Multimodal agility-based exercise training (MAT) versus strength and endurance training (SET) to improve multiple sclerosis-related fatigue and fatigability during inpatient rehabilitation: a randomized controlled pilot and feasibility study [ReFEx]

Florian Wolf[1,2]*, Jörn Nielsen[1,3], Jochen Saliger[1], Eva Hennecken[1], Philipp Kröber[1], Mareike Eschweiler[1], Ann-Kristin Folkerts[3], Hans Karbe[1] and Philipp Zimmer[2]

Abstract

Background Multimodal agility-based exercise training (MAT) is a group-based exercise training framework for persons with multiple sclerosis (pwMS) with a potential to impact fatigue and fatigability. In a mixed-methods design, this study evaluated the feasibility of implementing MAT in an inpatient rehabilitation setting and the feasibility of a randomized controlled trial (RCT) study protocol with 'traditional' strength and endurance training (SET) as an active control condition. Secondarily, preliminary outcome data was acquired.

Methods PwMS with low to moderate disability and self-reported fatigue were randomly allocated to either MAT or SET when starting inpatient rehabilitation (4–6 weeks). The MAT-participants exercised in a group following a MAT-manual (sessions were gym- (5x/week) and pool-based (3x/week)). SET-participants exercised individually 5x/week on a cycle ergometer, and 3x/week on strength training machines. Feasibility assessments focused on processes, resources, management, time, and scientific domains. Assessed clinical outcomes at admission and discharge included perceived fatigue, motor and cognitive fatigability, cognitive performance, motor function, and balance confidence. Perceived fatigue was reassessed 1, 4, and 12 weeks after discharge. Feasibility was determined regarding predetermined progression criteria.

Results Twenty-two participants were randomized. Both groups performed the minimum number of sessions (> 18), and retention was adequate (73–91%). SET-participants performed more sessions than MAT-participants (30.8 vs. 22.7) and stayed longer in the facility (34.2 vs. 31.6 days). Non-eligibility of admitted pwMS was high (74% non-eligible), mainly due to high EDSS and inability to attend pool-based sessions. Consequently, recruitment (1.8/month) was

*Correspondence:
Florian Wolf
florian.wolf@go.johanniter-kliniken.de

Full list of author information is available at the end of the article

slower than the predetermined progression criterium. Baseline assessments took longer than required (only 50% completed within 3 days). Short-term fatigue reduction was similar for both groups. Motor fatigability also improved in both groups, whereas cognitive fatigability deteriorated. In MAT, average improvement in walking endurance (43.9 m) exceeded minimal important change values for individuals (> 26.9 m).

Conclusions Progressing to a definitive RCT necessitates adaptation of eligibility criteria. In the present design it will also be difficult to attain similar dosing of interventions. A multicenter RCT focused only on gym-based MAT might be another option to assess the effect of MAT. The primary outcome measure should be able to measure change in perceived fatigue more robustly.

Trial registration German Clinical Trials Register: DRKS00023943, date of registration: 23 September 2021.

Keywords Agility, Exercise, Fatigue, Multiple sclerosis, Rehabilitation

Background

Multiple Sclerosis (MS) is the most common, non-traumatic, neurological disorder among middle aged adults. Initially characterized as an inflammatory demyelinating disease of the central nervous system, neurodegenerative processes lead to progressive disability during later stages [1]. In Germany, persons with MS (pwMS) frequently attend inpatient rehabilitation facilities for several weeks to improve their ability to work in a multidisciplinary setting [2]. 'Visible' symptoms such as mobility impairments play an obvious role in pwMS's ability to participate in the job market. However, 25% of pwMS are limited in their professional participation due to 'invisible' symptoms such as fatigue [3, 4].

The definition and conceptualization of fatigue has been changing and expanding for years, including recent updates [5, 6]. For the purpose of this study the term 'fatigue' refers to the 'subjective sensation of lack of energy and exhaustion' (p. E79) [7], retrospectively self-reported for a period of at least one week by a pwMS (i.e., the trait component of fatigue). The term 'fatigability' refers to objectively measured performance decrements on motor or cognitive tasks, corresponding to the taxonomy of Kluger et al. [8].

Contrary to its impact, pharmacological treatment options for fatigue are limited [9]. Consequently, many exercise and behavioral interventions have been evaluated [10]. One of the results concerning exercise is that endurance exercise, although frequently investigated, seems to have only a small effect on fatigue [11, 12]. Interventions broadly focused on 'balance' are less prevalent, but potentially with a more pronounced effect [10, 11]. Among exercise studies that explicitly addressed fatigue, almost none were conducted in an inpatient rehabilitation setting [11], which is characterized by a multidisciplinary environment, including various diagnostic and therapeutic components such as exercise, occupational therapy, health education, or neuropsychological assessment and training. Additionally, interactions between treatments as well as flexibility in the treatment schedule

are common [13, 14]. This leaves clinical practice with few results that could be applied directly to this setting.

We have recently described a group-based exercise training framework for pwMS (multimodal agility-based exercise training [MAT] [15]), which might comprise several aspects that have been proposed to be beneficial for fatigue reduction, e.g., (I) balance training for making 'navigating the environment' less effortful [11], (II) 'coordination of eye, head, and whole-body movements' to 'reduce the cognitive load associated with conscious compensatory strategies in dynamic environments' [16], and (III) 'improvement of sensory integration with a subsequent reduction of the cognitive load associated with motor processing' [17]. As the MAT approach also includes other aspects suitable for inpatient rehabilitation (i.e., group-based, applicable to other neurological conditions [18]) and is proposed to be beneficial for several symptoms, including fatigability [15], the ReFEx (Rehabilitation, Fatigue, and Exercise) project aims to compare MAT with a 'traditional' exercise approach, namely, strength and endurance training (SET) [19] during inpatient rehabilitation. Both, strength and endurance training can be considered standard elements in neurorehabilitation facilities in Germany (and for the current clinic) and are part of national-level MS exercise guidelines [20].

In a first step, the present feasibility study was conducted to determine whether all aspects of the trial were implementable in the clinical setting, and to inform the potential progression to a powered randomized controlled trial (RCT). Secondarily, preliminary clinical outcome data was acquired.

Methods
Design and setting
The study was located at the Neurological Rehabilitation Center (NRC) Godeshoehe GmbH in Bonn, Germany, which provides neurorehabilitation for all levels of disability. It is certified by the German MS Society as one of three MS rehabilitation centers in the state of North Rhine-Westphalia and treats around 120 pwMS each

Wolf *et al. BMC Neurology*   (2023) 23:388

Page 3 of 12

year, 75% coming from within-state. All study-related sessions were implemented within existing therapy services of the NRC.

The study had a two-armed, parallel-group, randomized-controlled design with 12 weeks follow-up (Fig. 1), pursuing a mixed-methods approach. The qualitative part will be reported elsewhere. We intended to recruit 12 participants per group [14], but no sample size calculation was performed as the feasibility evaluation was the primary aim.

Ethical approval was obtained from the Ethics Committee of the University of Bonn (reference number: 543/20). The study was prospectively registered in the German Clinical Trials Register (ID: DRKS00023943) on 23rd September 2021. For more details we refer to the published feasibility protocol [14].

Screening and recruitment

New admissions were screened for MS and eligibility criteria were evaluated in a joint effort by the treating neuropsychologists (JN, JS, EH) and the principal investigator (FW). Inclusion criteria were a relapsing-remitting or secondary-progressive disease course (2017 McDonald criteria [21, 22]), age between 18 and 67 years (age for retirement in Germany), Expanded Disability Status Scale (EDSS)≤5.0 [23], Fatigue Scale for Motor and Cognitive Functions (FSMC)≥53 (cut-off for 'moderate fatigue') [24], and written informed consent. Exclusion criteria included the inability to attend aquatic therapy, comorbidities, that prevented attending study sessions, chronic neurologic conditions other than MS, insufficient German language skills, and specific fatigue medication (Amantadine, Modafinil) started less than 3 months ago. If deemed eligible, pwMS were informed about the study verbally and in written form.

