

A Multicenter Randomized Comparison of High-Intensity Interval Training and Moderate-Intensity Exercise to Recover Walking Post-Stroke: Results of the HIT-Stroke Trial

Pierce Boyne PT, DPT, PhD, NCS;¹ Sandra A. Billinger PT, PhD;^{2,3,4,5} Darcy S. Reisman PT, PhD;⁶ Oluwole O. Awosika MD, MS;⁷ Sofia Buckley;¹ Jamiah Burson BS;¹ Daniel Carl PhD;¹ Matthew DeLange PT, DPT;¹ Sarah Doren PT, DPT;¹ Melinda Earnest PT, DPT, ATC;¹ Myron Gerson MD;⁸ Madison Henry PT, DPT;⁹ Alli Horning BS;¹ Jane Khoury PhD;^{10,11} Brett Kissela MD, MS;⁷ Abigail Laughlin BS;¹ Kiersten McCartney PT, DPT;⁶ Thomas McQuaid BS;¹ Allison Miller PT, DPT, PhD, NCS;⁶ Alexandra Moores MS;² Jacqueline A. Palmer PT, DPT, PhD;² Heidi Sucharew PhD;^{10,11} Elizabeth Thompson PT, PhD, NCS;⁶ Erin Wagner MS;¹ Jaimie Ward MS;² Emily Wasik PT, DPT;¹ Alicen A. Whitaker PT, DPT;⁹ Henry Wright PT, DPT;⁶ Kari Dunning PT, PhD¹

¹Department of Rehabilitation, Exercise and Nutrition Sciences, College of Allied Health Sciences, University of Cincinnati; Cincinnati, OH

²Department of Neurology, School of Medicine, University of Kansas Medical Center; Kansas City, KS

³Department of Cell Biology and Integrative Physiology, School of Medicine, University of Kansas Medical Center; Kansas City, KS

⁴University of Kansas Alzheimer's Research Disease Center, Fairway, KS

⁵Department of Physical Medicine and Rehabilitation, School of Medicine, University of Kansas Medical Center; Kansas City, KS

⁶Department of Physical Therapy, College of Health Sciences, University of Delaware; Newark, DE

⁷Department of Neurology and Rehabilitation Medicine, College of Medicine, University of Cincinnati; Cincinnati, OH

⁸Departments of Internal Medicine and Cardiology, College of Medicine, University of Cincinnati; Cincinnati, OH

⁹Department of Physical Therapy, Rehabilitation Sciences, and Athletic Training, School of Health Professions, University of Kansas Medical Center; Kansas City, KS

¹⁰Division of Biostatistics and Epidemiology, Cincinnati Children's Hospital Medical Center; Cincinnati, OH

¹¹Department of Pediatrics, College of Medicine, University of Cincinnati; Cincinnati, OH

Corresponding author:

Pierce Boyne, PT, DPT, PhD, NCS
Health Sciences Building, room 267
3225 Eden Ave.
Cincinnati, OH, 45267-0394
Pierce.Boyne@uc.edu
513-558-7499 (phone); 513-558-7474 (fax)

ABSTRACT

Introduction: For walking rehabilitation after stroke, training intensity and duration are critical dosing parameters that lack optimization. This trial aimed to determine the optimal training intensity (vigorous vs moderate) and minimum training duration (4, 8 or 12 weeks) needed to maximize immediate improvement in walking capacity in chronic stroke.

Methods: Persons with chronic post-stroke gait dysfunction at three centers were randomized to high-intensity interval training (HIT) or moderate intensity aerobic training (MAT), each involving 45 minutes of treadmill and overground walking exercise with a physical therapist, 3 times per week for 12 weeks. The HIT protocol used repeated 30 second bursts of walking at maximum safe speed, alternated with 30-60 second recovery periods, targeting an average aerobic intensity above 60% heart rate reserve (HRR). The MAT protocol used continuous walking with speed adjusted to maintain an initial target of $40 \pm 5\%$ HRR, progressing by 5% HRR every 2 weeks, up to 60% HRR as tolerated. Blinded assessment at baseline and after 4, 8 and 12 weeks of training included the 6-minute walk test (6MWT) as the primary measure of walking capacity.

Results: Randomized participants (N=55) attended 1,675 (85%) of 1,980 planned treatment sessions and 197 (90%) of 220 planned testing sessions. No serious adverse events related to study procedures occurred. Compared with MAT, HIT involved significantly higher training speeds (161% vs 96% baseline fastest 10-meter speed, $p < 0.0001$) and mean aerobic intensity (61% vs 46% HRR, $p < 0.0001$) across treatment visits. There was no significant between-group difference in 6MWT changes after 4 weeks of training (HIT +27 meters [95% CI: 6-48], MAT +12 meters [-9-33], $p = 0.28$), but randomization to HIT resulted in significantly greater gains than MAT after 8 weeks (+58 [39-76] vs +29 [9-48] meters, $p = 0.02$) and 12 weeks (+71 [49-94] vs +27 [3-50] meters, $p = 0.005$) of training. HIT also showed significantly greater improvements than MAT on some measures of gait speed, fatigue and exercise capacity.

Discussion: These findings show proof of concept that vigorous training intensity is a critical dosing parameter for walking rehabilitation. In chronic stroke, vigorous walking exercise can produce significant and meaningful gains in walking capacity with only 4 weeks of training, but at least 12 weeks are needed to maximize immediate gains.

Keywords: locomotion, gait, aerobic exercise, HIIT

Trial Registration: ClinicalTrials.gov Identifier: NCT03760016. First posted: November 30, 2018.
<https://clinicaltrials.gov/ct2/show/NCT03760016>

Competing Interests: The authors declare no competing interests.

Funding: Research reported in this publication was supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development (NICHD) of the National Institutes of Health under award number R01HD093694. This work was also supported by Clinical and Translational Science Award (CTSA) grants from the National Center for Advancing Translational Sciences (NCATS) awarded to the University of Cincinnati for the Cincinnati Center for Clinical and Translational Science and Training (UL1TR001425) and to the University of Kansas for Frontiers: University of Kansas Clinical and Translational Science Institute (TL1TR002368). SAB was also supported by P30 AG072973. AM was also supported by a Foundation for Physical Therapy Research Florence P. Kendall Doctoral Scholarship and Promotion of Doctoral Studies I and II Scholarships. AW was also supported by NICHD on T32HD057850. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding organizations.

