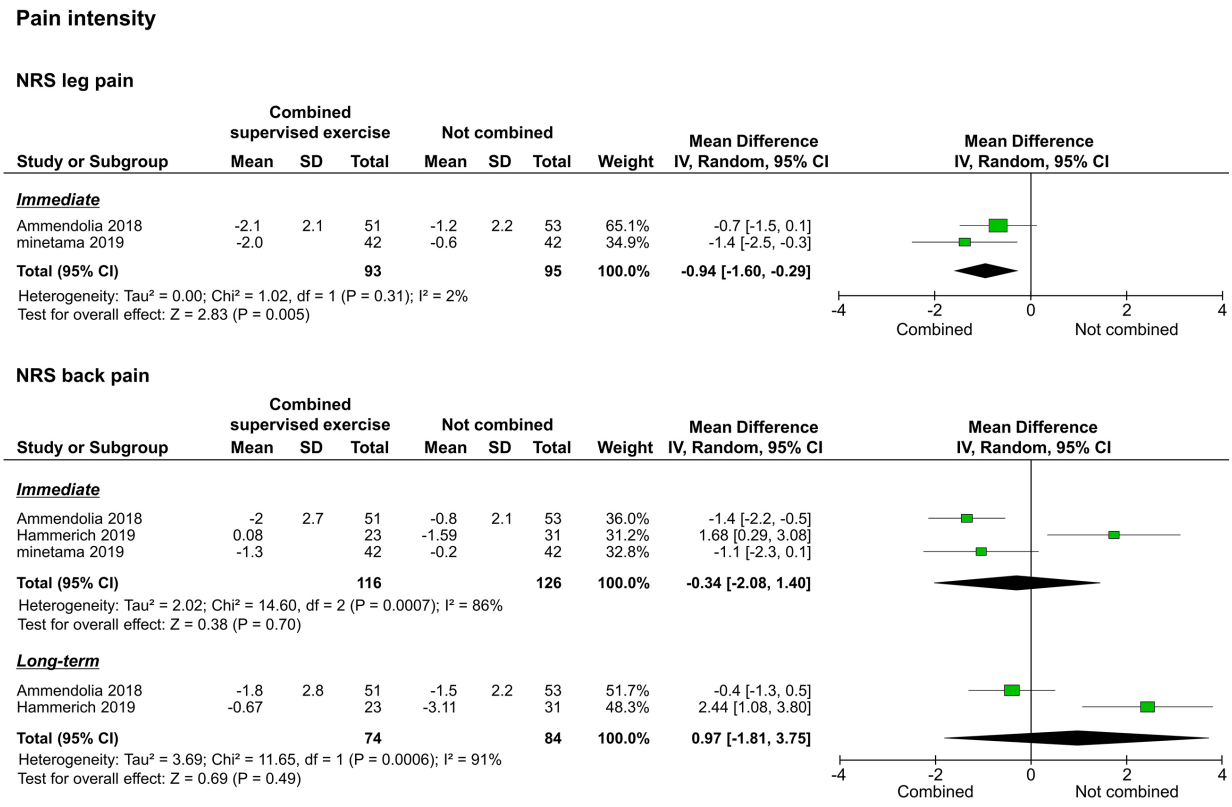


Fig. 3. Forest plots of pain intensity. Abbreviation. NRS, numerical rating scale.

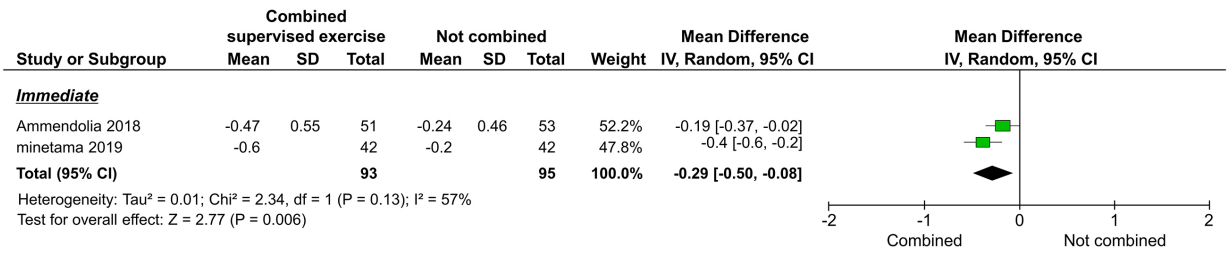


Fig. 4. Forest plot of symptom severity. Abbreviation. ZCQS, symptom severity domain of the Zurich claudication questionnaire.

ies [20–22]. We were able to pool the subscale results for all the studies. However, the meta-analyses performed to assess the long-term results of physical functioning (PF), role-physical, bodily pain (BP), vitality (VT), and social functioning (SF) were not performed because they were reported in only one study [20]. Regarding immediate results, PF ($n = 188$, MD [95% CI] $= 7.03$ [1.69 to 12.37], $p = 0.010$), general health (GH) ($n = 242$, MD [95% CI] $= 7.42$ [4.07 to 10.78], $p < 0.0001$), VT ($n = 188$, MD [95% CI] $= 6.61$ [1.17 to 12.05], $p = 0.02$), SF ($n = 188$, MD [95% CI] $= 6.02$ [0.06 to 11.99], $p < 0.05$), role-emotional (RE) ($n = 242$, MD [95% CI] $= 12.\ 03$ [0.79 to 23.27], $p = 0.04$), and mental health (MH) ($n = 242$, MD [95% CI] $= 6.40$ [1.76 to 11.04], $p = 0.007$) were significantly improved in the study group than in the control group (Figs 7 and 8). Regarding long-term results, MH ($n = 158$, MD [95% CI] $= 4.98$ [0.71 to 9.26], $p = 0.02$) was significantly improved in the study group than in the control group (Fig. 8). In addition, for the long-term

