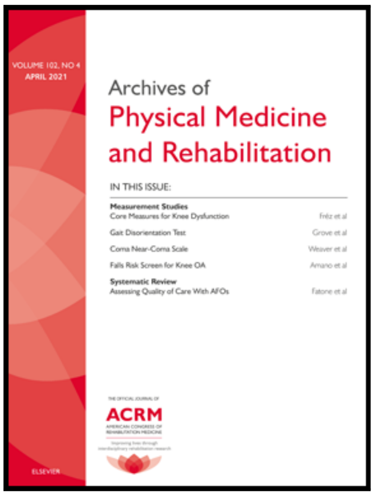
Journal Pre-proof



Is aerobic or resistance training the most effective exercise modality for improving lower extremity physical function and perceived fatigue in people with multiple sclerosis? A systematic review and meta-analysis.

Laurits Taul-Madsen MSc. , Luke Connolly PhD ,
Rachel Dennett BSc , Jenny Freeman PhD , Ulrik Dalgas PhD ,
Lars G. Hvid PhD

PII: S0003-9993(21)00307-5
DOI: https://doi.org/10.1016/j.apmr.2021.03.026
Reference: YAPMR 58199

To appear in: *Archives of Physical Medicine and Rehabilitation*

Received date: 28 January 2021
Revised date: 10 March 2021
Accepted date: 25 March 2021

Please cite this article as: Laurits Taul-Madsen MSc. , Luke Connolly PhD , Rachel Dennett BSc , Jenny Freeman PhD , Ulrik Dalgas PhD , Lars G. Hvid PhD , Is aerobic or resistance training the most effective exercise modality for improving lower extremity physical function and perceived fatigue in people with multiple sclerosis? A systematic review and meta-analysis., *Archives of Physical Medicine and Rehabilitation* (2021), doi: https://doi.org/10.1016/j.apmr.2021.03.026

Is aerobic or resistance training the most effective exercise modality for improving lower extremity physical function and perceived fatigue in people with multiple sclerosis? A systematic review and meta-analysis.

Laurits Taul-Madsen[1], MSc.

Luke Connolly[2], PhD

Rachel Dennett[2], BSc

Jenny Freeman[2], PhD

Ulrik Dalgas[1], PhD

Lars G. Hvid[1], PhD

1. Exercise Biology, Department of Public Health, Aarhus University, Denmark

2. School of Health Professions, Faculty of Health, University of Plymouth, UK.

Corresponding author:
Laurits Taul-Madsen
Dalgas Avenue 4
8000 Aarhus C
Denmark
ltm@ph.au.dk
+45 61656063

Objective: The purpose of this systematic review was to investigate whether aerobic training (AT) or resistance training (RT) is most effective in terms of improving lower limb physical function and perceived fatigue in persons with multiple sclerosis (pwMS).

Data sources: Nine databases (MEDLINE, EMBASE, CINAHL, AMED, PEDro, SPORTdiscus, PsycINFO, Web of Science and SCOPUS) were electronically searched in April 2020.

Study Selection: Included studies were randomized controlled trials (RCTs) involving pwMS attending one of two exercise interventions; AT or RT. Studies had to include at least one objective or self-reported outcome of lower extremity physical function and/or perceived fatigue.

Data Extraction: Data was extracted using a customized spreadsheet, which included detailed information on patient characteristics, interventions and outcomes. The methodological quality of the included studies was independently assessed by two reviewers using the TESTEX rating scale.

Data synthesis: Twenty-seven papers reporting data from 22 RCTS (AT=14, RT=8) including 966 pwMS. The two modalities were found to be equally effective in terms of improving short walk test (AT: ES=0.33 [-1.49: 2.06]; RT: ES=0.27 [0.07: 0.47]) and long walk test performance (AT: ES=0.37 [-0.04: 0.78]; RT: ES=0.36 [-0.35: 1.08]), as well as in reducing perceived fatigue (AT: ES=-0.61 [-1.10:-0.11]; RT: ES=-0.41 [-0.80: -0.02]). Findings on other functional mobility tests along with self-reported walking performance were sparse and inconclusive.