Acknowledgements: The authors thank Linda D'Silva, Rachel Gleason, Bill Hendry, Chun Kai-Huang, Katie Libby, Babette Northrop, Jenna Octavio, Jaclyn Reddington, Caio Sarmento, Eric Vidoni, Kaylah Williamson and Tamara Wright for assistance with data collection and Kelly Laipply for cardiology consultation. We also thank T George Hornby, Michael W O'Dell and Bin Huang for serving on the data safety monitoring board.

INTRODUCTION

While the majority of stroke survivors eventually regain the ability to walk without physical assistance from another person, most in the chronic phase of recovery still lack sufficient speed and endurance to resume normal daily activities.¹⁻⁵ The neuromotor impairment and aerobic deconditioning underlying these limitations in walking capacity can both be targeted simultaneously with locomotor exercise (i.e. task-specific walking practice at sufficient intensity to challenge the neuromotor and cardiopulmonary systems).⁶⁻⁸ With this intervention, training intensity appears to be a critical dosing parameter that strongly influences outcomes.⁹⁻²¹ However, the optimal intensity for improving walking recovery remains unknown.

Most randomized controlled studies of post-stroke locomotor exercise have tested moderate-intensity aerobic training (MAT), typically involving treadmill walking at average training heart rates (HR) between 40-60% heart rate reserve (HRR).²²⁻³³ These studies have found that MAT improves walking capacity and other outcomes significantly more than lower intensity walking practice or non-walking exercise. Consequently, this approach is currently recommended in stroke rehabilitation guidelines.⁶⁻⁸

Studies also suggest that a more vigorous training intensity (>60% HRR vs 40-60% HRR) could augment outcomes,⁹ but that a vigorous intensity can be difficult to achieve and sustain for many persons with stroke.³⁴⁻³⁶ Thus, high-intensity interval training (HIT) has emerged as a promising strategy for post-stroke locomotor exercise.^{37,38} This method involves bursts of fast walking alternated with recovery periods, and is designed to enable higher sustained intensities³⁹ at lower perceived exertion than high-intensity continuous exercise.⁴⁰ Initial stroke studies indicate that locomotor HIT is feasible, sufficiently safe for further study, can improve walking capacity and other outcomes,^{15,16,41-47} and has some early signs of potential for eliciting greater improvements than MAT.^{43,44}

However, no prior studies have been designed to compare outcomes of HIT and MAT post-stroke. In addition, most prior HIT studies post-stroke used short 4-week training durations. Therefore, it was unknown whether a longer 12-week training duration typical of many stroke and cardiac rehabilitation trials^{9,48} could yield better outcomes.

The purpose of this multi-center trial was to determine the optimal training intensity (vigorous vs moderate) and the minimum training duration (4, 8 or 12 weeks) needed to maximize immediate improvement in walking capacity in chronic stroke. We hypothesized that: 1) 4 weeks of HIT would elicit significantly greater improvement in walking capacity compared to 4 weeks of MAT; and 2) Compared with 4 and 8 weeks of HIT, 12 weeks of HIT would elicit significantly greater improvements in walking capacity and increased benefit over MAT.

METHODS

Study overview

This study was approved by the University of Cincinnati IRB and performed in rehabilitation and exercise laboratories at the University of Cincinnati, University of Delaware and University of Kansas Medical Center with participant enrollment from January 2019 to April 2022. It was prospectively registered at ClinicalTrials.gov (NCT03760016) and a detailed protocol was published.⁴⁹

Multisite standardization

Procedures were standardized across sites using a detailed manual of procedures, video-based online personnel training, 2-day in-person training at each site and monthly web-meetings. This study also used direct-electronic data entry into REDCap forms that were programmed to provide real-time protocol reminders, automated calculations, prompts and feedback during each study visit. In addition, a software application used during each treatment visit was programmed to prompt training start/stop times and provide real-time intensity feedback.

Participants

Participants were recruited from the community via outreach to clinicians and support groups, advertisement, existing databases and health record systems. Written informed consent was obtained prior to participation. Screening included a medical history and record review, physical examination, Fugl-Meyer lower limb motor function assessment,⁵⁰ depression questionnaire,⁵¹ 10-meter walk test,^{52,53} treadmill acclimation and screening test,⁵² resting electrocardiogram (ECG) and a treadmill graded exercise test with ECG monitoring.⁵⁴

Inclusion criteria were: 1) Age 40-80 years at the time of consent; 2) Single stroke for which the participant sought treatment, 6 months to 5 years prior to consent; 3) Walking speed ≤ 1.0 m/s on the 10-meter walk test;^{53,55} 4) Able to walk 10m over ground with assistive devices as needed and no continuous physical assistance from another person;⁵⁶ 5) Able to walk at least 3 minutes on the treadmill at ≥ 0.13 m/s (0.3 mph);⁵² 6) Stable cardiovascular condition (American Heart Association class B,⁵⁷ allowing for exercise capacity <6 METs); and 7) Able to communicate with investigators, follow a 2-step command and correctly answer consent comprehension questions.

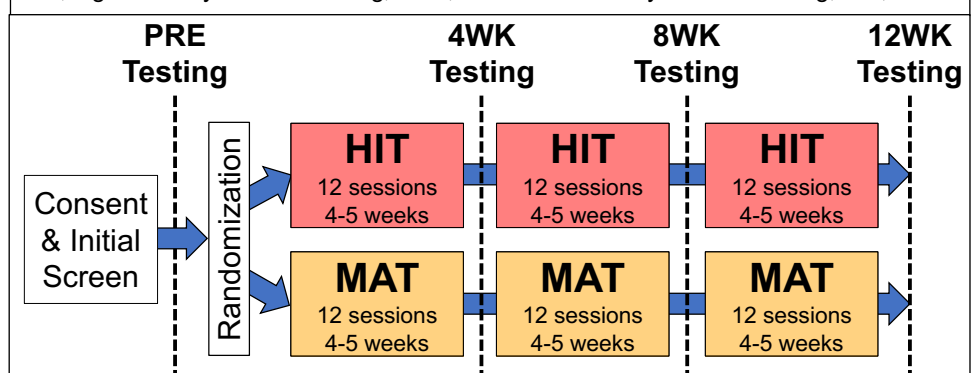
Exclusion criteria were: 1) Exercise testing uninterpretable for ischemia or arrhythmia;⁵⁷ 2) Evidence of significant arrhythmia or myocardial ischemia on treadmill ECG graded exercise test in the absence of recent (past year) more definitive clinical testing with negative result;⁵⁷ 3) Hospitalization for cardiac or pulmonary disease within the past 3 months; 4) Implanted pacemaker or defibrillator; 5) Significant ataxia or neglect (score of 2 on NIH stroke scale item 7 or 11);⁵⁸ 6) Severe lower limb spasticity (Ashworth >2);⁵⁹ 7) Recent history (<3 months) of illicit drug or alcohol abuse or significant mental illness; 8) Major post-stroke depression (Patient Health Questionnaire [PHQ-9] ≥ 10)⁵¹ in the absence of depression management by a health care provider;⁶⁰ 9) Currently participating in physical therapy or another interventional study; 10) Recent botulinum toxin injection to the paretic lower limb (<3 months) or planning to have lower limb botulinum toxin injection in the next 4 months; 11) Foot drop or lower limb joint instability without adequate stabilizing device, as assessed by a physical therapist; 12) Clinically significant neurologic disorder other than stroke; 13) Unable to walk outside the home prior to stroke; 14) Other significant medical condition likely to limit improvement or jeopardize safety as assessed by a physical therapist; 15) Pregnancy; and 16) Previous exposure to fast treadmill walking (>3 cumulative hours) during clinical or research therapy in the past year.