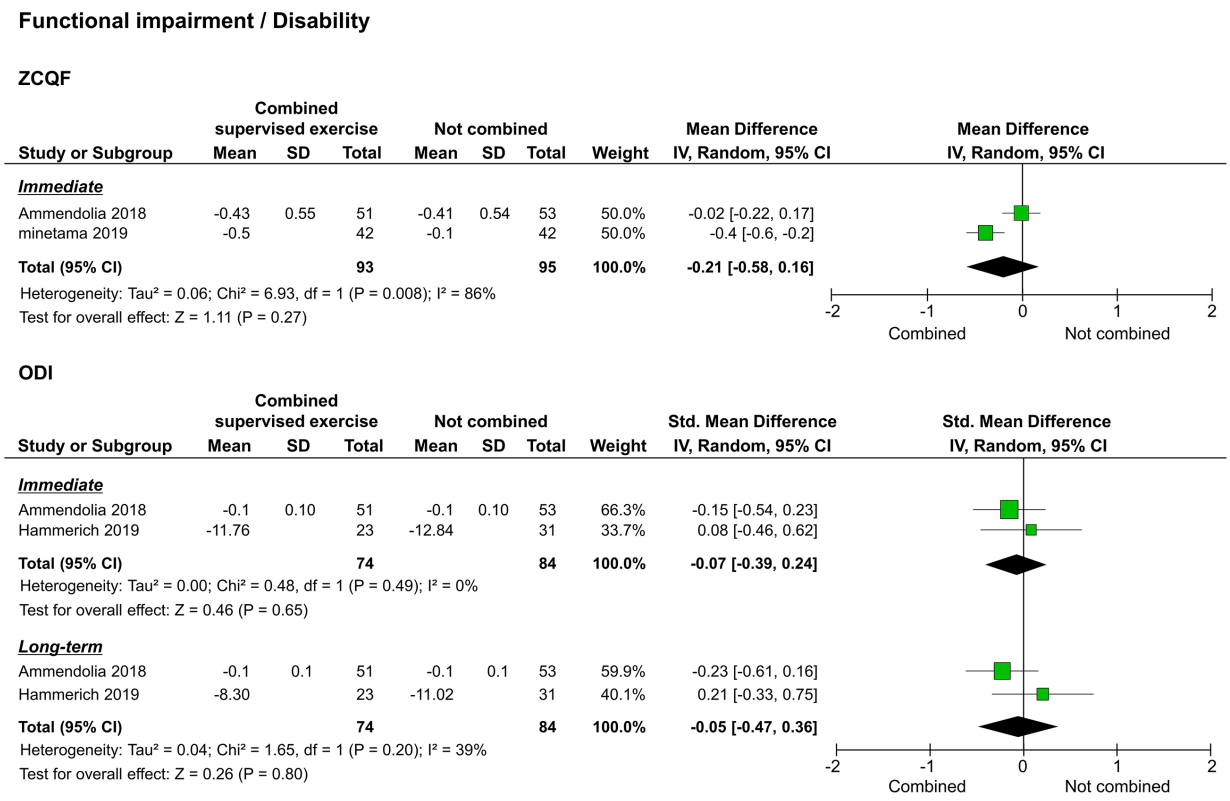


Fig. 5. Forest plots of functional impairment/disability. Abbreviation. ZCQF, physical function domain of the Zurich claudication questionnaire; ODI: Oswestry disability index.

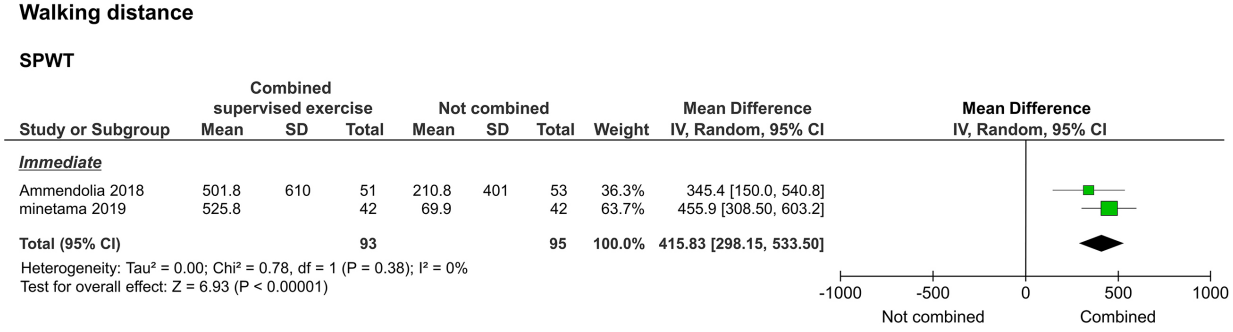


Fig. 6. Forest plot of walking distance. Abbreviation. SPWT, self-paced walking test.

results of PF ($n = 104$, MD [95% CI] = 8.20 [0.20 to 16.20], $p = 0.04$) and BP ($n = 104$, MD [95% CI] = 10.00 [2.10 to 17.90], $p = 0.01$), the group that underwent the self-directed training program combined with supervised exercise improved significantly compared to the group that underwent self-directed training program alone [20].