Randomization and blinding

After written informed consent, pwMS were randomly allocated (1:1) to MAT or SET, according to the minimization procedure [25], stratified by EDSS (≤3 or >3), Würzburg Fatigue Inventory for Multiple Sclerosis (WEIMuS, <38 or ≥38) [26], age (<45 or ≥45), and MS disease course (relapsing-remitting or secondary-progressive). The WEIMuS acted as a stratification factor, as it was the potential primary outcome for a future RCT. Randomization was provided by an independent researcher from the German Sport University Cologne using RITA ('Randomization-In-Treatment-Arms', Evident, Germany).

The neuropsychological staff conducting the cognitive tests were blinded to the study groups. Participants, therapists, and staff conducting the physical tests and analyzing the questionnaires were not blinded regarding group allocation. However, participants were blinded regarding which of the groups was the experimental condition.

Interventions

The intervention period lasted from admission (T_0) to discharge (T_1), comprising 4 to 6 weeks (based on medical indications, determined by the treating physician). MAT-participants performed five 30 min sessions of gym-based MAT, and three 30 min sessions of pool-based MAT per week in a group setting (including other neurological patients). SET comprised five 22 min sessions of endurance training on cycle ergometers, and three 30 min sessions of individual strength training. Endurance training was provided on cycle ergometers and not on a treadmill to enable more pronounced differences regarding the demand for sensory integration between MAT and SET [27]. Furthermore, it is the standard modality for endurance training in this clinic. Importantly, participants from both groups also attended a group on body awareness and relaxation techniques, which is part of usual care for pwMS in this clinic and which provided some social contact in the SET-group as well [14].

MAT consisted of three components: (I) standing balance exercises, (II) dynamic balance exercises including functional leg strength, and (III) agility-like exercises. Agility-like exercises have been defined as '[…] tasks, that require changes of direction, stop-and-go patterns, turns, and changing footwork strategies, with or without responding to a stimulus' [15]. For load management in the gym-setting, three sessions with higher physical demands (i.e., agility-like components and functional leg strength) were interspersed with two sessions of lower physical strain (i.e., standing balance and exercises with a cognitive focus).

In SET, endurance training was performed with 3 min of gradual increase, 17 min steady and 2 min cool-down on a cycle ergometer (ergoselect 5, ergoline GmbH, Bitz, Germany) with continuous monitoring of power output (W) and heart rate (ers.2 software, ergoline GmbH, Bitz, Germany). The first session started with a power output participants had rated 'light' (=11) to 'somewhat hard' (=13) (6–20 Rating of Perceived Exertion [RPE] – scale) during the baseline graded exercise test (GXT, see Sect. 2.5.2 and Supplement) and then continued within this range. In detail, in each of the cycling sessions, therapists asked participants to rate their perceived exertion on the 6–20 RPE scale after 8 min and/or 15 min of cycling. Two timepoints were chosen in case perceived exertion changed during the session. If pwMS gave two different RPE ratings for each timepoint, the average score was documented. Therapists regulated the power output so that participants stayed between 11 and 13 on the RPE-scale. The new session always continued with the training load from the previous session. The range of 11–13 was chosen based on recent evidence-based recommendations for pwMS with similar EDSS [19].
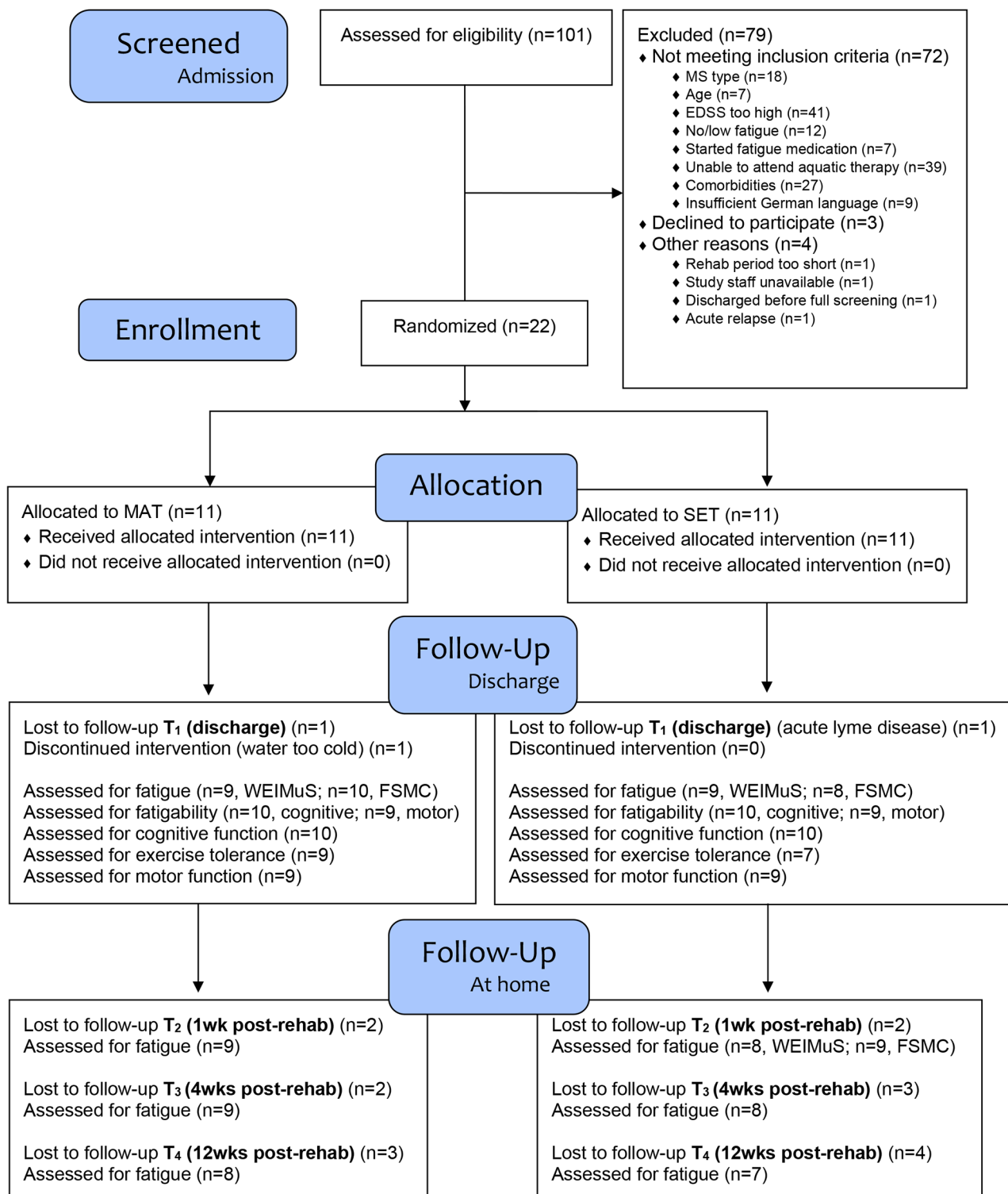
Screened
Admission

Assessed for eligibility (n=101)

Excluded (n=79)
- Not meeting inclusion criteria (n=72)
  - MS type (n=18)
  - Age (n=7)
  - EDSS too high (n=41)
  - No/low fatigue (n=12)
  - Started fatigue medication (n=7)
  - Unable to attend aquatic therapy (n=39)
  - Comorbidities (n=27)
  - Insufficient German language (n=9)
- Declined to participate (n=3)
- Other reasons (n=4)
  - Rehab period too short (n=1)
  - Study staff unavailable (n=1)
  - Discharged before full screening (n=1)
  - Acute relapse (n=1)

Enrollment

Randomized (n=22)

Allocation

Allocated to MAT (n=11)
- Received allocated intervention (n=11)
- Did not receive allocated intervention (n=0)

Allocated to SET (n=11)
- Received allocated intervention (n=11)
- Did not receive allocated intervention (n=0)

Follow-Up
Discharge

Lost to follow-up T_1 **(discharge)** (n=1)
Discontinued intervention (water too cold) (n=1)