Study design

Eligible participants were randomized to either HIT or MAT, with a target training volume of 45 minutes, 3 times per week for 12 weeks (Figure 1). This was divided into three 4-week (12-session) blocks separated by outcome testing. In each training block, up to one additional week was allowed to make up for missed sessions if needed. Outcomes were assessed by blinded raters

Figure 1. Study design

HIT, high-intensity interval training; MAT, moderate-intensity aerobic training; WK, week



before randomization (PRE) and after 4, 8 and 12 weeks of training (4WK, 8WK, 12WK), including measures of walking function, fatigue and aerobic capacity.

Randomization

After screening and PRE testing, eligible participants were randomized to HIT or MAT in a 1:1 ratio using the web-based REDCap randomization module.⁶¹ Randomization was stratified by site and baseline gait dysfunction (severe, <0.4 m/s; mild/moderate, 0.4-1.0 m/s),^{62,63} and used randomly-permuted block sizes of 2 or 4 to help ensure balanced groups while preventing personnel from being able to predict the last randomization within a block. The study statistician generated the random allocation sequence and kept it confidential from all other study personnel to ensure concealed allocation. When eligible participants arrived for the first training session, the treating therapist or study coordinator initiated the randomization within REDCap to irreversibly assign the participant to the next allocation in the randomized sequence before revealing the treatment assignment.

Interventions

Training sessions were directed by a physical therapist with support from a research assistant as needed for data collection. Habitual orthotic devices were used and occasionally supplemented by additional orthotics to protect against ankle inversion sprain or severe knee hyperextension during training, based on the judgment of the treating therapist. Participants were not physically assisted with stepping but were guarded as needed and assisted for injury prevention during instances of severe gait instability or loss of balance. Participants who enrolled after the beginning of the COVID-19 pandemic wore personal protective equipment during testing and treatment sessions according to the infection control protocols at each site.

During overground training bouts, participants walked back and forth in a corridor. Habitual assistive devices were used during initial bouts, then therapists attempted to progress the participant to less restrictive device(s) if it better enabled achievement of the target intensity. When training on a treadmill (LiteGait Gaitkeeper 2200T, Mobility Research, Tempe, AZ, USA; or L8 Rehab Treadmill, Landice, Randolph, NJ, USA), participants wore a fall protection harness without any weight support (Balance Harness, MASS Rehab, Dayton, OH) connected to an overhead system (RehabMODULE, MASS Rehab, Dayton, OH, USA) and used a handrail connected to the fall protection system for balance support.⁵² This handrail was individually height-adjusted during treadmill training to allow upright walking posture.

Training intensity was monitored and recorded using Bluetooth Smart heart rate (HR) transmitters (H7 and H10, Polar Electro Oy, Kempele, Finland; or Rhythm24, Scosche Industries, Inc., Oxnard, CA, USA) wirelessly connected to an iPod touch (Apple, Cupertino, CA, USA) using the iCardio application (FitDigits, Inc., Ventura, CA, USA).^{63,64} The training protocols and individual prescribed target HR zones were programmed into the application to time each training bout, signal burst/recovery transitions (for HIT) and provide real-time intensity guidance. A forearm-worn iPod mount enabled hands-free monitoring for the treating therapist while guarding the participant.

Training HR zones were calculated using the HR reserve (HRR) method: $(HR_{\text{peak}} - HR_{\text{rest}}) \times \% \text{HRR target} + HR_{\text{rest}}$.⁵⁷ During training, HR_{peak} was taken as the highest instantaneous HR obtained in any prior exercise testing session for that participant. HR_{rest} was measured at the beginning of each session in both sitting and standing. For this study, the HR_{rest} value in standing was used for target HR calculations. This was based on our pilot testing in which some participants with more severe deconditioning were found to have large HR increases just from standing that would have otherwise exceeded some of the target HR ranges and precluded walking practice if we had used the true HR_{rest} value for target HR calculation.

Each session, the training protocol for both groups included: 1) a 3-minute warm-up of overground walking at a speed that increased HR to 30-40% HR reserve (HRR); 2) a 10-minute bout of overground HIT or MAT; 3) a 20-minute bout of treadmill HIT or MAT; 4) another 10-minute bout of overground HIT or MAT; and 5) a 2-minute cool down at 30-40% HRR (Figure 2).