3.3.6. Adverse events

Adverse events were reported in two studies [20, 21]. One of them reported adverse events related to supervised exercise therapy [20]. Specifically, back pain exacerbation ($n = 5/51$) and knee, hip, and ankle joint pain exacerbation ($n = 4/51$) were reported during the 8-week intervention period.

3.4. Certainty of evidence

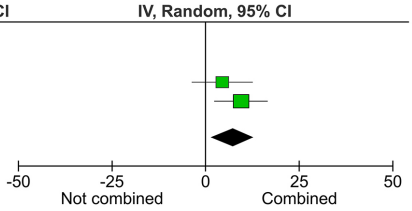
Results of the certainty of evidence are presented in Tables 3 and 4. For the item of risk of bias, the patient-

Quality of life

SF-36 Physical functioning

Study or Subgroup	Combined supervised exercise Mean	SD	Total	Not combined Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Immediate								
Ammendolia 2018	19.9	22.3	51	14.9	21.8	53	43.4%	4.2 [-3.9, 12.4]
minetama 2019	11.2		42	2.0		42	56.6%	9.2 [2.1, 16.3]
Total (95% CI)			**93**			**95**	**100.0%**	**7.03 [1.69, 12.37]**

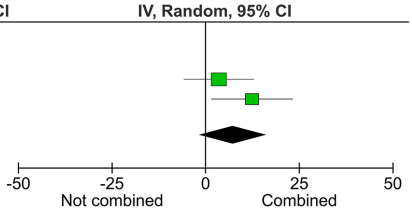
Heterogeneity: Tau² = 0.00; Chi² = 0.83, df = 1 (P = 0.36); I² = 0%
Test for overall effect: Z = 2.58 (P = 0.010)



SF-36 Role physical

Study or Subgroup	Combined supervised exercise Mean	SD	Total	Not combined Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Immediate								
Ammendolia 2018	12.9	26.3	51	10.8	30.2	53	55.3%	3.3 [-6.0, 12.5]
minetama 2019	17.4		42	5.2		42	44.7%	12.2 [1.3, 23.1]
Total (95% CI)			**93**			**95**	**100.0%**	**7.28 [-1.40, 15.95]**

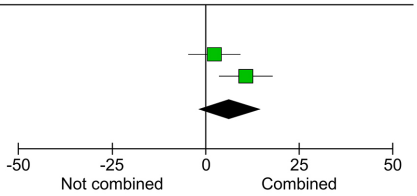
Heterogeneity: Tau² = 12.88; Chi² = 1.48, df = 1 (P = 0.22); I² = 33%
Test for overall effect: Z = 1.64 (P = 0.10)



SF-36 Bodily pain

Study or Subgroup	Combined supervised exercise Mean	SD	Total	Not combined Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Immediate								
Ammendolia 2018	14.4	17.8	51	9.7	20.0	53	50.5%	2.0 [-4.9, 8.9]
minetama 2019	15.8		42	5.4		42	49.5%	10.40 [3.3, 17.5]
Total (95% CI)			**93**			**95**	**100.0%**	**6.16 [-2.07, 14.39]**

Heterogeneity: Tau² = 22.52; Chi² = 2.77, df = 1 (P = 0.10); I² = 64%
Test for overall effect: Z = 1.47 (P = 0.14)



SF-36 General health

Study or Subgroup	Combined supervised exercise Mean	SD	Total	Not combined Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Immediate								
Ammendolia 2018	4.7	13.8	51	0.3	17.6	53	34.6%	6.0 [0.3, 11.7]
Hammerich 2019	8.27		23	-0.72		31	16.7%	8.99 [0.78, 17.20]
minetama 2019	8.9		42	1.0		42	48.8%	7.9 [3.1, 12.7]
Total (95% CI)			**116**			**126**	**100.0%**	**7.42 [4.07, 10.78]**

Heterogeneity: Tau² = 0.00; Chi² = 0.42, df = 2 (P = 0.81); I² = 0%
Test for overall effect: Z = 4.34 (P < 0.0001)

Long-term								
Ammendolia 2018	4.0	15.8	51	3.0	15.3	53	65.2%	2.1 [-3.9, 8.0]
Hammerich 2019	2.98		23	-2.12		31	34.8%	5.10 [-3.11, 13.31]
Total (95% CI)			**74**			**84**	**100.0%**	**3.14 [-1.70, 7.99]**

Heterogeneity: Tau² = 0.00; Chi² = 0.33, df = 1 (P = 0.56); I² = 0%
Test for overall effect: Z = 1.27 (P = 0.20)