Conclusions: AT and RT appear equally highly effective in terms of improving lower extremity physical function and perceived fatigue in pwMS. Clinicians can thus use either modality to target impairments in these outcomes. In a future perspective, head-to-head exercise modality studies are warranted. Future MS exercise studies are further encouraged to adapt a consensus ΨŎƵǊŜ ōŀ·űűŜǊèΩ ƵŦ ǇKèăƛŎŀ·Ŧ ŦdzyŎűƛƵƴ űŜăűă űƵ ŦŀŎƛŧƛűŀ·űŜ ŀ· ŘŜűŀ·ƛŧŜŘ ŎƵƳǇŀ·ǊƛăƵƴ ƵŦ ǊŜădzŧűă ŀ·ŎǊƵăă modalities.

Keywords:
Rehabilitation; Exercise; Multiple Sclerosis; Systematic Review.

Abbreviations:

MS: Multiple Sclerosis

pwMS: Person with multiple sclerosis

RCT: Randomized controlled trial

AT: Aerobic training

RT: Resistance training

RM: Repetition maximum

HR: Heart rate

Testex: Tool for assessment of study quality for reporting on exercise

RPE: ratings of perceived exertion

ES: Effect size

CI: Confidence interval

EDSS: Expanded disability status scale

6MWT: Six minute walk test

MSWS: 12-item multiple sclerosis walking scale

SSST: Six spot step test

FSS: Fatigue severity scale

Introduction

Multiple sclerosis (MS) is a chronic, autoimmune, and inflammatory disease of the central nervous system, exemplified through demyelination and axonal loss[1]. As a consequence, multiple symptoms can appear[1-3], with fatigue and walking limitations reported to be among the most debilitating[4-7]. Moreover, an estimated 50% of persons with MS (pwMS) will require a walking aid within 15-25 years after disease onset[8,9]. Since physical function is associated with lowered quality of life at the individual level along with a greater economic burden at a health service and societal level [10,11], it is crucial to diminish progression of disability[12].

While pharmacological treatments appear to have limited beneficial effect on fatigue and walking limitations[13], exercise has proven to be a potent non-pharmacological treatment option, being both safe and eliciting numerous beneficial effects in pwMS [14,15]. Specifically, exercise is an effective way of reducing fatigue[16,17] and improving walking performance [18,19], with the latter often considered to be clinically meaningful[20,21].

Exercise constitutes a number of different modalities known to elicit different physiological adaptations (such as neuromuscular function or cardiovascular function) that in most cases are paralleled by (and perhaps even translated into) improved physical function[22]. A recent review investigating randomized controlled trials (RCTs) of exercise interventions in pwMS reported that the two most applied exercise modalities were aerobic training (AT) and resistance training (RT)[23]. Several studies have reported positive effects of both AT[24-26] and RT[27,28] on parameters directly related to lower extremity physical function (e.g. walking performance, chair rise, stair negotiation) as well as on parameters indirectly related to lower extremity physical function, such as perceived fatigue. However, based on the existing literature it currently remains unknown which of these two common exercise modalities is the most effective in terms of improving physical function and perceived fatigue in pwMS. Despite the somewhat impossible task of matching AT and RT on traditional exercise parameters such as duration, frequency, and intensity, understanding the specific effectiveness of the two different exercise modalities is an important factor for consideration in optimizing exercise prescription in pwMS.

Therefore, the objectives of this systematic review were to investigate which of the two exercise modalities (AT or RT) are the most effective in terms of improving lower extremity physical function and reducing perceived fatigue in pwMS.

Methods:

The present systematic review follows the Preferred Reporting Items for Systematic Reviews and Meta Analyses (PRISMA) guidelines on systematic reviews of RCTs[29]. Search strategy, study selection, eligibility criteria, methodology assessment, data extraction and analysis were performed in accordance with a protocol pre-registered in PROSPERO (CRD42020189855).

Definitions:

In this review the following definitions were applied:

Exercise: A form of physical activity that is planned, structured and repetitive, and is undertaken with the objective of improving or maintaining at least one aspect of physical fitness, comprising strength, flexibility or aerobic endurance[30].

Physical activity: Any bodily movement produced by skeletal muscles that requires energy expenditure above resting levels[30].

Physical function: The ability of an individual to perform physical activities of daily living. For the purposes of this systematic review, this particularly relates to lower extremity tasks (e.g. simple/complex/endurance walking, chair rise, stair negotiation) [31].

Perceived fatigue: Subjective sensations of weariness, increasing sense of effort, mismatch between effort expended and actual performance or exhaustion[32].

Resistance training: Performed with external resistance of varying degrees relative to maximal strength provided by either free weights, machines, bodyweight, or some other implements (e.g., resistance bands), either with single or multiple sets of repetitions which may or may not be performed to momentary failure (but are often performed to a relatively high effort)[33].