High-intensity interval training (HIT) protocol

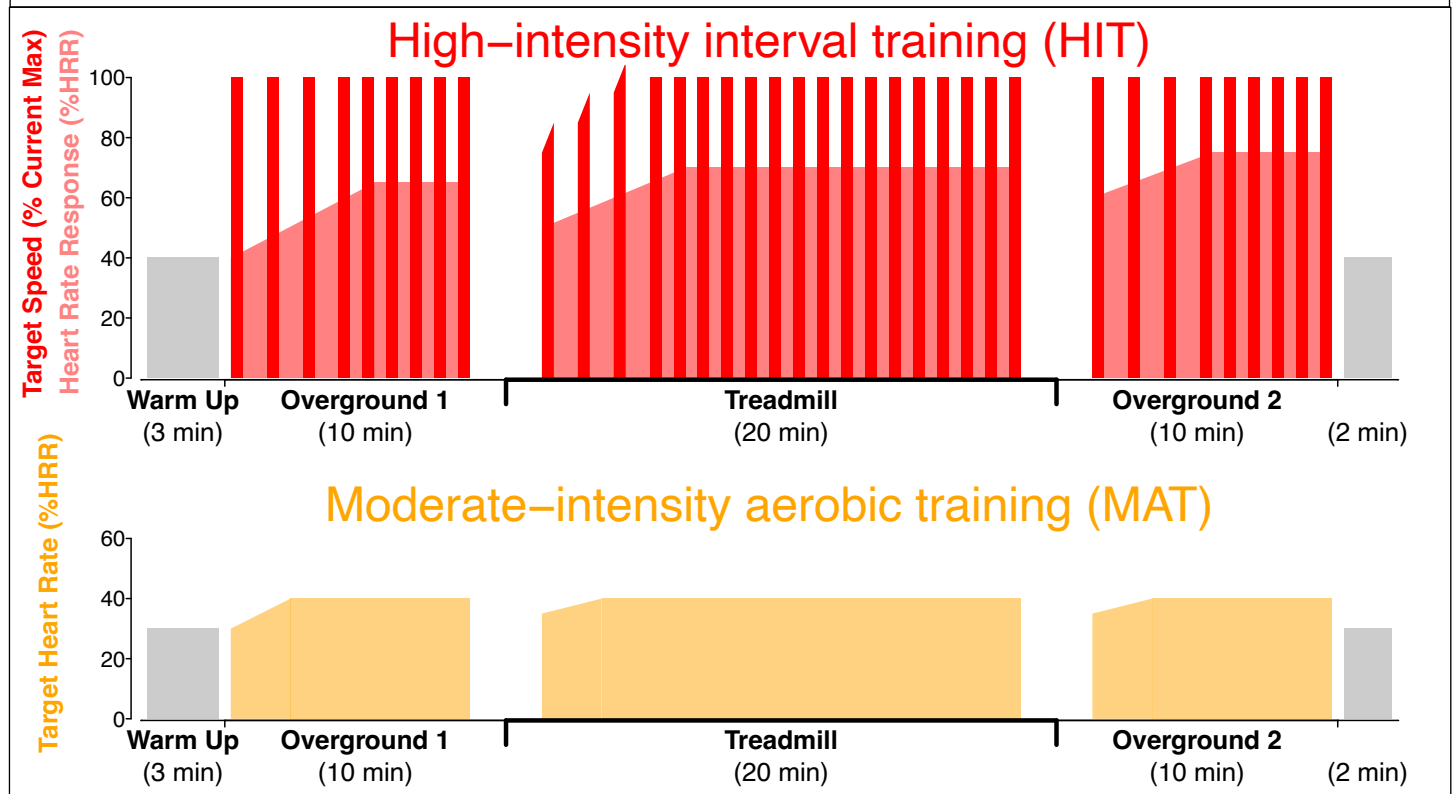
The HIT group used a 'short-interval' HIT protocol that was specifically developed for locomotor exercise post-stroke.^{43,63-65} It involved repeated 30 second bursts of walking at maximum safe speed, alternated with 30-60 second passive recovery periods, targeting an average aerobic intensity above 60% HRR. During overground HIT, burst speed was maximized by placing cones or beanbags on the floor to mark the distance covered during each burst and encouraging participants to continually work towards making it past the marker before time ran out. During treadmill HIT, burst speed was selected to provide the maximum safe

challenge for that participant and was progressed as able (or regressed as needed) throughout each session based on performance criteria to maintain constant challenge.⁶⁴

Moderate-intensity aerobic training (MAT) protocol

Following current best-practice guidelines for stroke rehabilitation,^{6,7} the MAT group performed continuous walking practice with speed adjusted to maintain an initial target HR of $40 \pm 5\%$ HRR, progressing by 5% HRR every 2 weeks, up to 60% HRR as tolerated.⁵²

Figure 2. Training Protocol Schematics. HIT (upper panel) uses 30 second bursts at 100% maximum safe speed (dark red bars) alternated with 30-60 second passive recovery periods and within-session speed progression as able (not shown). This typically elicits a mean aerobic intensity $>60\%$ heart rate reserve (HRR; light red shading). The first three treadmill bursts depict example within-burst speed increases to find the current maximum safe treadmill speed. MAT (lower panel) uses continuous walking with speed adjusted as needed to maintain the moderate aerobic intensity target, which starts at 40% HRR (shown) and progresses by 5% HRR every 2 weeks, up to 60% HRR (not shown).



Treatment fidelity data collection

- *Training volume* was measured by the number of sessions attended, the number of treadmill and overground training bouts started and the total number of training minutes performed.
- *Neuromotor intensity* was measured by peak training speed in each bout. Overground speeds were captured using a stopwatch at the beginning and end of each overground training bout as the participants crossed over markings for the 10-meter walk test in the overground training corridor. Treadmill speeds were recorded from the treadmill display.
- *Aerobic intensity* was measured by HR data collected continuously by the Bluetooth HR monitors and iPod touch iCardio application during each training bout. These data were exported at 1/3 Hz and processed to calculate mean steady-state HR (excluding the first 3 minutes) and max HR for each bout as %HRR.
 - For real-time treatment monitoring, %HRR values were calculated using standing HR_{rest} and the most current HR_{peak} , as described above. However, the use of standing HR_{rest} makes such %HRR values incomparable to previous studies and allowing HR_{peak} updates at 4WK or 8WK could distort %HRR comparisons across sessions and/or treatment groups. Therefore, to facilitate these comparisons, %HRR values were also calculated using true HR_{rest} (the smaller of the two HR_{rest} values obtained from each session, which was typically the sitting HR_{rest}) and HR_{peak} from PRE testing only.
 - HR data preprocessing included spike-filtering and plotting to visually check for invalid data (signal dropout, residual non-physiologic spikes, episodes of supraventricular tachycardia and large elevations clearly attributed to coughing or laughing), while referencing notes and other data from

each bout. Where possible, invalid data segments were filled using linear interpolation. When a bout began with invalid data, the first value at the beginning of the adjacent bout(s) in the same session was used to initialize the interpolation, to approximate the typical HR increase from a low starting value at the beginning of the bout. When a bout ended with invalid data, the final valid data point was carried forward, to approximate the typical steady mean trajectory at the end of the bout. However, when this did not reasonably approximate the likely mean HR trajectory, the HR data from that bout were discarded as invalid (e.g. when >50% of the HR data from a bout were invalid/missing or the mean trajectory was steeply sloped). If resting HR was missing or invalid for a session (e.g. due to an episode of supraventricular tachycardia that subsided during the session), it was imputed for HRR calculation by taking the mean resting HR across sessions for that participant.

- *Anaerobic intensity* was measured by blood lactate concentration after the treadmill training portion of the middle training session in each training week, using a finger stick and a point-of-care blood lactate analyzer (Lactate Plus, Nova Biomedical, Waltham, MA, USA).
- *Practice repetition* was measured by step count, recorded each session using a Stepwatch Activity Monitor (Modus Health, LLC, Edmonds, WA, USA) on the non-paretic ankle.⁶⁶
- *Rating of perceived exertion* (RPE) for the session was measured by the Borg 6-20 scale⁵⁷ at the end of each session.