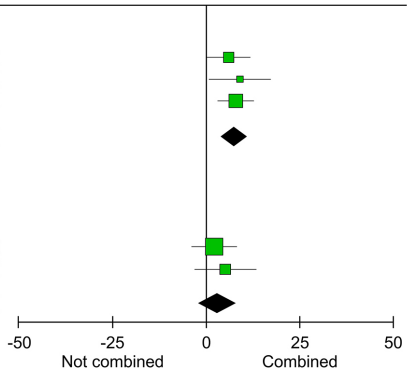
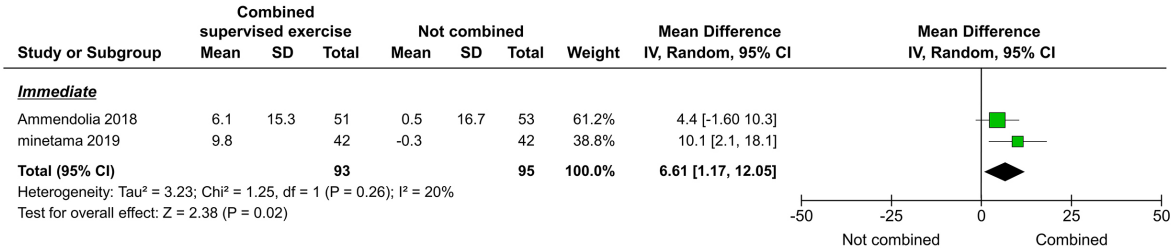
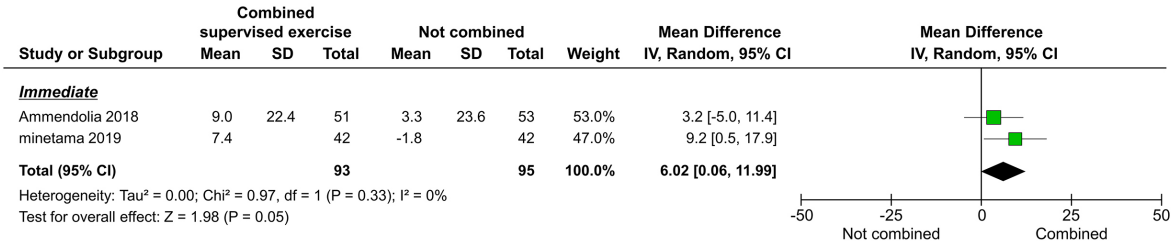


Fig. 7. Forest plots of QOL. Abbreviation. SF-36, 36-Item Short-Form Health Survey.

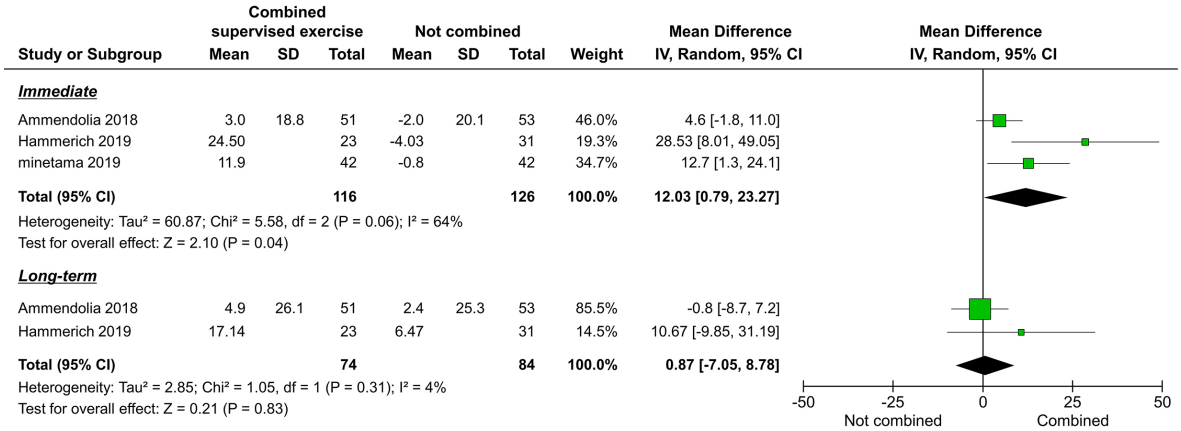
SF-36 Vitality



SF-36 Social functioning



SF-36 Role emotional



SF-36 Mental health

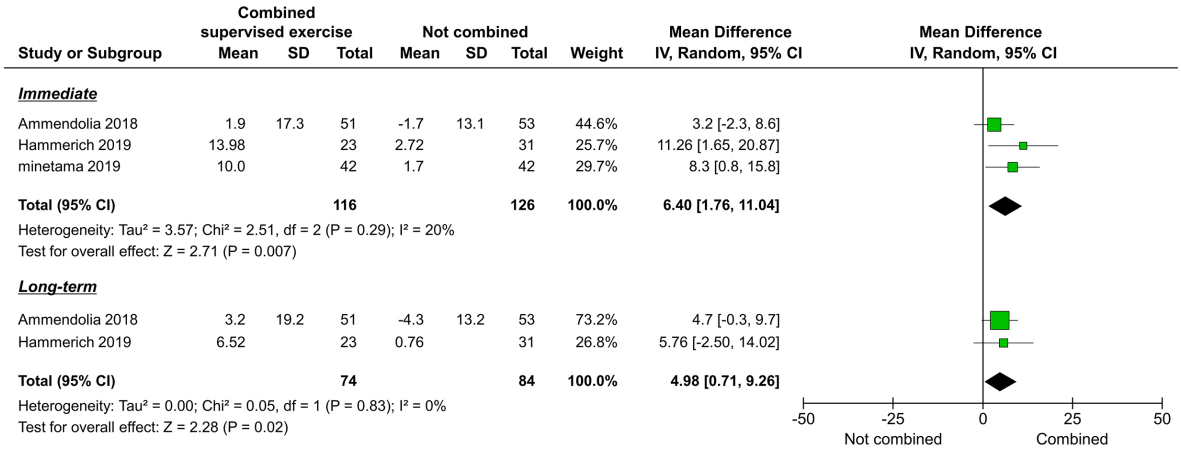


Fig. 8. Continuation of Fig. 7.