Aerobic training: Performed using locomotor or ergometer tasks (e.g., walking, jogging, running, cycling, rowing, etc.) in a continuous or intermittent fashion with respect to duration at submaximal intensities of effort, commonly determined relative to either maximal heart rate, heart rate reserve, VO_{2max}, or sometimes using ratings of perceived effort scales[33].

Exercise intensity: For ATΣ ŜÈŜ̈ŎλăŜ Ж 63% of Heartrate max (HRmax) was defined as low intensity, 64-76% of HRmax as moderate intensity, I·yŘ ж 77 % of HRmax as high intensity[34]. For RT, ŜÈŜ̈ŎλăŜ ж 16 Repetition Maximum (RM) was defined as low intensity (Ж 64% of 1RM), 9-15 RM as moderate intensity (65-79% of 1RM) I·yŘ Ж 8 RM as high intensity (ж yɫ ƻŦ ᴍRM)[35,36]

Searches

An original search was carried out as part of another review by the same authors in 2018, having the aim to summarize reported adherence and drop-out data from RCT studies of exercise interventions in pwMS.[23]

This search was updated in April 2020. Furthermore, in March 2020, The World Health hƌƎI·yλÎI·ŰλƻyΩă Lyŕ̂ŜyЛ·ŰλƻyЛ·ƒ /ƒλyЛŎI·ƒ ¢λI·ƒă wŜƎλăŰè tƒI·ŰŦƻ̈Ƴ όLCTRP) http://apps.who.int/trialsearch/, which comprise the 16 primary registries of the WHO registry network and ClinicalTrials.gov, was searched for relevant ongoing trials investigating a head-to-head comparison of AT and RT in pwMS.

Data sources and search strategy

LƴōλŜŦΣ ŰKŜ ăŜI·ŎK ăŰI·ŰŜƎè ǿI·ă ōI·ăŜŘ ƻƴ ŰKŜ ƚŜè ŰŜƳă όƳƨŰλLJŦŜ ăŎƚŜƻăλăέ hw a{ ! b5 exerciăŜ hw όLJKèăλŎI·ƒ I·ŎŰλǾλŰèέΦ CƻƎ full search strategy please see Dennett et al. 2020[23].

The original search was carried out in October 2018 and updated in April 2020.

Two reviewers (LM and RD) conducted the original search in the electronic databases MEDLINE, EMBASE, CINAHL, AMED, PEDro, SPORTdiscus, PsycINFO, Web of Science and SCOPUS limited to scientific research papers being published between January 1993 and October 2018. The same databases were searched from September 2018 to March 2020 by two reviewers (LM and LC) in April 2020. All searches were supplemented by hand searches of reference lists.

Study selection

The following PICO (population, intervention, comparison, outcome) question guided the ăŜI·ǋŎK I·yŘ ʎyŎƭdzăʎƨyˁăűǋI·űŜ∃èɸ ά ² KʎŎK ŜȅŜǋŎʎăŜ ϒƧŘI·ƭʎűèƩ AT or RT, is most effective in improving physical function (specifically lower extremity tasks such as simple/complex/endurance walking, chair rise, stair negotiation) and perceived fatigue in pwMSKέ

Eligibility criteria

RCT studies involving adults over the age of 18 with a definite diagnosis of MS, regardless of gender, disease duration, MS phenotype or level of disability were considered eligible for inclusion. While all identified studies could be included regardless of location, group/ individual structure, level of supervision, intervention duration, session duration, intensity, progression, frequency, the content had to be either AT or RT; with or without a follow-up period.

Control interventions had to include non-training controls only or active control conditions having no expected effects on the cardiovascular system or the musculoskeletal system, for example stretching were accepted.

Studies had to include at least one objective or self-reported measure of lower extremity physical function (such as simple/complex/endurance walking, chair rise, stair negotiation) and/or perceived fatigue. If reported, measures of cardiovascular function (i.e. maximal oxygen uptake) and neuromuscular function (i.e. maximal muscle strength or muscle power) were also extracted, as these outcomes could, (1) help verify the effectiveness of interventions, and (2) are likely mediators of adaptations in lower extremity physical function.

Data management and selection process

The original search resulted in 93 papers included in the previous review, all of which were considered for inclusion in the present review (see figure 1).