Adverse event (AE) monitoring

Each study visit, participants were asked about AEs in general then specifically queried about falls, injuries, pain and fatigue. During exercise testing and training, safety monitoring included HR, blood pressure, and continuous observation for other signs or symptoms of cardiorespiratory insufficiency, worsening neurologic impairments or orthopedic injury, using accepted stopping criteria.^{57,67} AEs were systematically categorized and graded using Common Terminology Criteria for Adverse Events,^{68,69} by a centralized physician blinded to group randomization, based on standardized information provided by the research team member who discovered the AE.

Outcome assessment

Outcomes were assessed by blinded raters at PRE, 4WK, 8WK and 12WK. All measures were assessed in a single visit and were arranged in standardized order to mitigate fatigue effects whenever possible, with overground walking tests followed by seated questionnaires then treadmill exercise testing.⁴⁹ Testing visits were scheduled 2-7 days after the last treatment visit in the preceding block and subsequent treatment blocks were scheduled to begin within 10 days of the last treatment visit. An individual participant was tested by the same rater at all time points whenever possible. Similar to training, participants used habitual assistive and orthotic devices during testing, wore a fall protection harness during treadmill walking and were not physically assisted but were guarded by a physical therapist and provided injury prevention assistance if needed.

Primary outcome measure

The primary outcome measure was walking capacity, measured by distance walked during the 6-minute walk test (6MWT).^{33,70} The 6MWT was the primary outcome measure because it can be influenced by both gait speed and aerobic fitness,^{6,7} the two primary factors contributing to limited walking capacity post-stroke and the two primary targets of HIT and MAT.⁹ Further, this test explains more variance in home and community ambulation than any other laboratory measure for persons with stroke,⁷¹⁻⁷³ is a primary characteristic distinguishing between home and community ambulators,⁷⁴ and is significantly associated with post-stroke quality of life.⁷⁵ The 6MWT has also been accepted by the FDA as a primary outcome for registered trials across various neurologic and cardiopulmonary diagnoses.⁷⁶⁻⁸⁰ After stroke, it has excellent test-retest reliability (ICC: 0.97-0.99)⁸¹⁻⁸⁴ and adequate inter-rater reliability (ICC: 0.78).⁸⁵ The minimal clinically important difference is 20 meters and 50 meters is a large improvement.⁸⁶ It was standardized according to guidelines,⁷⁰ allowing minor common modifications to the testing course at each site.⁴⁹

Secondary outcome measures

- Gait speed at self-selected and fastest speeds, measured by the 10-meter walk test.^{52,53}
- Self-reported fatigue over the past 7 days, measured by the Patient Reported Outcomes Measurement Information System (PROMIS)-Fatigue Scale version 8a.^{87,88}
- Aerobic capacity, measured by oxygen consumption rate (VO₂) at the ventilatory threshold during a treadmill graded exercise test (GXT).⁸⁹ The ventilatory threshold is a submaximal transition point that

represents the upper intensity limit of prolonged aerobic activity,⁹⁰⁻⁹² beyond which anaerobic metabolism becomes more dominant, limiting sustainability.^{90,91,93} Compared with peak VO_2 during an exercise test, which is often confounded by motor impairment after stroke,^{89,90,93,94} the ventilatory threshold appears to provide a more specific measure of aerobic capacity in this population.⁸⁹

Treadmill exercise testing

The treadmill GXT was done with ECG monitoring, blood pressure testing and gas exchange analysis, using a metabolic cart (TrueOne 2400, ParvoMedics, Salt Lake City, UT, USA) with a facemask interface. A blinded physical therapist monitored gait safety while additional study team member(s) monitored ECG, blood pressure and metabolic data. The symptom-limited GXT protocol was standardized to be 0.3 mph at 0% grade for the first 3 minutes, then to increase by 0.1 mph every 30 seconds. If a participant reached 3.5 mph, that speed was maintained, and the incline started increasing by 0.5% every 30 seconds. The test continued until the participant requested to stop, drifted backward on the treadmill and was unable to recover, had severe gait instability judged to pose an imminent safety risk, or reached a cardiovascular safety limit.⁵⁷ At least 10 minutes after the GXT, participants also attempted a 3-minute verification test without respiratory gas collection to help ensure that the highest possible HR was reached to guide individualized training intensity prescription. This test was done at the peak successful speed and grade from the GXT or at a higher speed/grade if the blinded testing therapist judged that it would be feasible. It was stopped before 3 minutes if one of the GXT stop criteria occurred sooner.

Ventilatory threshold determination

Ventilatory threshold identification from the GXT metabolic data was done using a semi-automated method. First, breath-by-breath data from each test were resampled to 5 second averaging and truncated to the GXT duration, including VO_2 , the rate of carbon dioxide production (VCO_2) and the expiratory air flow volume (VE). Excess CO_2 was calculated as $(\text{VCO}_2^2 / \text{VO}_2^2) - \text{VCO}_2^2$, with all values in L/min, and the ventilatory equivalents of O_2 and CO_2 were calculated as VE/VO_2 and VE/VCO_2 .⁹⁵⁻⁹⁷ Automated threshold detection was then applied separately to the excess CO_2 and VE/VO_2 time series⁹⁵⁻⁹⁷ using the R⁹⁸ package 'mcp'⁹⁹ to identify the posterior distribution likelihood of the ventilatory threshold at each 5 second time point. This was performed separately after smoothing the data using 0, 10 and 20 second rolling averages. The median of the posterior distributions across all methods was taken as an automated point estimate for time to ventilatory threshold.

Two raters then independently verified or adjusted this automated estimate while blinded to group assignment. Raters viewed graphs of VO_2 , VCO_2 and VE by time, excess CO_2 by time, VE/VO_2 and VE/VCO_2 by time and VCO_2 by VO_2 (V slope), across all testing timepoints for each participant. Automated posterior distribution likelihoods and point estimates were also displayed on the graphs. The ventilatory threshold was identified by an upward break point in the excess CO_2 and VE/VO_2 data, disproportionate to any VE/VCO_2 increase, where the slope of the VCO_2 by VO_2 relationship increased above 1.⁹⁵⁻⁹⁷ If no threshold occurred during the test, time to threshold was set at the end of the GXT. Inter-rater reliability was excellent ($\text{ICC}_{2,1}$: 0.98 [95%CI: 0.97-0.99], mean absolute error: 0.47 ± 0.87 minutes) despite a small systematic difference in mean time-to-threshold between raters (10.2 ± 4.9 versus 10.5 ± 4.7 minutes, $p < 0.0001$). Final time-to-threshold values were determined by agreement (10.4 ± 4.7 minutes) and VO_2 at the ventilatory threshold was calculated by averaging the VO_2 data in a 20 second window around this time point.