Table 3
Summary of findings of certainty of the evidence for each outcome (Immediate)

	No. of studies	No. of participants	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Effect (95% CI)	Certainty
Pain intensity									
NRS leg pain	2	188	Serious[a]	Not serious	Not serious	Very serious[d,e]	Strongly suspected[g]	MD **−0.94** (−1.60 to −0.29)	VERY LOW
NRS back pain	3	242	Serious[a]	Very serious[c]	Not serious	Very serious[d,f]	Strongly suspected[g]	MD **−0.34** (−2.08 to 1.40)	VERY LOW
Symptom severity									
ZCQS	2	188	Serious[a]	Serious[b]	Not serious	Very serious[d,e]	Strongly suspected[g]	MD **−0.29** (−0.50 to −0.08)	VERY LOW
Functional impairment/disability									
ZCQF	2	188	Serious[a]	Very serious[c]	Not serious	Very serious[d,f]	Strongly suspected[g]	MD **−0.21** (−0.58 to 0.16)	VERY LOW
ODI	2	158	Serious[a]	Not serious	Not serious	Very serious[d,f]	Strongly suspected[g]	SMD **−0.07** (−0.39 to 0.24)	VERY LOW
Walking distance									
SPWT	2	188	Not serious	Not serious	Not serious	Serious[d]	Strongly suspected[g]	MD **415.83** (298.15 to 533.50)	LOW
Quality of life									
SF-36									
Physical function	2	188	Serious[a]	Not serious	Not serious	Very serious[d,f]	Strongly suspected[g]	MD **7.03** (1.69 to 12.37)	VERY LOW
Role physical	2	188	Serious[a]	Not serious	Not serious	Very serious[d,f]	Strongly suspected[g]	MD **7.28** (−1.40 to 15.95)	VERY LOW
Bodily pain	2	188	Serious[a]	Very serious[c]	Not serious	Very serious[d,f]	strongly suspected[g]	MD **6.16** (−2.07 to 14.39)	VERY LOW
General health	3	242	Serious[a]	Not serious	Not serious	Very serious[d,f]	Strongly suspected[g]	MD **7.42** (4.07 to 10.78)	VERY LOW
Vitality	2	188	Serious[a]	Not serious	Not serious	Very serious[d,f]	Strongly suspected[g]	MD **6.61** (1.17 to 12.05)	VERY LOW
Social functioning	2	188	Serious[a]	Not serious	Not serious	Very serious[d,f]	Strongly suspected[g]	MD **6.02** (0.06 to 11.99)	VERY LOW
Role emotional	3	242	Serious[a]	Very serious[c]	Not serious	Very serious[d,f]	Strongly suspected[g]	MD **12.03** (0.79 to 23.27)	VERY LOW
Mental health	3	242	Serious[a]	Not serious	Not serious	Very serious[d,f]	Strongly suspected[g]	MD **6.40** (1.76 to 11.04)	VERY LOW

a, Risk of bias of some concerns; b, Unexplained moderate heterogeneity ($I^2 = 30$–60%); c, Unexplained substantive heterogeneity ($I^2 = 50$–90%); d, Small sample size ($n < 400$); e, Lower limit of 95% CI less than threshold (MCID); f, Confidence interval with no effect or very small effect estimated; g, Possibility of publication bias. Abbreviation. RCT, randomized controlled trial; CI, confidence interval; MD, mean difference; SMD, standardized mean difference.

Table 4
Summary of findings of certainty of the evidence for each outcome (long term)

	No. of studies	No. of participants	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Effect (95% CI)	Certainty
Pain intensity									
NRS leg pain	1	104	Serious[a]	Not serious	Not serious	Very serious[d,e]	Strongly suspected[f]	MD **−0.50** (−1.60 to 0.60)	VERY LOW
NRS back pain	2	158	Serious[a]	Very serious[c]	Not serious	Very serious[d,e]	Strongly suspected[f]	MD **0.97** (−1.81 to 3.75)	VERY LOW
Symptom severity									
ZCQS	1	104	Serious[a]	Not serious	Not serious	Very serious[d,e]	Strongly suspected[f]	MD **−0.22** (−0.47 to 0.02)	VERY LOW
Functional impairment/disability									
ZCQF	1	104	Serious[a]	Not serious	Not serious	Very serious[d,e]	Strongly suspected[f]	MD **−0.27** (−0.49 to −0.04)	VERY LOW
ODI	2	158	Serious[a]	Serious[b]	Not serious	Very serious[d,e]	Strongly suspected[f]	SMD **−0.05** (−0.47 to 0.36)	VERY LOW
Walking distance									
SPWT	1	104	Not serious	Not serious	Not serious	Serious[d]	Strongly suspected[f]	MD **473.20** (203.90 to 742.40)	LOW
Quality of life									
SF-36									
Physical functioning	1	104	Serious[a]	Not serious	Not serious	Very serious[d,e]	Strongly suspected[f]	MD **8.20** (0.20 to 16.20)	VERY LOW
Role physical	1	104	Serious[a]	Not serious	Not serious	Very serious[d,e]	Strongly suspected[f]	MD **5.20** (−4.90 to 15.40)	VERY LOW
Bodily pain	1	104	Serious[a]	Not serious	Not serious	Very serious[d,e]	Strongly suspected[f]	MD **10.00** (2.10 to 17.90)	VERY LOW
General health	2	158	Serious[a]	Not serious	Not serious	Very serious[d,e]	Strongly suspected[f]	MD **3.14** (−1.70 to 7.99)	VERY LOW
Vitality	1	104	Serious[a]	Not serious	Not serious	Very serious[d,e]	Strongly suspected[f]	MD **1.20** (−4.10 to 6.50)	VERY LOW
Social functioning	1	104	Serious[a]	Not serious	Not serious	Very serious[d,e]	Strongly suspected[f]	MD **1.20** (−5.90 to 8.40)	VERY LOW
Role emotional	2	158	Serious[a]	Not serious	Not serious	Very serious[d,e]	Strongly suspected[f]	MD **0.87** (−7.05 to 8.78)	VERY LOW
Mental health	2	158	Serious[a]	Not serious	Not serious	Very serious[d,e]	Strongly suspected[f]	MD **4.98** (0.71 to 9.26)	VERY LOW