Assessed for fatigue (n=9, WEIMuS; n=10, FSMC)
Assessed for fatigability (n=10, cognitive; n=9, motor)
Assessed for cognitive function (n=10)
Assessed for exercise tolerance (n=9)
Assessed for motor function (n=9)

Lost to follow-up T_1 **(discharge)** (acute lyme disease) (n=1)
Discontinued intervention (n=0)

Assessed for fatigue (n=9, WEIMuS; n=8, FSMC)
Assessed for fatigability (n=10, cognitive; n=9, motor)
Assessed for cognitive function (n=10)
Assessed for exercise tolerance (n=7)
Assessed for motor function (n=9)

Follow-Up
At home

Lost to follow-up T_2 **(1wk post-rehab)** (n=2)
Assessed for fatigue (n=9)

Lost to follow-up T_3 **(4wks post-rehab)** (n=2)
Assessed for fatigue (n=9)

Lost to follow-up T_4 **(12wks post-rehab)** (n=3)
Assessed for fatigue (n=8)

Lost to follow-up T_2 **(1wk post-rehab)** (n=2)
Assessed for fatigue (n=8, WEIMuS; n=9, FSMC)

Lost to follow-up T_3 **(4wks post-rehab)** (n=3)
Assessed for fatigue (n=8)

Lost to follow-up T_4 **(12wks post-rehab)** (n=4)
Assessed for fatigue (n=7)

Fig. 1 CONSORT diagram. EDSS=Expanded Disability Status Scale; FSMC=Fatigue Scale for Motor and Cognitive Functions; MAT=multimodal agility-based exercise training; MS=multiple sclerosis; SET=strength and endurance training; WEIMuS= Würzburg Fatigue Inventory for Multiple Sclerosis

Strength training was adapted from published protocols [17]: each session started with a 5 min warm-up, followed by three to four lower extremity exercises. Sessions 1 through 5 included 3×10 repetitions (intensity: 15 repetitions maximum [RM]) and Session 6 through discharge included 3×12 repetitions (12RM). More details and treatment manuals are displayed in the protocol [14].

Outcomes
Feasibility
Feasibility domains were based on Thabane et al. [28] and evaluated according to prespecified progression criteria (Table 3). Evaluation of *processes* included the eligibility, recruitment, refusal, and retention rates, intervention adherence and fidelity. *Resources*-related outcomes focused on the number of days needed to complete baseline assessments, time requirements for the physical testing blocks at T_0 and T_1, and for the preparation of MAT sessions. *Data management* feasibility determined missing items from the WEIMuS and FSMC and missing assessments at T_0 and T_1. The *scientific* domain evaluated adverse events and perceived exertion. Adverse events were defined as events related to the interventions that led to early termination of a session. Perceived exertion was based on the session-RPE scale (0–10 scale, where 0 indicates 'rest', 10 indicates 'maximum' intensity) [29, 30]. See Table 4 in the protocol for details and rationales [14] and the Supplement for specifics of session-RPE application.

Six participants from each study arm were interviewed face-to-face at T_1 regarding feasibility objectives (will be reported elsewhere).

Potential clinical outcomes
Perceived fatigue was assessed with the WEIMuS [26] and FSMC [24] questionnaires at T_0 and T_1. Participants were followed up via e-mail to fill out online versions 1, 4, and 12 weeks after discharge (T_2-T_4, Fig. 1). Change in WEIMuS total score from T_0 to T_2 was evaluated as the potential primary endpoint for a future RCT [14].

Other potential clinical outcomes for a definitive RCT were assessed only at admission and discharge and included cognitive fatigability (circadian changes in tonic alertness measures, assessed with the Test Battery of Attention Performance – Alertness [TAP-Alert] [31, 32] before 11 a.m. and after 3 p.m.), motor fatigability (Distance Walked Index [DWI] [33]), cognitive performance (California Verbal Learning Test [CVLT], Symbol Digit Modalities Test [SDMT] [34]), exercise tolerance (GXT on a cycle ergometer, protocol: start 25 W, progression 10 W/min, for details, see [35] and Supplement), motor function (6-Minute Walk Test [6MWT] [36], Timed 25-Foot Walk Test [T25FW] [37], Six Spot Step Test [SSST] [38], Functional Gait Assessment [FGA] [39]),

and balance confidence (Activities-specific Balance Confidence scale [ABC] [40]).

Data analysis
Descriptive statistics were used to summarize baseline sample characteristics, feasibility, and clinical outcomes, using IBM SPSS Statistics 29. Baseline differences between groups were examined using independent samples t-tests for normally distributed continuous variables or Mann-Whitney U tests for non-normally distributed and ordinal variables, and Fisher's Exact test for categorical variables, with $p<0.05$ indicating significant differences.

Change scores from baseline were calculated for clinical outcomes for each of the measurement timepoints, as was the frequency of participants in each group with a relevant improvement related to the WEIMuS ($\leq$-6) [14] and FSMC ($\leq$-10) [41] total scores at T_2. Since this was a small-scale feasibility study, hypothesis testing of within- or between-group treatment effects was not performed [42, 43]. For the same reasons, no effect sizes were estimated [44]. However, we compared the feasibility data to the prespecified progression criteria (Table 3).

Results
Participants
Flow of participants is depicted in the CONSORT diagram (Fig. 1). Due to maintenance work of the pool starting in October 2022 we had to reduce the sample size from 24 to 22 participants. Baseline sociodemographic and clinical characteristics are described in Table 1.

Feasibility
Results regarding the *a priori* defined progression requirements are shown in Table 3.

Processes
Twenty-five of 101 (26%) patients screened were eligible and three of 25 (12%) declined to participate. Two were interested but overwhelmed with their current situation or wanted to focus on their own primary goals. One declined because not wanting to be restrained to one intervention. 'EDSS' and 'able to attend aquatic therapy' produced the most negative cases regarding eligibility, followed by comorbidities and disease course (Fig. 1, Table S1). It took 12 months (11/2021 to 11/2022) to randomize 22 participants, equaling 1.8 randomizations per month, which is below the progression requirement (Table 3).

Retention between T_0 and T_1 was 91% for each group. One MAT-participant dropped out during the intervention period, because of the pool being too cold. One SET-participant was excluded from the follow-up analysis as he developed acute lyme disease and was unable to attend

Table 1 Baseline sociodemographic and clinical characteristics

	MAT (n = 11)	SET (n = 11)	*p*-value
Age mean (SD, min-max)	45.6 (10.1, 26–56)	53.3 (9.3, 31–64)	**0.019**[a]
Sex f:m	10:1	8:3	0.586[b]
BMI mean (SD, min-max)	25.0 (5.0, 20.0–36.9)	28.9 (7.6, 21.6–48.0)	0.116[a]
Work status (n)	• Unfit for work (1) • Retired (3) • 3-6 h/d (3) • > 6 h/d (4)	• Unfit for work (1) • Retired (5) • 3-6 h/d (3) • > 6 h/d (2)	0.921[b]
MS type RR:SP	9:2	9:2	1.000[b]
TSD mean (SD, min-max)	9.5 (7.0, 2–28)	9.7 (8.6, 0–27)	0.699[a]
EDSS median (min-max)	3.0 (1.5–4.5)	2.5 (2.0–4.5)	0.748[a]
Walking device (n)	0	0	1.000[b]
CES-D mean (SD, min-max)	24.7 (11.7, 4–40)	25.7 (8.1, 13–39)	0.818[c]
DMT (n)	• None (4) • Glatiramer acetate (1) • Natalizumab (1) • Ofatumumab (1) • Teriflunomide (1) • Cladribine (1) • Fingolimod (1) • Siponimod (1)	• None (2) • Glatiramer acetate (2) • Natalizumab (1) • Ofatumumab (1) • Teriflunomide (2) • Dimethyl fumarate (1) • Interferone beta-1b (1) • Ocrelizumab (1)	0.929[b]
Fatigue medication (n)	• Amantadine (1)[*] • None (10)	• None (11)	1.000[b]