Other exercise test measures to facilitate interpretation and exploratory outcomes

- Exercise capacity (not necessarily specific to aerobic fitness) was measured by GXT duration, peak GXT speed, peak GXT VO_2 , peak GXT HR and HR_{peak} (highest HR between GXT and verification test). Peak VO_2 was calculated from 5 second averaged data after smoothing with a 20 second rolling average. Peak HR was the instantaneous peak from the ECG data.
- To facilitate interpretation of whether maximal *aerobic* capacity was reached during exercise testing:
 - HR_{peak} was also expressed as % age-predicted maximal HR, calculated using $164 - (0.7 \times \text{age} [\text{years}])$ ¹⁰⁰ for participants taking a β -blocker medication and $206.9 - (0.67 \times \text{age} [\text{years}])$ ¹⁰¹ otherwise. Values are typically $\geq 90\%$ at maximal aerobic capacity.⁵⁷
 - Peak respiratory exchange ratio (VCO_2/VO_2) was calculated, using the same preprocessing as peak VO_2 . Values are typically at least ≥ 1.05 at maximal aerobic capacity.⁵⁷
 - Rating of perceived exertion (Borg 6-20) was obtained at the end of the GXT. Values are typically ≥ 17 at maximal aerobic capacity.⁵⁷
- Treadmill speed at the ventilatory threshold was measured to facilitate interpretation of any changes in VO_2 at the ventilatory threshold. This represents the fastest speed that should have prolonged sustainability

without accumulating anaerobic metabolites. It was calculated by averaging the protocol speed values in a 20 second window around the ventilatory threshold time.

- Metabolic cost of treadmill gait [$\text{mL O}_2/\text{kg/m}$] during the GXT was calculated as $\text{VO}_2 [\text{mL/kg/min}] / \text{speed} [\text{m/min}]$, where lower values represent more efficient gait.¹⁰² Values were averaged in the last 3 minutes of each GXT. However, since faster speeds are typically more efficient after stroke,¹⁰² we also averaged the metabolic cost data in the last 3 minutes of the *shortest* test for each participant. This matched the speeds across time points to assess how much any changes in efficiency were related to (or independent from) any changes in speed.¹⁰³
- Gait stability was measured with the functional ambulation category,¹⁰⁴ based on participant performance during the 10-meter and 6-minute walk tests. Since we did not observe participants walking on nonlevel surfaces, the highest category (independent on nonlevel surfaces) was collapsed into the second highest category (independent on level surfaces), yielding a score range from 0-4+.
- Perceived balance confidence was measured with the activities-specific balance confidence (ABC) scale, a questionnaire with 16 items averaged to yield a total confidence score from 0-100%.^{105,106}
- Quality of life related to mobility, self-care, usual activities, pain/discomfort and anxiety/depression were measured with the EuroQOL-5D-5L questionnaire.^{107,108} A total 'misery score' was calculated by summing the 5 item scores.¹⁰⁸ Individual item scores were also analyzed for mobility and usual activities, given that these constructs were targeted by the interventions. Higher values represent worse problems for each of these scores.

Statistical analysis

Analyses followed a prespecified plan,⁴⁹ used an *intent-to-treat* approach and were done with SAS[®], version 9.4 (SAS Institute, Cary, NC, USA). The study statistician remained blinded to treatment groups until after the primary analysis. To assess randomization performance, baseline participant characteristics were compared between groups with independent t-tests and Fisher exact tests. Baseline data were also expressed as a percentage of normative predicted values for self-selected gait speed,⁵⁵ 6MWT distance,¹⁰⁹ VO_2 at the ventilatory threshold^{92,110} and $\text{VO}_{2\text{-peak}}$.¹¹⁰ The HR_{peak} increase attributable to the verification test was quantified by comparing the GXT peak HR to HR_{peak} with a paired t-test.

Treatment fidelity analysis

Training intensity, repetition and perceived exertion data were compared between groups to evaluate treatment fidelity. Each of these dependent variables was tested in a separate statistical model, which included fixed effects for treatment group, session number (modeled as a categorical effect), training bout (for variables collected at overground 1, treadmill and overground 2 each session) and all possible interactions. Repeated sessions within the same participant were modeled with compound symmetry covariance and repeated bouts within the same session were modeled with unconstrained covariance. Variances were not constrained to be the same between sessions or between treatment groups. Missing data were handled with the method of maximum likelihood.

Adverse event (AE) analysis

To assess the relative odds of harms, post-randomization AEs were compared between treatment groups with logistic regression. Separate models were tested for overall AEs and for each AE categorization, using the number of participants with an AE in that category as the dependent variable and fixed effects for [treatment group], [study site] and [baseline gait dysfunction (<0.4, 0.4-1.0 m/s)]. If only one group had AE(s), a continuity correction was added of 0.5 AEs to each group to permit calculation of the odds ratio. In this case, it was not possible to adjust for study site or baseline dysfunction.

Hypothesis testing

Primary hypothesis testing used a linear model with the 6MWT as the dependent variable and fixed effects for treatment group, testing time point (PRE, 4WK, 8WK, POST), group by time interaction, study site, study site by time interaction, baseline gait dysfunction and baseline gait dysfunction by time interaction, with unconstrained covariance between repeated testing time points within the same participant and a significance threshold of $p < 0.05$. Secondary outcomes were tested using the same model, with false discovery rate (FDR) correction¹¹¹ to the significance threshold to control for multiple testing time points (4WK Δ , 8WK Δ , 12WK Δ) and the 5 primary or secondary outcome measures. Other exercise test measures and exploratory outcomes were tested with FDR correction across testing time points and all tested measures. These analyses handled any missing data with the method of maximum likelihood, which assumes that data were missing at random.¹¹²

Analysis stratified by baseline gait dysfunction subgroup

To preliminarily examine whether the primary results differed for participants with severe versus mild/moderate baseline gait dysfunction, we added a baseline gait dysfunction by group by time interaction effect to the primary model. In chronic stroke, persons with mild/moderate gait dysfunction have been found to have better responsiveness to locomotor exercise.^{62,63}

Sensitivity analysis for missing data assumptions

While the 'missing at random' assumption is common in the analysis of clinical trials, it is not testable and could be violated in many typical circumstances, like when outcome data are missing because of AE-related participant withdrawal.^{112,113} To assess how much the results depended on the 'missing at random' assumption, we repeated the primary analysis assuming that participants with missing outcomes due to AE-related withdrawal had poor outcomes, similar to Duncan and colleagues.⁶⁰ For any outcome data point missing because of AE-related withdrawal, the true value was assumed to be distributed around either the baseline or the last observation for that participant, whichever was smaller. If the data point missingness was not AE-related, the true value was assumed to be distributed around the last observation for that participant. These distributions for the true values were assumed to be normal with a standard deviation of 15 m, matching the observed standard deviation of 6MWT changes after 4 weeks of no intervention in a similar population.⁶³ Using the general multiple imputation framework,^{112,113} 50 datasets were generated by random sampling to impute the missing values, each dataset was analyzed, and the model estimates were pooled across datasets.