a, Risk of bias of some concerns; b, Unexplained moderate heterogeneity (I^2 = 30–60%); c, Unexplainable significant heterogeneity (I^2 = 75–100%); d, Small sample size ($n < 400$); e, Confidence interval with no effect or very small effect estimated; f, Possibility of publication bias. Abbreviation. RCT, randomized controlled trial; CI, confidence interval; MD, mean difference; SMD, standardized mean difference.

reported outcomes were downgraded due to some concerns. For the imprecision, all evidence was downgraded due to the sample size being less than the OIS. Furthermore, the immediate results of NRS leg pain and ZCQS were further downgraded because their effect estimates did not exceed the MCID. Regarding publication bias, the possibility could not be ruled out, and thus all evidence was downgraded. Eventually, NRS leg pain, NRS back pain, ZCQS, ZCQF, ODI, and SF-36 showed very low evidence in both immediate and long-term results. In addition, SPWT showed low evidence for both immediate and long-term results.

4. Discussion

Our systematic review evaluated RCTs comparing non-surgical treatment with and without supervised exercise for LSS. To the best of our knowledge, this is the first systematic review evaluating the benefits of combining supervised exercise therapy with non-surgical treatment for LSS. According to our review, the combination of supervised exercise had better immediate results for NRS leg pain, ZCQS, ZCQF, SPWT, and SF-36 (PF, GH, VT, SF, RE, and MH) and long-term results for ZCQF, SPWT, and SF-36 (PF, BP, and MH) than that had by non-combined treatment. However, the certainty of all the evidence was low, and the strength of the effect was small. This result was contrary to our hypothesis.

This review identified two studies that reported that a combination of supervised exercise and a voluntary training program was superior to the voluntary training program alone in immediate leg pain intensity and symptom severity results. However, a meta-analysis of the pooled data detected only small differences, and the certainty of the evidence was very low. Our findings were similar to those of previous studies in both effect direction and magnitude. Two RCTs included in the meta-analysis were also included in a prior study [7], and this could explain the similarity between the prior study and our study.

Back pain intensity did not improve significantly in both immediate and long-term results. Although the results were heterogeneous (Fig. 3), due to the small number of studies identified, the causes could not be determined. These trends were similar for the ODI (Fig. 5). To explore this inconsistent result, there are factors to consider. The three studies included in this review prescribed supervised exercises that promote lumbar flexion, including knee-holding lumbar paraspinal muscle stretching, abdominal muscle exercises, pelvic tilt exercises, and spinal manipulation [20–22]. These exercises may have been unsuitable for some patients. Padmanabhan et al. [23] reported that low back pain and ODI scores improved when exercises that encouraged lumbar extension in patients with LSS who had challenges with lumbar flexion were performed. It has been suggested that the motor approach to LSS should focus not only on radiological findings but on the direction of movement and posture in which symptoms disappear or are alleviated [24]. Therefore, clinicians and therapists may need to be cautious when introducing supervised exercises that promote lumbar flexion.

Regarding functional impairment/disability as assessed by the ZCQF, two studies compared the voluntary training programs alone versus a combination of the training program and supervised exercise. ZCQF score is a useful clinical index to determine the severity of lower extremity symptoms and neuropathic intermittent claudication and reflects the subjective walking ability of patients [25]. Of the two studies that reported immediate results, one supported the combination of supervised exercise with the training program, while the other showed no statistical difference. In the meta-analysis, no statistical differences were detected to support the superiority of the combination of supervised exercise and the training program. In one study that examined long-term outcomes, beneficial results were observed after the combination of supervised exercise and the training program [20]. However, the size of the effect was small. In addition, SPWT results, reflecting objective walking ability, suggested that a combination of supervised exercise and voluntary training is significantly better than voluntary training alone in achieving both immediate and long-term results. Although there are limitations in interpreting the effect size of SPWT, the improvement in walking ability due to the combination of supervised exercise and voluntary training may differ between subjective and objective assessments.