[a]Mann-Whitney U test, [b]Fisher's Exact test, [c]Independent samples t-test, bold=significant difference between groups. BMI=Body Mass Index; CES-D=Centre for Epidemiological Studies Depression Scale (German version); DMT=disease-modifying treatment; EDSS=Expanded Disability Status Scale; f=female; m=male; MAT=multimodal agility-based exercise training; max=maximum value; min=minimum value; MS=multiple sclerosis; n=number of patients; RR=relapsing-remitting multiple sclerosis; SD=standard deviation; SET=strength and endurance training; SP=secondary progressive multiple sclerosis; TSD=time since diagnosis in years

[*]this patient had started taking Amantadine more than three months ago and therefore, was not excluded

Table 2 Adherence results are shown for MAT, SET and both groups combined ('total'). Results are also given separately for gym/pool sessions and strength/endurance sessions

	Total n = 20	MAT n = 10	Gym n = 10	Pool n = 10	SET n = 10	Strength n = 10	Endurance n = 10
No. appointments (completed/scheduled (rate attended))	535/596 (90%)	228/270 (84%)	148/165 (90%)	80/105 (76%)	308/327 (94%)	117/123 (95%)	191/204 (94%)
Average completed sessions/participant (min-max)	26.8 (12–38)	22.7 (12–33)	14.8 (9–20)	8.0 (3–14)	30.8 (18–38)	11.7 (8–14)	19.1 (10–25)

MAT=multimodal agility-based exercise training; SET=strength and endurance training

most of the sessions and follow-up assessments. At T_2, nine (82%, MAT-group) and eight (73%, SET-group) participants completed the WEIMuS, respectively. Thus, retention-related progression requirements were mostly fulfilled (Table 3).

Average length of stay in the rehabilitation facility were 31.6 (SD=5.2, min-max=25–41, n=10) full days for the MAT-group and 34.2 (SD=6.2, min-max=22–41, n=10) full days for the SET-group and both groups managed to attain the required minimum number of sessions (Table 3). However, on average, the SET-group performed more sessions than the MAT-group and adherence for the pool-based training was lower (76%) than for all other sessions (90–95%) (Table 2).

Regarding fidelity, a total of 122 gym-based and 76 pool-based MAT-sessions were logged and analyzed for MAT-components, as noted by the respective therapists.

In an average week, 18.5%/17.9% (gym/pool) of training content targeted standing balance, 46.2%/50.1% dynamic balance/functional leg strength, and 35.3%/30.2% targeted agility-like exercises, showing that therapists provided all three MAT components. Average heart rate during all tracked gym-based sessions was 93.7 bpm (SD=11.3, min-max=78.8-114.1, n=11), average maximum heart rate was 116.9 bpm (SD=11.8, min-max=99.1-132.3, n=11).

In SET, strength sessions included an average 2.8 exercises (goal: three), and participants performed an average 99.5% of the prescribed load. Endurance sessions lasted for an average 21min26s (goal: 22 min) and average heart rate corresponded to 107.1 bpm (SD=14.4, min-max=90.5-127.9, n=10). The average prescribed training intensity for cycling sessions was 57.5 W (SD=24.2, min-max=35–120, n=10), while average actual load was

Table 3 *A priori* progression requirements [14] and results

Requirement	Results
quantitative	
1. Adherence	
Average of at least 18 therapy sessions completed during the stay per group	+ MAT: 22.7 (12–33)
	+ SET: 30.8 (18–38)
2. Recruitment	
4 participants/month	− 1.8/month
<25% non-eligible pwMS	− 74% non-eligible
<10% eligible but unwilling to participate	− 12%
3. Retention	
$T_1 > 90\%$ per group	+91% (both groups)
$T_2 > 80\%$ per group	+82% (MAT)
	− 73% (SET)
4. Time	
>80% able to complete all baseline assessments within the first 3 days of therapy	− 50%
qualitative	
5. Interviews	
Statements indicate that the interventions and study processes are acceptable	There were no major acceptability issues. Still, some adaptations to the study protocol were identified and will be reported separately.

Adherence data is presented as mean (min-max); + =requirement fulfilled; - =requirement not fulfilled; MAT=multimodal agility-based exercise training; pwMS=persons with multiple sclerosis; SET=strength and endurance training; T_1=discharge; T_2=one-week post-discharge

53.5 W (SD=34.2, min-max=24.5-139.6, n=10), corresponding to an average 88% completion of the prescribed load (SD=20, min-max=54–116, n=10). Average RPE (6-20) during cycling was 12.8 (SD=0.5, min-max=11.7–13.2, n=10) (goal: 11–13).

Resources
Days needed to complete the baseline assessments were 4.1 (SD=1.5, min-max=3–9, n=22) and 50% of participants completed all assessments within the first 3 days of therapy, which is below the required 80% (Table 3).

At T_0 and T_1, average time requirements for the physical assessments were 45 min (T_0) and 43 min (T_1) for motor function, and 30 min (T_0) and 33 min (T_1) for the GXT. To prepare MAT-sessions, therapists needed an average 3.7 min (gym-based, min-max=1–12, n=102 sessions) and 2.6 min (pool-based, min-max=1–12, n=62 sessions).

Data management
Data management revealed no missing items for WEIMuS and FSMC questionnaires and no missing assessments at T_0. At T_1, 7/22 (32%) participants had at least one missing assessment, with the GXT missing the most (six participants).
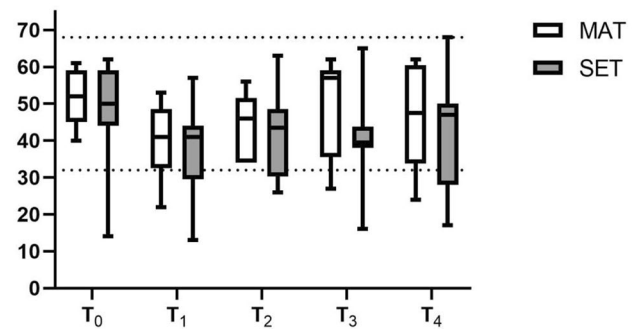

WEIMuS

Fig. 2 WEIMuS total scores for both groups. Lower scores indicate less fatigue. Box plot: line=median, whiskers=min-max. Upper dotted line=maximum WEIMuS total score (=68); lower dotted line=cut-off for fatigue (=32); MAT = multimodal agility-based exercise training; SET = strength and endurance training; T_0=admission; T_1=discharge; T_2=1 week post-discharge; T_3=4 weeks post-discharge; T_4=12 weeks post-discharge

Scientific
No adverse events occurred in the MAT-group. During the cycling sessions, therapists noted six minor adverse events occurring in three participants (knee pain, severe fatigue, dizziness, low blood pressure).

Average session-RPE was 4.7 (gym-based, min-max=2.3–6.8, n=11, 141 sessions), 3.6 (pool-based, min-max=2.3–5.4, n=11, 61 sessions), 4.0 (strength, min-max=1.7–5.4, n=10, 105 sessions), and 3.8 (endurance, min-max=2.1–5.8, n=9, 136 sessions), respectively. According to the session-RPE scale a score of 3 indicates 'moderate' intensity, 4 is 'somewhat hard', and 5 is 'hard' [30].

Clinical outcomes
Fatigue
Both groups showed a comparable median reduction in WEIMuS total scores at T_1 and T_2 (Fig. 2, scores are reported in Table S2 as Supplement), with the MAT group having a higher percentage of participants with a relevant improvement at T_2 (7/9, 78% [MAT] vs. 5/8, 63% [SET]). At T_3, the MAT group displayed a sharp rise in scores, which dropped again at T_4. From T_0 to T_4, individual WEIMuS trajectories showed an increasing variability for participants with a full data set (Fig. 3).

FSMC total scores are reported in the Supplement (Figure S1, Table S2). The proportion of participants with a relevant improvement at T_2 was considerably lower than for the WEIMuS (0/9, 0% [MAT] vs. 3/9, 33% [SET]).