Sample size

This study was powered to detect a between-group difference of 20 m in 6MWT change⁸⁶, using the software 'GLIMMPSE'.¹¹⁴ We estimated that the MAT group would improve by 15 m every 4 weeks⁴³ and set the HIT estimate 20 m larger at each time point. Variance & covariance parameters were estimated by pooling data across our prior 4-week pilot studies and extrapolating the repeated measures correlations involving the later 8WK and 12WK time points using the highest suggested exponential decay rate (0.5).¹¹⁴ These calculations indicated a total target sample size of 40 for $\geq 80\%$ power at a two-sided significance threshold of 0.05. To account for up to 20% attrition, we initially planned to enroll and randomize 50 participants. However, after having to withdraw four participants directly due to COVID-19 shutdown, we opted to increase the enrollment target to 55. This decision was made before any analysis of outcome data.

RESULTS

Recruitment and retention

We consented and screened 74 participants to find 55 who were eligible for enrollment and randomized (Figure 3). No baseline characteristic significantly differed between groups (Table 1). In total, participants attended 1,675 (85%) of 1,980 planned treatment sessions and 197 (90%) of 220 planned testing sessions.

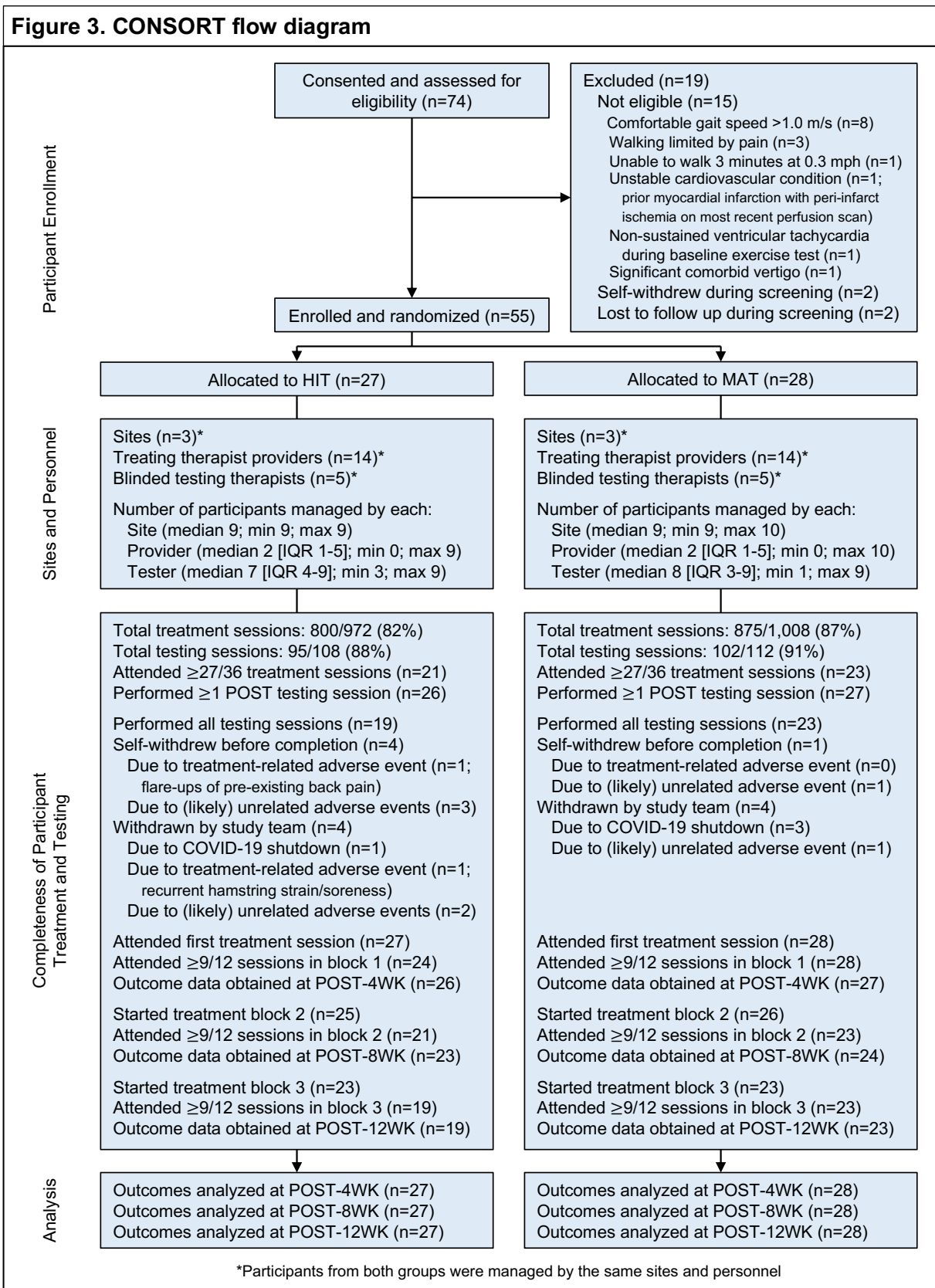


Table 1. Baseline Participant Characteristics. <i>Values are mean (SD) or N (%).</i>			
	HIT (N=27)	MAT (N=28)	p-value
Age, years	63.8 (9.9)	61.5 (9.9)	0.38
Females, N (%)	11 (40.7%)	8 (28.6%)	0.40
Race, N (%)			0.09
American Indian	0 (0.0%)	1 (3.6%)	
Asian	0 (0.0%)	3 (10.7%)	
Black / African American	4 (14.8%)	7 (25.0%)	
White	23 (85.2%)	17 (60.7%)	
Hispanic/Latinx ethnicity, N (%)	1 (3.7%)	0 (0.0%)	0.49
Side of paresis, N (%)			1.00
Left	14 (51.9%)	14 (50.0%)	
Right	13 (48.1%)	14 (50.0%)	
Stroke type, N (%)			0.41
Ischemic	15 (55.6%)	19 (67.9%)	
Hemorrhagic	12 (44.4%)	9 (32.1%)	
Stroke chronicity, years	2.7 (1.4)	2.2 (1.2)	0.13
Aphasia present, N (%)	7 (25.9%)	7 (25.0%)	1.00
Comorbid medical conditions, N (%)	26 (96.3%)	27 (96.4%)	1.00
History of myocardial infarction	1 (3.7%)	2 (7.1%)	1.00
Atrial fibrillation	2 (7.4%)	5 (17.9%)	0.42
History of lightheadedness/syncope	4 (14.8%)	5 (17.9%)	1.00
Chronic obstructive pulmonary disease	3 (11.1%)	1 (3.6%)	0.35
Diabetes mellitus	8 (29.6%)	9 (32.1%)	1.00
Insulin dependent	2 (25.0%)	2 (22.2%)	1.00
Arthritis	8 (29.6%)	10 (35.7%)	0.78
History of cancer	4 (14.8%)	3 (10.7%)	0.70
Prescribed medication, N (%)			
Anti-hypertensive	23 (85.2%)	20 (71.4%)	0.33
β-blocker	7 (25.9%)	11 (39.3%)	0.39
Statin	22 (81.5%)	23 (82.1%)	1.00
Anti-spasmodic	11 (40.7%)	12 (42.9%)	1.00
Anti-depressant	20 (74.1%)	13 (46.4%)	0.054
Pain relief	11 (40.7%)	11 (39.3%)	1.00
Reported falling in previous 3 months, N (%)	8 (29.6%)	7 (25.0%)	0.77
Body mass index, kg/m ²	28.9 (5.3)	28.7 (4.6)	0.87
Functional ambulation category, 2-4	3.3 (0.6)	3.4 (0.8)	0.90
2: Dependent, Level 1, N (%)	2 (7.4%)	5 (17.9%)	0.19
3: Dependent, Supervision, N (%)	14 (51.9%)	8 (28.6%)	
≥4: Independent, Level Surfaces, N (%)	11 (40.7%)	15 (53.6%)	
Orthotic and/or assistive device use, N (%)	18 (66.7%)	19 (67.9%)	1.00
Orthotic device use, N (%)	12 (44.4%)	12 (42.9%)	1.00
Solid ankle foot orthosis	4 (14.8%)	2 (7.1%)	0.42
Articulated/flexible ankle foot orthosis	7 (25.9%)	7 (25.0%)	1.00
Assistive device use, N (%)	16 (59.3%)	18 (64.3%)	0.78