Results from the updated search were exported to EndNote, where duplicates were removed. The remaining papers were imported into Rayyan data management system (rayyan.qcri.org) where titles and abstracts were independently screened for eligibility by two reviewers (LC and LTM). If papers were included at this stage, a full-text reading by the same two reviewers was performed, and any discrepancies were discussed with a third party (LGH). Reasons for excluding full text RCTs were recorded.

Data extraction

Data was extracted using the same spreadsheet as the previous review[23], which included detailed information on participant characteristics (age, gender, disease duration, MS phenotype, disability level, and fatigue as a symptom); modality of the intervention (setting, group/individual structure, level of supervision, intervention duration, session duration, intensity, frequency); content of the intervention (aerobic or resistance); report of adverse events, % drop-out, and adherence during the intervention period and at any follow up. Furthermore, an additional customized spreadsheet was made to extract information on all outcomes of lower extremity physical function, perceived fatigue and measures of cardiovascular and neuromuscular function. Data extraction was completed by two reviewers (LC and LTM).

Quality assessment

The methodological quality of the included studies was independently assessed by two reviewers (LTM and LC) using the "Tool for assessment of study quality for reporting on exercise" (TESTEX) rating scale[37]. Any discrepancies were discussed and resolved between the two reviewers.

Synthesis of results

In addition to the qualitative analysis (summary of identified studies and their data), we also performed quantitative analysis by calculating sample-size weighted averages across selected studies. A minimum of two studies was required in order to conduct a meta-analysis. Random effects meta-analyses comprising data on physiological adaptations, short walking tests, long walking tests and perceptions of fatigue were conducted by using *Meta-Essentials version 1.5* designed for Excel.[38] Intervention effect sizes (ES) (between-group differences) for different outcomes at post-treatment, were calculated using Hedges' g statistic, along with 95% confidence intervals (CIs) around the estimated effect-size. Also, if data was available and adequate, we performed a weighted regression of all study ES as a function of intervention duration and frequency (weeks and number of sessions) as well as intervention intensity, as these factors were hypothesized to impact the outcomes[39]. Of note, this approach was done to establish specific within-modality information only. ES were interpreted as follows: small = 0.14, moderate = 0.31, large = 0.61 based on empirical data from 99 meta-analyses examining the effects of rehabilitation/exercise[41]. {űl·űλăűλŎl·ƭ ƘŜűŜǊƶƎŜƴŜλűè ǿl·ă ljdzl·ƴűλŦλŜŘ dzăλƴƎ l λƎƎλƴăΩ L^2 statistic, and was interpreted as follows: heterogeneity: > 50%, no or limited heterogeneity: < 50%[42].

If studies reported on more than one outcome in each domain (e.g. physiological adaptations such as knee extensor and knee flexor muscle strength as well as perceptions of fatigue using different questionnaires), an average was calculated and used for the meta-analyses.

[INSERT FIGURE 1]

Results

Study characteristics

As depicted in figure 1, the search yielded 2117 hits. After removal of duplicates, 1538 papers remained for the screening process, with 12 of these assessed for full-text reading. Five papers were included, which with the addition of 22 papers from the previous review, resulted in a total of 27 papers being included in the qualitative and quantitative synthesis.

The 27 papers reported 22 RCTs (AT (n=14), RT (n=8)) which involved a total of 966 pwMS. As seen in table 1, Expanded disability status scale (EDSS) ranged from 1.5-7 while disease duration ranged from 2.7-18.6 years. The duration of AT interventions ranged from 3-26 weeks (involving 9-48 sessions) with the intensity being deemed moderate (n=5)[43-47], high (n=4)[26,48-50], or unknown (no information, n=5)[25,51-54]. The duration of RT interventions ranged from 8-24 weeks (involving 15-48 sessions) with the intensity being deemed moderate (n=1)[55], high (n=4)[28,56-58], or unknown (no information, n=3)[40,59,60]. Due to the missing information and the use of divergent scales of exercise intensity for both AT (e.g. % of HR_{max}, RPE, % of $VO2_{max}$, % of Peak Power) and RT (% of 1RM, % of bodyweight, absolute weights), we were unable to perform weighted (moderator) analysis using this parameter. Two[25,60] of the 22 identified RCTs reported a primary outcome that was not based on a sample size calculation. Ten papers[26,28,44,48-50,52,54,55,57] of the 22 identified RCTs reported a primary outcome based on a sample size calculation, with five of these having a primary outcome aligned with the purpose of the present systematic review.