Fatigability
Of the whole sample, three MAT- and two SET-participants revealed clinically relevant (i.e., at least −10% decrease in meters walked between the first minute of
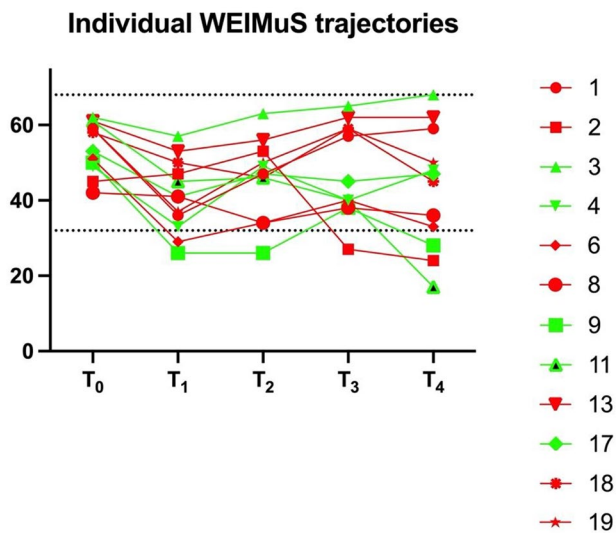
Individual WEIMuS trajectories

Fig. 3 Individual WEIMuS total score trajectories for participants with a full data set from T_0 to T_4 (n = 12). A decrease in scores indicates less fatigue. Red = multimodal agility-based exercise training; green = strength and endurance training; upper dotted line = maximum WEIMuS total score (68); lower dotted line = indicates cut-off for fatigue [32]; T_0 = admission; T_1 = discharge; T_2 = 1 week post-discharge; T_3 = 4 weeks post-discharge; T_4 = 12 weeks post-discharge

the 6MWT and the last minute) [33] walking fatigability at T_0 and T_1, respectively. Both groups improved their DWI (i.e., less drop-off in meters walked between minute 1 and 6, Table 4) with several participants showing substantial improvements (e.g., +14.6%). However, as most participants also walked further at T_1, some also showed a worse DWI at T_1.

Cognitively, tonic alertness reaction times increased in both groups between morning and afternoon assessments at T_0 and T_1, reflecting cognitive fatigability [45]. Unexpectedly, in both groups, afternoon alertness measures deteriorated between admission and discharge, as did the difference between morning and afternoon assessments. Only in the SET reaction times in the morning were faster at T_1. Overall, variability was high.

Other clinical outcomes

Peak power output in the GXT increased in both groups, with higher change scores in SET. Validity criteria [46] for cardiorespiratory fitness testing were mostly not attained, except for perceived exertion (Supplement). Therefore, we changed our terminology to 'exercise tolerance' [47].

All motor function measures also increased in both groups. For the 6MWT, improvements (43.9 m [MAT], 26.3 m [SET]) exceeded measurement error on group-level (≥11.1 m) and in the MAT also substantially exceeded minimal important change values for patients with mild disability (≥26.9 m) [48].

Discussion

This study examined the feasibility of a trial comparing a 'new' group-based exercise framework for pwMS (MAT) [15] with 'traditional' exercise (SET) in an inpatient rehabilitation setting. Main clinical outcomes of interest for a definitive RCT were fatigue and fatigability [8].

Among the four predefined quantitative progression categories, one category was fulfilled (adherence), one was mixed, but approached a positive value (retention), and two were negative (recruitment, time). Therefore, changes to the study design are necessary.

Descriptively, favorable changes in fatigue were observable in both groups at the end of rehabilitation. Mobility related outcomes also improved in both groups, with the most pronounced changes in walking endurance, where MAT elicited clinically relevant changes.

Feasibility

Advantages of inpatient rehabilitation include the high frequency of exercise sessions and high retentions rates, as patients are on-site. However, pwMS are also embedded in a multidisciplinary setting, with study-unrelated appointments possibly interfering with study-related sessions. Adherence rates were lower than the ones described by Zimmer et al. [49] (100%, n=57) in a similar setting, but this trial also included fewer study-related sessions/week (3–5/week for 3 weeks vs. 8/week for >4 weeks). Thus, the present results on adherence still indicate that a high frequency of sessions was possible on an organizational level and regarding the capacities of this (fatigued) patient collective. One exception was the pool session (only 76% adherence), which should be moved to a spot with a longer break from lunch, as indicated by pwMS in the interviews.

Another option would be to cut the pool-based sessions to (I) allow for more time to recover, (II) increase recruitment (as eligibility for pool-based sessions was low), (III) avoid the mix-up of effects from gym- and pool-based MAT, and (IV) lower the barrier for replication, without the need for a pool.

Even though both groups performed the minimum number of sessions, the SET-group attended substantially more sessions (228 [MAT] vs. 308 [SET]), as appointments were much more flexible due to individual scheduling. Consequently, MAT had considerably lower 'dosing'. This is interesting, as one hypothesis of the MAT framework is that it might be more time-efficient [15]. Still, to evaluate differences in treatment effects it is important to ensure similar amounts of sessions performed, first [50].

Only about 1.8 participants were randomized per month instead of the intended four. There are examples of other trials conducted in similar settings, reporting equally low eligibility [49, 51]. If 'able to attend aquatic

Table 4 Descriptive data on clinical outcomes at admission, discharge, and change scores (T_1-T_0).

Domain	Outcome	MAT		SET	
		Mean (SD)	Min-max	Mean (SD)	Min-max
Fatigability	**DWI (%)**				
	T_0[a]	-7.8 (6.5)	-17.6–0.0	-7.6 (5.8)	-20.0–3.3
	T_1[a]	-6.5 (5.2), n = 9	-15.5–0.0	-5.4 (5.3), n = 9	-11.6–3.1
	T_1-T_0 (change) ↑	1.7 (8.5)	-9.8–14.6	3.4 (4.3)	-2.0–9.6
	TAP-Alertness (afternoon-morning difference, ms)				
	T_0[b]	12.2 (60.2)	-108–119	8.2 (79.7)	-62–240
	T_1[b]	36.2 (136.3), n = 10	-97–364	10.6 (50.7), n = 10	-26–148
	T_1-T_0 (change) ↓	22.6 (130. 9)	-154–332	25.6 (56.2)	-19–167
Cognitive performance	**CVLT (n words)**				
	T_0[c]	56.9 (10.7)	37–71	53.0 (10.0)	28–65
	T_1[c]	53.0 (12.5), n = 10	38–74	53.7 (11.3), n = 10	36–70
	T_1-T_0 (change) ↑	-3.9 (7.5)	-18–3	0.8 (11.4)	-20–22
	SDMT(n pairs)				
	T_0[d]	49.7 (9.1)	30–62	49.2 (9.3)	35–63
	T_1[d]	52.9 (10.6), n = 10	36–67	52.0 (8.8), n = 10	38–69
	T_1-T_0 (change) ↑	1.2 (5.8)	-10–9	3.6 (6.1)	-5–13
Exercise tolerance	**GXT (W_{peak})**				
	T_0[c]	106.8 (37.1)	55–175	110.5 (42.7)	55–225
	T_1[c]	101.7 (32.8), n = 9	65–165	130.7 (51.6), n = 7	95–245
	T_1-T_0 (change) ↑	2.2 (6.7)	-10–10	14.3 (16.1)	0–40
Motor function & confidence	**6MWT (m)**				
	T_0[c]	499.0 (90.1)	366–631	511.5 (89.6)	315–648
	T_1[c]	544.8 (101.9), n = 9	399–700	537.1 (87.0), n = 9	365–660
	T_1-T_0 (change) ↑	43.9 (30.0)	-25–77	26.3 (32.8)	-14–81
	T25FW (s)				
	T_0[d]	5.20 (0.99)	4.05–7.10	4.98 (0.98)	3.90–7.05
	T_1[d]	5.01 (0.79), n = 9	3.95–6.5	4.80 (0.55), n = 9	4.35–6.10
	T_1-T_0 (change) ↓	-0.10 (0.40)	-0.61–0.65	-0.30 (0.61)	-1.10–0.50
	SSST (s)				
	T_0[d]	8.19 (2.42)	4.95–11.37	7.48 (1.80)	5.52–11.62
	T_1[d]	7.55 (2.21), n = 9	4.91–10.72	6.58 (1.33), n = 9	4.90–8.85
	T_1-T_0 (change) ↓	-0.51 (1.12)	-2.45–1.52	-0.92 (1.07)	-2.77–0.43
	FGA(total score)				
	T_0[c]	23.0 (18.0–26.0)[*]	14–28	24.0 (23.0–27.0)[*]	12–28
	T_1[c]	24.0 (19.0–26.5), n = 9	14–30	25.0 (23.0–28.0), n = 9	15–29
	T_1-T_0 (change) ↑	2.0 (-1.5–4.5)	-4–6	2.0 (0.0–2.0)	-1–3
	ABC (total score)				
	T_0[c]	70.5 (21.3)	24.1–99.1	75.8 (14.9)	55.0–95.6
	T_1[c]	67.8 (19.4), n = 10	33.4–97.2	72.1 (13.2), n = 8	57.5–96.3
	T_1-T_0 (change) ↑	0.2 (7.2)	-12.5–10.3	-2.6 (6.8)	-16.4–5.6