Regarding QOL assessed by the SF-36, two studies compared a voluntary training program alone and its combination with supervised exercise, and one study compared ESI alone and its combination with supervised exercise. In the meta-analysis of the pooled data, improvements in several subcomponents were shown. The benefits of exercise and physical activity on well-being have been reported from several neuroscientific perspectives [26,27]. Increased exercise or physical activity due to the combination of supervised exercise with other non-surgical treatments may have influenced these improvements.

According to the results of our review, there was no firm evidence that the combination of supervised exercise notably improves clinical outcomes in LSS. In previous reviews, it has been emphasized that supervised exercise (and/or a combination of manual therapy) improves pain, symptom severity, and physical function in LSS more than medical, voluntary training, or group exercise [7,8]. Although supervised exercise is superior to other exercise modalities, the benefits of combining it may be less. The review results suggest the need for clinicians and therapists to consider which non-surgical treatments and supervised exercise combinations are optimal.

This systematic review has some limitations. First, the number of included studies was limited, and the sample size was small. The lack of an additional meta-analysis and the assessment of publication bias may have affected the weak estimates of evidence. In addition, the number of studies that contributed to the long-term results tended to be even smaller than that of studies that contributed to the immediate results. Long-term outcomes of non-surgical treatment have also been reported in previous large cohort studies [28]. In the future, a systematic review that includes cohort studies in the eligibility criteria should be considered. Second, all the studies that contributed evidence had a potential risk of bias. For patient-reported outcomes, the risk of bias is higher if patient-blinding is not ensured [16]. Although double-blinding is difficult when comparing exercise therapies, the influence of implementation bias and detection bias on the results cannot be ignored. Third, our review included only conditions that combine voluntary training and ESI with supervised exercise; therefore, the findings of this study cannot be generalized to the combinations of supervised exercise with other non-surgical treatments.

5. Conclusion

The number of studies on this topic was small and limited. Current evidence suggests that the combination of non-surgical treatment and supervised exercise probably does not provide significant benefit. Many studies are needed in the future to reach a certain conclusion on this topic.

Ethical approval

Not applicable.

Funding

The authors report no funding.

Informed consent

Not applicable.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

The authors have no acknowledgments.

Author contributions

All authors contributed to the conception and design of the study. Material preparation, data collection, and analysis were performed by RU, TI and AS. The first draft of the manuscript was written by RU, and all authors commented on the previous versions of the manuscript. All authors read and approved the final manuscript.

Supplementary data

The supplementary files are available to download from http://dx.doi.org/10.3233/BMR-220220.

References

[1] Norden J, Smuck M, Sinha A, Hu R, Tomkins-Lane C. Objective measurement of free-living physical activity (performance) in lumbar spinal stenosis: Are physical activity guidelines being met? Spine J. 2017; 17(1): 26–33. doi: 10.1016/j.spinee.2016.10.016.

[2] Ishimoto Y, Yoshimura N, Muraki S, Yamada H, Nagata K, Hashizume H, et al. Prevalence of symptomatic lumbar spinal stenosis and its association with physical performance in a population-based cohort in Japan: The Wakayama Spine Study. Osteoarthritis Cartilage. 2012; 20(10): 1103–1108. doi: 10.1016/j.joca.2012.06.018.

[3] United Nations Department of Economic and Social Affairs Population Division; 2019. World population Prospects 2019: highlights.ST/ESA/SER.A/423 [updated 2022 Jul 11; cited 2021 May 14]. Available from: https://population.un.org/wpp/Publications.

[4] Macedo LG, Hum A, Kuleba L, Mo J, Truong L, Yeung M, et al. Physical therapy interventions for degenerative lumbar spinal stenosis: A systematic review. Phys Ther. 2013; 93(12): 1646–1660. doi: 10.2522/ptj.20120379.

[5] Ammendolia C, Stuber K, Tomkins-Lane C, Schneider M, Rampersaud YR, Furlan AD, et al. What interventions improve walking ability in neurogenic claudication with lumbar spinal stenosis? A systematic review. Eur Spine J. 2014; 23(6): 1282–1301. doi: 10.1007/s00586-014-3262-6.

[6] Slater J, Kolber MJ, Schellhase KC, Patel CK, Rothschild CE, Liu X, et al. The Influence of Exercise on Perceived pain and disability in patients with lumbar spinal stenosis: A systematic review of randomized controlled trials. Am J Lifestyle Med. 2016; 10(2): 136–147. doi: 10.1177/1559827615571510.

[7] Rousing R, Jensen RK, Fruensgaard S, Strøm J, Brøgger HA, Degn JDM, et al. Danish national clinical guidelines for surgical and nonsurgical treatment of patients with lumbar spinal stenosis. Eur Spine J. 2019; 28(6): 1386–1396. doi: 10.1007/s00586-019-05987-2.

[8] Jacobi S, Beynon A, Dombrowski SU, Wedderkopp N, Witherspoon R, Hébert JJ. Effectiveness of conservative nonpharmacologic therapies for pain, disability, physical capacity, and physical activity behavior in patients with degenerative lumbar spinal stenosis: A systematic review and meta-analysis. Arch Phys Med Rehabil. 2021; 102(11): 2247–2260.e7. doi: 10.1016/j.apmr.2021.03.033.

[9] Bagley C, MacAllister M, Dosselman L, Moreno J, Aoun SG, El Ahmadieh TY. Current concepts and recent advances in understanding and managing lumbar spine stenosis. F1000Res 2019; 8: F1000 Faculty. F1000Res. 2019; 8: Rev–137. doi: 10.12688/f1000research.16082.1.