The median TESTEX score of the included studies was nine out of 15. Detailed information on the scores can be found in Table 2.

[INSERT TABLE 1]

Table 1: Characteristics of included studies

Study	Active/ passive control	Recruited (n)	Analysed (n)	Sex (% female)	Age (Years mean)	MS type (% RRMS)	EDSS (0-10)	Disease duration (mean)	Individual/ group	Supervised/ unsupervised	Frequency (d/wk) and duration (wks)	Intensity	Length of training session (min)
Aerobic training													
Ahmadi, 2010[55] 2013[72] and Arastoo 2013[73]	Waitlist control	Int: 11 Con :10	NR NR	100 100	32.2 36.7	NR NR	2.0 2.3	4.7 5.0	NR N/A	Supervised N/A	3d/wk *8wk NR	40-75 % of HR_{max} N/A	30 N/A
Bahmani, 2019[56]	Active control (attention control)	Int: 31 Con:31	26 21	100 100	38.0 37.9	NR	2.5 2.0	6.9 7.2	Group Group	Supervised Supervised	3d/wk *8wk 3d/wk *8wk	NR N/A	30-45 30-45
Baquet, 2018[41]	Waitlist control (usual activity)	Int: 34 Con:34	34 34	62 74	38.2 39.6	100 100	1.7 1.8	6.8 5.7	Group N/A	Supervised N/A	2.5d/ wk*12wk N/A	RPE of 8 NR	27-69 N/A
Dettmers, 2009[25]	Active control (stretching and relaxation)	Int: 15 Con:15	15 15	67 73	45.8 39.7	77	2.6 2.8	10.7 10.5	Group Group	Supervised Supervised	3d/wk *3wk 3d/wk *3wk	NR NR	45 45
Feys, 2017[44]	Waitlist control (usual activity)	Int: 21 Con: 21	18 17	95 86	36.6 44.4	NR	NR	8.1 9.2	Indiv N/A	Unsupervised N/A	3d/wk *12wk N/A	NR N/A	NR N/A
Heine, 2017[26]	Active control (consultation with MS nurse)	Int1: 43 Con: 46	33 30	74 72	48.8 48.4	79 74	2.5 3.0	7.0 12.0	Group NR	Unsupervised NR	3d/wk *16wks NR	3min 40%, 1 min 60% and 1min at 80% of PPO. N/A	30 NR
Hebert, 2011[45]	Waitlist control (usual activity)	Int: 13 Con: 13	13 12	85 85	42.6 50.	85 92	NR NR	5.1 9.1	NR N/A	Supervised N/A	2d/wk *6wk N/A	65-75% of HR_{max} N/A	60 N/A

					2									
Mokhtarzade, 2017[49]	Passive control	Int : 25 Con:20	2218	100100	32.031.3	100100	1.841.57	2.693.47	NR N/A	NR N/A	3d/wk *8wk N/A	60-75% of W_{max} N/A	42-66 N/A
Langeskov-Christensen, 2020[46]	Waitlist control (habitual activity)	Int : 43 Con:43	4343	6060	44.045.6	9579	2.72.8	10.98.6	Group N/A	Supervised N/A	2d/wk *24wk N/A	65-95% of HR_{max} N/A	30-60 N/A
Mostert, 2002[57]	Active control (usual activity)	Int: 18 Con:18	1313	7785	45.243.9	30.838.5	4.64.5	11.211.6	NR N/A	Supervised N/A	5d/wk *4wk N/A	NR N/A	30 N/A
Oken, 2004[47]	Waitlist control (usual activity)	Int: 21 Con: 22	1520	87100	48.848.4	NR	2.93.1	NR	Indiv N/A	Supervised N/A	1d/wk *26wk N/A	NR N/A	NR N/A
Petajan, 1996[50]	Waitlist control (usual activity)	Int: 21 Con:25	2125	7164	41.139.0	NR	3.82.9	9.36.2	NR N/A	Supervised N/A	3d/wk *15wk N/A	60% of VO_{2max} N/A	50 N/A
Schulz, 2004[51]	Waitlist control (usual activity)	Int::15 Con:13	1513	7362	39.042.0	NRNR	2.02.5	NRNR	NR NR	NR NR	2d/wk *8wk N/A	75% of W_{max} NR	30 N/A
Tollar, 2020[48]	Waitlist control (usual activity)	Int: 14 Con:12	1412	9392	48.144.4	5066	5-65-6	13.214.0	Group N/A	Supervised Supervised	5d/wk *5wk N/A	80% of age-predicted HR_{max} N/A	60 N/A
Progressive resistance training													
Callesen, 2019[42]	Passive control (usual activity)	Int: 23 Con:20	1718	7080	52.056.	7065	4.03.5	15.011.	Group N/A	Supervised N/A	2d/wk *10wk N/A	10 sets at 15 RM ς 8 sets at 8 RM N/A	NR N/A