All outcomes had n = 11 at T_0, sample size for each outcome at T_1 is reported in the table. All scores are presented as mean (SD, min-max), except for the FGA, presented as median (interquartile range). Arrows indicate direction for improvement. [a] = scores closer to zero indicate less fatigability; [b] = lower scores indicate smaller increase in fatigability from morning to afternoon; [c] = higher scores are favorable; [d] = lower scores are favorable; [*]FGA scores are presented as median (interquartile range); 6MWT = 6-Minute Walk Test; ABC = Activities-specific Balance Confidence scale; CVLT = California Verbal Learning Test; DWI = distance walked index; FGA = Functional Gait Assessment; GXT = graded exercise test; MAT = multimodal agility-based exercise training; SDMT = Symbol Digit Modalities Test; SET = strength and endurance training; SSST = Six Spot Step Test; T25FW = Timed 25-Foot Walk Test; TAP = Test Battery of Attention Performance; T_0 = admission; T_1 = discharge; W_{peak} = peak power output

therapy' and 'relapsing-remitting/secondary-progressive disease course' would be excluded as criteria, eligibility could be increased from 26 to 35%. Still, this would not be sufficient. A preliminary sample size calculation for a clinically relevant difference regarding the WEIMuS retrieved a sample size of n = 66, which would take about 46 months to recruit with the present results, or about 32 months with the adapted inclusion criteria. A multicenter

trial might be an option to progress in a more efficient amount of time. A limitation of the eligibility assessment could have been the COVID-19 pandemic, which might have resulted in fewer applications for rehabilitation from pwMS with lower EDSS, because of COVID-19 restrictions.

Interestingly, gym-based MAT elicited the highest ratings of session-based perceived exertion, which on

average approached the rating of 'hard' - despite the fact, that average heart rate values were low. This warrants several considerations. First, the present heart rate values only give a very broad impression regarding intensity as they do not take maximum heart rate into account. As age differed between groups, heart rates indicating relative intensity would also be different. Furthermore, it is important to consider that even 'high physical strain' gym-based MAT sessions have many standing breaks, e.g., while the therapist gives demonstration and instructions, which profoundly reduces average heart rate for a complete session and leads to a more interval-like training stimulus. Values of average maximum heart rate might also be misleading, as 'low physical strain' MAT sessions blunted maximum values from the 'high physical strain' sessions. For example, four MAT-participants had individual gym-sessions with maximum heart rates above 150 bpm. Lastly, what we can only describe anecdotally is that MAT-participants repeatedly differentiated their session-based perceived exertion between physical and cognitive exertion (but were then prompted to give an overall score). Supposedly, this could reflect the unique content of MAT, and cognitive elements might increase session-RPE scores.

Clinical outcomes

Regarding the performed assessments, this study is among the first which broadly focused on balance/motor control and conducted several follow-up measurements for fatigue [10]. Retention from end-of-treatment to 3 months follow-up was acceptable in both groups (78–89%), suggesting the online fatigue assessment to be a viable option for future follow-up assessments.

Observed (descriptive) changes in fatigue (i.e., WEIMuS) from admission to discharge/1 week after discharge were higher than for short-term treatment with fampridine, for example [52]. The reason for the rise in scores at T_3 for MAT is unclear, but could be attributed to the small sample size, as scores dropped again at T_4.

Proper assessment of fatigue seems to be one of the biggest challenges regarding the development of treatment strategies. The WEIMuS and FSMC differed in their classification of responders 1 week after discharge and clinically relevant change values have not been rigorously determined for these measures [7]. Therefore, switching to one of the 'new generation' questionnaires (e.g., PROMIS Fatigue (MS) 8a [53]) might be considered for a future trial [53, 54].

Similarly, no gold standard exists to quantify walking fatigability in pwMS [55]. Assessment of DWI is one option and was easily implementable in the present setting. Nevertheless, instrumented walking fatigability assessment has been advocated recently as another option and might be selected for a future RCT [55, 56].

Unexpectedly, cognitive fatigability tended to be worse at T_1. A reason for this might be that the cognitive load between the morning/afternoon assessments was not rigorously standardized. In addition, it has been shown that a number of other disease-specific factors (e.g., daytime sleepiness, hand motor impairment), personality-related behavioral factors in dealing with performance situations (e.g., energy management, test anxiety), or psychometric aspects of test use (e.g., practice effects) could contribute directly or indirectly to this relationship, complicating interpretation [57].

Limitations

As mentioned, advantages of inpatient rehabilitation are the high volume and frequency of exercise sessions and high retentions rates. However, some general disadvantages of the inpatient rehabilitation setting should be reconsidered. First, the intervention duration is restricted and, as MAT is a group-based, therapist-led intervention (i.e., no home-based or digital training) [15], the setting precludes the possibility of evaluating long-term MAT effects. If MAT is framed as a reserve-building activity long-term studies are needed [15, 58], which would favor an outpatient setting. Second, the multidisciplinary setting might affect some of the outcome measures. As it would be unethical to withhold a certain kind of therapy, patients can receive, for example, varying degrees of complementary computerized cognitive training, making it challenging to distill the effect of the study-related training. This feeds into the wider discussion on the 'black box' of usual care and treatment components in rehabilitation [59]. Third, the length of stay in a neurorehabilitation facility in Germany is not fixed but is determined during the stay by the treating physician based on medical indications and can vary between 4 and 6 weeks. Therefore, intervention duration was not matched. However, this allows the design to reflect the actual clinical setting and the same follow-up periods post-discharge. Lastly, all results from the present feasibility study, must be seen in light of the small sample size, and potentially 'delivery agent bias' [60], i.e., the newly evaluated intervention (MAT) was partially provided by a developer (FW) [14], which might have increased its effects and compromised blinding.

Conclusions

Substantial changes to the study design are needed, especially to increase recruitment. Going forward, a multicenter trial focused on gym-based MAT might be another option. 'New generation' fatigue questionnaires, instrumented motor fatigability, and alertness assessments with a standardized cognitive load are candidates for improved outcome assessments.

Wolf *et al. BMC Neurology*   (2023) 23:388

Page 11 of 12

List of Abbreviations

6MWT	6-Minute Walk Test
ABC	Activities-specific Balance Confidence scale
CVLT	California Verbal Learning Test
DWI	Distance Walked Index
FGA	Functional Gait Assessment
FSMC	Fatigue Scale for Motor and Cognitive Functions
GXT	Graded exercise test
MAT	Multimodal agility-based exercise training
NRC	Neurological Rehabilitation Center
ReFEx	Rehabilitation, Fatigue, and Exercise
pwMS	Persons with multiple sclerosis
RPE	Rating Perceived Exertion
SET	Strength and endurance training
SDMT	Symbol Digit Modalities Test
SSST	Six Spot Step Test
T25FW	Timed 25-Foot Walk Test
TAP-Alert	Test Battery of Attention Performance – Alertness
WEIMuS	Würzburg Fatigue Inventory for Multiple Sclerosis

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12883-023-03436-8.