[10] Hageman D, Fokkenrood HJ, Gommans LN, van den Houten MM, Teijink JA. Supervised exercise therapy versus home-based exercise therapy versus walking advice for intermittent claudication. Cochrane Database Syst Rev. 2018; 4(4): CD005263. doi: 10.1002/14651858.CD005263.pub4.

[11] Velthuis MJ, Agasi-Idenburg SC, Aufdemkampe G, Wittink HM. The effect of physical exercise on cancer-related fatigue during cancer treatment: A meta-analysis of randomised controlled trials. Clin Oncol (R Coll Radiol). 2010; 22(3): 208–221. doi: 10.1016/j.clon.2009.12.005.

[12] Moseng T, Dagfinrud H, Smedslund G, Østerås N. The importance of dose in land-based supervised exercise for people with hip osteoarthritis. A systematic review and meta-analysis. Osteoarthritis Cartilage. 2017; 25(10): 1563–1576. doi: 10.1016/j.joca.2017.06.004.

[13] Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ. 2021; 372: n71. doi: 10.1136/bmj.n71.

[14] Urata R, Igawa T, Suzuki A, Kubo A. Effects of supervised exercise on lumbar spinal stenosis: A systematic review and meta-analysis of randomized controlled trials. Prospero. 2020; CRD42020199232. Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020199232.

[15] Hansen S, Aaboe J, Mechlenburg I, Overgaard S, Mikkelsen LR. Effects of supervised exercise compared to non-supervised exercise early after total hip replacement on patient-reported function, pain, health-related quality of life and performance-based function – a systematic review and meta-analysis of randomized controlled trials. Clin Rehabil. 2019; 33(1): 13–23. doi: 10.1177/0269215518791213.

[16] Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: A revised tool for assessng risk of bias in randomised trials. BMJ. 2019; 366: l4898. doi: 10.1136/bmj.l4898.

[17] Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines. GRADE guidelines: 7. Rating the quality of evidence-inconsistency. J Clin Epidemiol. 2011; 64(12): 1294–1302. doi: 10.1016/j.jclinepi.2011.03.017.

[18] Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol. 2011; 64(4): 401–406. doi: 10.1016/j.jclinepi.2010.07.015.

[19] Cleland JA, Whitman JM, Houser JL, Wainner RS, Childs JD. Psychometric properties of selected tests in patients with lumbar spinal stenosis. Spine J. 2012; 12(10): 921–31.

[20] Ammendolia C, Côté P, Southerst D, Schneider M, Budgell B, Bombardier C, et al. Comprehensive nonsurgical treatment versus self-directed care to improve walking ability in lumbar spinal stenosis: A randomized trial. Arch Phys Med Rehabil. 2018; 99(12): 2408–2419.e2. doi: 10.1016/j.apmr.2018.05.014.

[21] Hammerich A, Whitman J, Mintken P, Denninger T, Akuthota V, Sawyer EE, et al. Effectiveness of physical therapy combined with epidural steroid injection for individuals with lumbar spinal stenosis: A randomized parallel-group trial. Arch Phys Med Rehabil. 2019; 100(5): 797–810. doi: 10.1016/j.apmr.2018.12.035.

[22] Minetama M, Kawakami M, Teraguchi M, Kagotani R, Mera Y, Sumiya T, et al. Supervised physical therapy vs. home exercise for patients with lumbar spinal stenosis: A randomized controlled trial. Spine J. 2019; 19(8): 1310–1318. doi: 10.1016/j.spinee.2019.04.009.

[23] Padmanabhan G, Sambasivan A, Desai MJ. Three-step treadmill test and McKenzie mechanical diagnosis and therapy to establish directional preference in a patient with lumbar spinal stenosis: A case report. J Man Manip Ther. 2011; 19(1): 35–41. doi: 10.1179/2042618610Y.0000000002.

[24] Longtin C, Busseau Y, Jetté M, Cabana-Boucher G, Ouellet C, Lam OT, et al. Systematic flexion-based approach for patients with radiological signs of lumbar spinal stenosis: myth or reality? A retrospective study. Ann Phys Rehabil Med. 2018; 61(4): 270–272. doi: 10.1016/j.rehab.2017.03.005.

[25] Tomkins-Lane CC, Battié MC, Macedo LG. Longitudinal construct validity and responsiveness of measures of walking capacity in individuals with lumbar spinal stenosis. Spine J. 2014; 14(9): 1936–1943. doi: 10.1016/j.spinee.2013.11.030.

[26] Di Liegro CM, Schiera S, Proia P, Di Liegro I. Physical activity and brain health. Genes. 2019; 10(9): 720. doi: 10.3390/genes10090720.

[27] Matta Mello Portugal E, Cevada T, Sobral Monteiro-Junior R, Teixeira Guimarães T, da Cruz Rubini E, Lattari E, et al. Neuroscience of exercise: From neurobiology mechanisms to mental health. Neuropsychobiology. 2013; 68(1): 1–14. doi: 10.1159/000350946.

[28] Fritz JM, Lurie JD, Zhao W, Whitman JM, Delitto A, Brennan GP, et al. Associations between physical therapy and long-term outcomes for individuals with lumbar spinal stenosis in the SPORT study. Spine J. 2014; 14(8): 1611–1621. doi: 10.1016/j.spinee.2013.09.044.