					0			0					
Dalgas, 2009[28] and 2010[58,74]	Passive control (waitlist usual activity)	Int: 19 Con: 19	1516	6662	47.750.4	100	3.73.9	6.68.1	Group N/A	Supervised N/A	2d/wk*12wk N/A	3-4 sets of 8-12 repetitions at 8-15 RM N/A	NR N/A
DeBolt, 2004[40]	Passive control (usual activity)	Int: 19 Con: 18	1917	7978	51.647.8	4744	4.03.5	15.113.0	Indiv N/A	Unsupervised N/A	3d/wk*8wk N/A	2-3 sets of 8-12 repetitions wearing a weighted vest (0.5% of BW) increasing by 0.5-1.5% of BW every 2 wk	35-50 N/A
Dodd, 2011[43]	Passive control (Usual activity + social program)	Int: 39 Con: 37	3635	7274	47.750.4	100	NR	NR	Group Group	Supervised Supervised	2d/wk*10wk 1d/wk*10wk	2 sets of 10-12 repetitions at 10-12 RM N/A	4560
Harvey, 1999[36]	Passive control (usual activity)	Int : 7 Con: 5	65	8380	38.043.0	100	NR	510	Indiv N/A	Unsupervised N/A	2d/*8wk N/A	5 sets of 10 leg extensions using 0.5 or 1kg ankle weights N/A	NR N/A
Hosseini, 2018[52]	Passive control (usual activity)	Int: 9 Con: 8	88	5550	32.933.0	NR	NR	NR	Indiv N/A	Unsupervised N/A	3d/wk*8wk N/A	1% of BW fastened to body increasing by 0.5-1% every 2 wk N/A	35-50 N/A
Kjolhede, 2015[53] and Jørgensen 2019[75]	Passive control (waitlist usual activity)	Int: 18 Con: 17	1715	NR	43.2	100	3	5	NR N/A	Supervised N/A	2d/wk*24wk N/A	3-5 sets of 6-10 repetitions at 6-15 RM N/A	NR N/A
Moradi, 2015[54]	Passive control (usual activity)	Int: 10 Con: 10	810	00	34.433.1	6260	3.03.0	8.16.5	NR N/A	Supervised N/A	3d/wk*8wk N/A	1 set of 6-15 repetitions at 50-80% of 1 RM N/A	30 N/A

Abbreviations: MS: Multiple sclerosis; RRMS: Relapse remitting multiple sclerosis; Int: Intervention; Con: Control; PPO: Peak power output achieved during incremental exercise test to exhaustion; VO_{2max}: Maximal oxygen consumption; VO_{2peak}: Peak oxygen consumption; HR: Heart rate; RM: Repetition maximum; RPE: ratings of perceived exertion; BW: Body weight; MVC: Maximal voluntary contraction;

MIP: Maximal inspiratory pressure; TMW: Tolerated maximum workload; W: Watts; Indiv: Individual; N/A: not applicable; NR: not reported.