> Supplementary Material 1

Authors' contributions

FW, JN, ME and PZ designed the overall study. FW and JN designed the feasibility study. FW, JN, JS and EH implemented the screening and assessment procedures. PK assisted with data collection and strength sessions. FW conducted statistical analysis and drafted the manuscript. All authors read and revised the manuscript.

Data Availability

The raw data were generated at the Neurological Rehabilitation Center Godeshoehe GmbH. The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

Ethical approval was obtained from the Ethics Committee of the University of Bonn (reference number: 543/20). Participants were informed about the study verbally and in written form and all provided written consent. All methods were carried out in accordance with relevant guidelines and regulations in the declaration of Helsinki.

Consent for publication

Not applicable.

Competing interests

FW: none. JN: received honoraria from the MS Society Vienna, Austria, and grants from the Pharos Foundation, Dortmund, Germany JS: none. EH: none. PK: none. ME: none. A-KF: none. HK: none. PZ: received grants from the German Multiple Sclerosis Society NRW.

Author details

[1]Neurological Rehabilitation Center Godeshoehe GmbH, Bonn, Germany
[2]Division of Performance and Health, Institute for Sport and Sport Science, Technical University Dortmund, Dortmund, Germany
[3]Department of Medical Psychology | Neuropsychology and Gender Studies, Center for Neuropsychological Diagnostics and Intervention (CeNDI), Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany

References

1. Thompson AJ, Baranzini SE, Geurts J, Hemmer B, Ciccarelli O. Multiple sclerosis. Lancet. 2018;391(10130):1622–36.
2. Dettmers C, Schmidt R, Jöbges M. Fatigue und fatigability bei Multipler Sklerose – Leistungsbeurteilung. Die Rehabilit. 2020;59:327–9.
3. Kobelt G, Langdon D, Jönsson L. The effect of self-assessed fatigue and subjective cognitive impairment on work capacity: the case of multiple sclerosis. Mult Scler. 2019;25(5):740–9.
4. Nielsen J, Schnell R, Markett S, Saliger J, Arling V, Karbe H. Are invisible MS symptoms (fatigue, cognitive dysfunction, depression) significant in the professional performance assessment of multiple sclerosis patients two years after the inpatient primary evaluation? Z Neuropsychol. 2020;31(1):11–25.
5. Behrens M, Gube M, Chaabene H, Prieske O, Zenon A, Broscheid KC, et al. Fatigue and human performance: an updated Framework. Sports Med. 2023;53(1):7–31.
6. Enoka RM, Almuklass AM, Alenazy M, Alvarez E, Duchateau J. Distinguishing between fatigue and fatigability in multiple sclerosis. Neurorehabil Neural Repair. 2021;35(11):960–73.
7. Sander C, Voelter H-U, Schlake H-P, Eling P, Hildebrandt H. Assessment of fatigue in multiple sclerosis. Neurol Int Open. 2017;1(02):E79–E85.
8. Kluger BM, Krupp LB, Enoka RM. Fatigue and fatigability in neurologic illnesses: proposal for a unified taxonomy. Neurology. 2013;80(4):409–16.
9. Nourbakhsh B, Revirajan N, Morris B, Cordano C, Creasman J, Manguinao M, et al. Safety and efficacy of amantadine, modafinil, and methylphenidate for fatigue in multiple sclerosis: a randomised, placebo-controlled, crossover, double-blind trial. Lancet Neurol. 2021;20(1):38–48.
10. Harrison AM, Safari R, Mercer T, Picariello F, van der Linden ML, White C, et al. Which exercise and behavioural interventions show most promise for treating fatigue in multiple sclerosis? A network meta-analysis. Mult Scler. 2021;27(11):1657–78.
11. Moss-Morris R, Harrison AM, Safari R, Norton S, van der Linden ML, Picariello F, et al. Which behavioural and exercise interventions targeting fatigue show the most promise in multiple sclerosis? A systematic review with narrative synthesis and meta-analysis. Behav Res Ther. 2021;137:103464.
12. Wolf F, Rademacher A, Joisten N, Proschinger S, Schlagheck ML, Bloch W, et al. The aerobic capacity – fatigue relationship in persons with multiple sclerosis is not reproducible in a pooled analysis of two randomized controlled trials. Mult Scler Relat Disord. 2022;58:103476.
13. Voigt-Radloff S, Stemmer R, Korporal J, Horbach A, Ayerle G, Schäfers R et al. Forschung zu komplexen Interventionen in der Pflege- und Hebammenwissenschaft und in den Wissenschaften der Physiotherapie, Ergotherapie und Logopädie: ein forschungsmethodischer Leitfaden zur Ermittlung, Evaluation und Implementation evidenzbasierter Innovationen; Version 2. 2016.
14. Wolf F, Nielsen J, Saliger J, Hennecken E, Eschweiler M, Folkerts A-K et al. Randomised controlled pilot and feasibility study of multimodal agility-based exercise training (MAT) versus strength and endurance training (SET) to improve multiple sclerosis-related fatigue and fatigability during inpatient rehabilitation (ReFEx): study protocol. BMJ Open. 2022;12(9).
15. Wolf F, Eschweiler M, Rademacher A, Zimmer P. Multimodal Agility-based Exercise training for persons with multiple sclerosis: a New Framework. Neurorehabil Neural Repair. 2022;36(12):777–87.
16. Hebert JR, Corboy JR, Vollmer T, Forster JE, Schenkman M. Efficacy of Balance and Eye-Movement exercises for persons with multiple sclerosis (BEEMS). Neurology. 2018;90(9):e797–e807.
17. Callesen J, Cattaneo D, Brincks J, Kjeldgaard Jørgensen ML, Dalgas U. How do resistance training and balance and motor control training affect gait performance and fatigue impact in people with multiple sclerosis? A randomized controlled multi-center study. Mult Scler. 2020;26(11):1420–32.

18. O'Malley N, Clifford AM, Conneely M, Casey B, Coote S. Effectiveness of interventions to prevent falls for people with multiple sclerosis, Parkinson's Disease and Stroke: an umbrella review. BMC Neurol. 2021;21(1):378.

19. Kalb R, Brown TR, Coote S, Costello K, Dalgas U, Garmon E, et al. Exercise and lifestyle physical activity recommendations for people with multiple sclerosis throughout the Disease course. Mult Scler. 2020;26(12):1459–69.

20. Tholen R, Dettmers C, Henze T, Höthker S, Flachenecker P, Lamprecht S, et al. Bewegungstherapie zur verbesserung der mobilität von patienten mit multipler sklerose: konsensusfassung für die s2e-leitlinie der dgnr in zusammenarbeit mit physio deutschland – Deutscher verband für physiotherapie (ZVK) e. V. Neurol Rehabil. 2019;25(1):3–40.

21. Lublin FD, Reingold SC, Cohen JA, Cutter GR, Sørensen PS, Thompson AJ, et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. Neurology. 2014;83(3):278–86.

22. Thompson AJ, Banwell BL, Barkhof F, Carroll WM, Coetzee T, Comi G, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. Lancet Neurol. 2018;17(2):162–73.

23. Andreu-Caravaca L, Ramos-Campo DJ, Chung LH, Rubio-Arias J. Dosage and effectiveness of aerobic training on Cardiorespiratory Fitness, Functional Capacity, Balance, and fatigue in people with multiple sclerosis: a systematic review and Meta-analysis. Arch Phys Med Rehabil. 2021;102(9):1826–39.

24. Penner IK, Raselli C, Stocklin M, Opwis K, Kappos L, Calabrese P. The Fatigue Scale for Motor and cognitive functions (FSMC): validation of a new instrument to assess multiple sclerosis-related fatigue. Mult Scler. 2009;15(12):1509–17.

25. Scott NW, McPherson GC, Ramsay CR, Campbell MK. The method of minimization for allocation to clinical trials. A review. Control Clin Trials. 2002;23(6):662–74.

26. Flachenecker P, König H, Meissner H, Müller G, Rieckmann P. Fatigue in multiple sclerosis: validation of the WEIMuS scale (Würzburger Erschöpfungs-Inventar Bei Multipler Sklerose). Neurol Rehabil. 2008;14:299–306.

27. Sandroff BM, Motl RW, Reed WR, Barbey AK, Benedict RHB, DeLuca J. Integrative CNS plasticity with Exercise in MS: the PRIMERS (PRocessing, integration of Multisensory Exercise-related stimuli) conceptual Framework. Neurorehabil Neural Repair. 2018;32(10):847–62.

28. Thabane L, Ma J, Chu R, Cheng J, Ismaila A, Rios LP, et al. A tutorial on pilot studies: the what, why and how. BMC Med Res Methodol. 2010;10:1.

29. Foster C, Florhaug JA, Franklin J, Gottschall L, Hrovatin LA, Parker S, et al. A new approach to monitoring exercise training. J Strength Cond Res. 2001;15(1):109–15.

30. Haddad M, Stylianides G, Djaoui L, Dellal A, Chamari K. Session-RPE Method for Training load monitoring: validity, ecological usefulness, and influencing factors. Front Neurosci. 2017;11:612.

31. Neumann M, Sterr A, Claros-Salinas D, Gütler R, Ulrich R, Dettmers C. Modulation of alertness by sustained cognitive demand in MS as surrogate measure of fatigue and fatigability. J Neurol Sci. 2014;340(1–2):178–82.

32. Claros-Salinas D, Bratzke D, Greitemann G, Nickisch N, Ochs L, Schröter H. Fatigue-related diurnal variations of cognitive performance in multiple sclerosis and stroke patients. J Neurol Sci. 2010;295(1–2):75–81.

33. Van Geel F, Veldkamp R, Severijns D, Dalgas U, Feys P. Day-to-day reliability, agreement and discriminative validity of measuring walking-related performance fatigability in persons with multiple sclerosis. Mult Scler. 2020;26(13):1785–9.

34. Strober L, DeLuca J, Benedict RH, Jacobs A, Cohen JA, Chiaravalloti N, et al. Symbol Digit modalities Test: a valid clinical trial endpoint for measuring cognition in multiple sclerosis. Mult Scler. 2019;25(13):1781–90.

35. Schlagheck ML, Bansi J, Wenzel C, Kuzdas-Sallaberger M, Kiesl D, Gonzenbach R, et al. Complexity and pitfalls in maximal exercise testing for persons with multiple sclerosis. Eur J Neurol. 2023;30(9):2726–35.

36. Goldman MD, Marrie RA, Cohen JA. Evaluation of the six-minute walk in multiple sclerosis subjects and healthy controls. Mult Scler. 2008;14(3):383–90.

37. Motl RW, Cohen JA, Benedict R, Phillips G, LaRocca N, Hudson LD, et al. Validity of the timed 25-foot walk as an ambulatory performance outcome measure for multiple sclerosis. Mult Scler. 2017;23(5):704–10.

38. Nieuwenhuis MM, Van Tongeren H, Sørensen PS, Ravnborg M. The six spot step test: a new measurement for walking ability in multiple sclerosis. Mult Scler. 2006;12(4):495–500.

39. Thieme H, Ritschel C, Zange C. Functional Gait Assessment – deutsche version. Physioscience. 2009;5:5–11.

40. Schott N. [German adaptation of the activities-specific balance confidence (ABC) scale for the assessment of falls-related self-efficacy]. Z Gerontol Geriatr. 2008;41(6):475–85.

41. D'Hooghe M, Van Gassen G, Kos D, Bouquiaux O, Cambron M, Decoo D, et al. Improving fatigue in multiple sclerosis by smartphone-supported energy management: the MS TeleCoach feasibility study. Mult Scler Relat Disord. 2018;22:90–6.

42. Sim J. Should treatment effects be estimated in pilot and feasibility studies? Pilot Feasibility Stud. 2019;5:107.

43. Leon AC, Davis LL, Kraemer HC. The role and interpretation of pilot studies in clinical research. J Psychiatr Res. 2011;45(5):626–9.

44. Kistin C, Silverstein M. Pilot studies: a critical but potentially misused component of Interventional Research. JAMA. 2015;314(15):1561–2.

45. Harrison AM, das Nair R, Moss-Morris R. Operationalising cognitive fatigability in multiple sclerosis: a Gordian knot that can be cut? Mult Scler. 2017;23(13):1682–96.

46. Langeskov-Christensen M, Langeskov-Christensen D, Overgaard K, Moller AB, Dalgas U. Validity and reliability of VO(2)-max measurements in persons with multiple sclerosis. J Neurol Sci. 2014;342(1–2):79–87.

47. Valet M, Stoquart G, de Broglie C, Francaux M, Lejeune T. Simplified indices of exercise tolerance in patients with multiple sclerosis and healthy subjects: a case-control study. Scand J Med Sci Sports. 2020;30(10):1908–17.

48. Baert I, Freeman J, Smedal T, Dalgas U, Romberg A, Kalron A, et al. Responsiveness and clinically meaningful improvement, according to disability level, of five walking measures after rehabilitation in multiple sclerosis: a European multicenter study. Neurorehabil Neural Repair. 2014;28(7):621–31.

49. Zimmer P, Bloch W, Schenk A, Oberste M, Riedel S, Kool J, et al. High-intensity interval exercise improves cognitive performance and reduces matrix metalloproteinases-2 serum levels in persons with multiple sclerosis: a randomized controlled trial. Mult Scler. 2018;24(12):1635–44.

50. Hecksteden A, Faude O, Meyer T, Donath L. How to Construct, Conduct and analyze an Exercise Training Study? Front Physiol. 2018;9:1007.

51. Flachenecker P, Meissner H, Frey R, Guldin W. Neuropsychological training of attention improves MS-Related fatigue: results of a Randomized, Placebo-Controlled, double-blind pilot study. Eur Neurol. 2017;78(5–6):312–7.

52. Broicher SD, Filli L, Geisseler O, Germann N, Zörner B, Brugger P, et al. Positive effects of fampridine on cognition, fatigue and depression in patients with multiple sclerosis over 2 years. J Neurol. 2018;265(5):1016–25.

53. Kamudoni P, Johns J, Cook KF, Salem R, Salek S, Raab J, et al. A comparison of the measurement properties of the PROMIS fatigue (MS) 8a against legacy fatigue questionnaires. Mult Scler Relat Disord. 2022;66:104048.

54. Watson C, Scippa K, Barlev A, Kresa-Reahl K, Cole JC. Results from patient interviews on fatigue in Progressive multiple sclerosis and evaluation of fatigue patient-reported outcome (PRO) Instruments. Neurol Ther. 2022;11(2):725–39.

55. Van Geel F, Moumdjian L, Lamers I, Bielen H, Feys P. Measuring walking-related performance fatigability in clinical practice: a systematic review. Eur J Phys Rehabil Med. 2020;56(1):88–103.

56. Broscheid KC, Behrens M, Bilgin-Egner P, Peters A, Dettmers C, Jöbges M et al. Instrumented Assessment of Motor Performance Fatigability during the 6-Min Walk Test in mildly affected people with multiple sclerosis. Front Neurol. 2022;13.

57. Weinrich JO, Saliger J, Eschweiler M, Karbe H, Kalbe E, Nielsen J. The relationship between diurnal measures of Tonic alertness and self-reported fatigue in persons with multiple Sclerosis-A Retrospective Data Analysis. Arch Clin Neuropsychol. 2023.

58. Brandstadter R, Katz Sand I, Sumowski JF. Beyond rehabilitation: a prevention model of reserve and brain maintenance in multiple sclerosis. Mult Scler. 2019;25(10):1372–8.

59. Arienti C, Buraschi R, Pollet J, Lazzarini SG, Cordani C, Negrini S, et al. A systematic review opens the black box of usual care in Stroke rehabilitation control groups and finds a black hole. Eur J Phys Rehabil Med. 2022;58(4):520–9.

60. Beets MW, von Klinggraeff L, Weaver RG, Armstrong B, Burkart S. Small studies, big decisions: the role of pilot/feasibility studies in incremental science and premature scale-up of behavioral interventions. Pilot Feasibility Stud. 2021;7(1):173.

Publisher's Note