Table 2: TESTEX scores

Paper	Eligibility criteria	Randomi-zation	Allocation concealed	Baseline data	assessor primary OM	OM in >85% patients	AE reported	Exercise attendance	Intention-to-treat	stats primary OM	stats secondary OM	Outcomes point estimates	Control Physical activity	Exercise load titrated	calculated	Total
Ahmadi 2010	1	0	1	1	0	1	0	0	0	0	1	1	0	1	1	8
Ahmadi 2013a	1	0	1	1	0	1	0	0	0	0	1	1	0	0	0	6
Ahmadi 2013b	1	0	1	1	0	1	0	0	0	0	1	1	0	1	1	8
Bahmani 2019	1	1	1	1	0	0	0	0	0	0	1	1	0	0	0	6
Baquet 2018	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	13
Callesen 2019	1	0	1	1	1	0	1	1	1	1	1	1	0	1	1	12
Dalgas 2009	1	1	1	1	1	0	0	1	0	1	1	1	0	1	1	11
Dalgas 2010a	1	0	1	1	0	0	0	1	0	1	1	1	0	1	1	9
Dalgas 2010b	1	0	1	1	1	0	0	1	0	1	1	1	0	1	1	10
DeBolt 2004	1	0	0	1	0	1	0	1	0	1	1	1	0	1	1	9
Dettmers 2009	1	1	0	1	0	1	0	0	0	1	1	0	0	0	0	6
Dodd 2011	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	13
Feys 2017	1	0	0	1	0	0	1	1	1	1	1	1	0	0	1	9
Harvey 1999	1	1	0	1	0	1	0	1	0	1	1	0	0	0	0	7
Hebert 2011	1	0	1	1	1	0	1	1	1	1	1	1	0	0	1	11
Heine 2017	1	1	0	0	1	0	0	1	1	1	1	1	1	1	1	11
Hosseini 2018	1	1	0	1	0	1	0	0	0	0	0	1	0	0	0	5
Jørgensen 2019	1	0	1	1	0	1	0	0	0	0	1	0	0	1	1	7
Kjølhede 2015	1	1	1	0	0	1	0	1	1	1	1	1	0	1	1	11
Langeskov-Christensen 2020	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	13
Mokhtarzade 2017	1	0	1	1	0	1	0	0	0	0	0	1	0	1	1	7
Moradi 2015	1	1	0	1	1	1	1	0	0	1	1	1	0	1	1	11
Mostert 2002	1	0	1	1	0	0	0	1	0	0	0	0	0	0	0	4
Oken 2004	1	0	1	1	1	0	1	1	0	1	0	1	0	0	0	8
Petajan 1996	1	0	0	1	1	1	0	1	0	0	0	1	0	1	1	8
Schulz 2004	1	0	1	1	0	0	0	0	0	1	1	1	0	0	1	7
Tollar 2020	1	1	1	1	1	1	1	1	0	1	1	1	0	0	0	11

Abbreviations: OM: Outcome measure

Table 3: Effect sizes of all outcomes.

Study	Strength/ VO$_{2peak}$ (positive ES = improvement)		Short walk (Positive ES = improvement)			Long walk (Positive ES = improvement)			Other walking (Negative ES = improvement)	Functional mobility (other) (Negative ES = improvement)				Perceived fatigue (Negative ES = improvement)			
	Strength	VO$_{2peak}$	T25FW	T10MW	T50MW	2minW	6minW	Distance	MSWS	TUG	SSST	5-STS	Stair climb	FSS	MFIS	CIS20r	FSMC
Aerobic training																	
Ahmadi, Arastoo 2010+13				1.29		1.40								-2.73			
Dettmers 2009								0.47									
Bahmani 2019														-0.37			
Baquet 2018		0.32	0.00				-0.14		0.09								
Feys 2017		1.06	0.00				0.33		-0.29			-0.38					0.0
Heine 2017		0.30												-0.29	-0.29	-0.52	
Hebert 2011							0.10								-0.45		
Langeskov-Christensen 2020		0.49					0.36										
Mokhtarzade 2017		1.40												-0.71			
Mostert 2002														-0.31			
Oken 2004															-0.94		
Petajan 1996		2.18															
Schulz 2004		0.69													-0.63		
Tollar 2020							0.73										
Resistance training																	
Callesen 2019	0.73		0.20				1.07		0.47		-0.18				-0.60		
Dalgas 2009+10	0.42			0.34			-0.06					-0.83	-0.52	-0.26			
DeBolt 2004										-0.35							
Dodd 2011	0.35					0.27									-0.37		
Hosseini 2018	0.29			0.49													
Harvey 1999*	0.07			n.c	n.c.												
Kjølhede 2015, Jørgensen 2019	2.03		0.08			0.27			-0.35			-2.27	-1.84				
Moradi 2015	2.59			0.42						-0.							

87

Abbreviations: ES: Effect size; $VO2_{peak}$: Peak oxygen consumption; T25FW: Timed 25 foot walk; T10MW: Timed 10 meter walk; T50MW: Timed 50 meter foot walk; TUG: Timed up and go; SSST: Six spot step test; 5-STS: 5 times sit to stand; FSS: Fatigue Severity Scale; MFIS: Modified Fatigue Impact Scale; CIS20r: Checklist Individual Strength. *ES was non computable as no standard deviation was reported.
Bold indicates that the paper has reported a statistically significant between-group change.