Single point cane	8 (29.6%)	11 (39.3%)	0.33
Narrow-based quad cane	5 (18.5%)	1 (3.6%)	
Wide-based quad cane	3 (11.1%)	5 (17.9%)	
Front-wheeled walker	0 (0.0%)	1 (3.6%)	
Fugl-Meyer lower limb motor score, 0-34	23.8 (5.1)	22.8 (5.1)	0.44
Self-selected gait speed, m/s	0.65 (0.29)	0.62 (0.33)	0.75
Self-selected gait speed, % predicted	50.5 (23.3)	47.3 (25.1)	0.63
Self-selected gait speed <0.4 m/s, N (%)	7 (25.9%)	7 (25.0%)	1.00
6-minute walk test, m	248 (136)	230 (130)	0.61
6-minute walk test, % predicted	48.5 (26.3)	44.3 (26.8)	0.56
Ventilatory threshold VO ₂ , mL/kg/min	12.1 (3.9)	11.6 (3.9)	0.59
Ventilatory threshold VO ₂ , % predicted	93.1 (24.4)	82.5 (26.2)	0.12
Peak VO ₂ , mL/kg/min	14.3 (4.4)	14.0 (4.8)	0.79
Peak VO ₂ , % predicted	60.4 (15.9)	55.1 (19.3)	0.27

Between-group p-values are from independent t-tests or Fisher exact tests.

Treatment fidelity

In total during the attended treatment sessions, participants initiated 5,006 (~100%) of 5,025 planned overground and treadmill bouts (HIT, 2,383/2,400 [99%]; MAT, 2,623/2,625 [~100%]), and performed 66,598 (99%) of 67,000 planned training minutes (HIT, 31,662/32,000 [99%]; MAT, 34,936/35,000 [~100%]), including intermittent rest breaks.

- Peak training speed was recorded during 4,970 (99%) of the 5,006 initiated bouts (HIT, 2,364/2,383 [99%]; MAT, 2,606/2623 [99%]). Compared with MAT, HIT involved significantly faster training speed during all bouts and training blocks (Figure 4; Table 2). For HIT, training speed was significantly faster on the treadmill versus overground. Conversely for MAT, speed was significantly faster overground versus the treadmill. Training speed significantly increased across sessions for both protocols.
- Valid HR recordings were obtained for 65,666 (99%) of the 66,598 performed training minutes (HIT, 31,211/31,662 [99%]; MAT, 34,456/34,936 [99%]). Compared with MAT, HIT involved significantly higher mean and max HR during all bouts and training blocks. For both HIT and MAT, HR was significantly higher on the treadmill vs overground and HR significantly increased across sessions when expressed relative to baseline HR_{peak} .
- Blood lactate was recorded after the treadmill bout in the middle session of 527 (94%) of 563 attended training weeks (HIT, 246/268 [92%]; MAT, 281/295 [95%]). Compared with MAT, HIT elicited a significantly higher lactate response overall.
- Step counts were recorded during 1,464 (87%) of the 1,675 attended treatment sessions (HIT, 712/800 [89%]; MAT, 752/875 [86%]). Compared with MAT, HIT produced significantly lower step counts during all training blocks. For both HIT and MAT, step counts significantly progressed across sessions, but significantly less so for HIT versus MAT.
- RPE was recorded after 1,632 (97%) of the 1,675 attended treatment sessions (HIT, 761/800 [95%]; MAT, 871/875 [~100%]). Compared with MAT, HIT elicited significantly higher RPE overall.

Figure 4. Treatment intensity. Values are model estimates for each bout of each session, averaging data across participants. Upper panel: Speed data are peak values within a bout. Lower panel: Steady-state mean heart rate (HR) excludes the first 3 minutes of the bout. For treatment (Tx) monitoring, % HR reserve (HRR) was relative to standing HR_{rest} and the most current HR_{peak} , which could increase after sessions 12 and 24. Shading shows aerobic intensity zones: light orange, moderate; light red, vigorous.

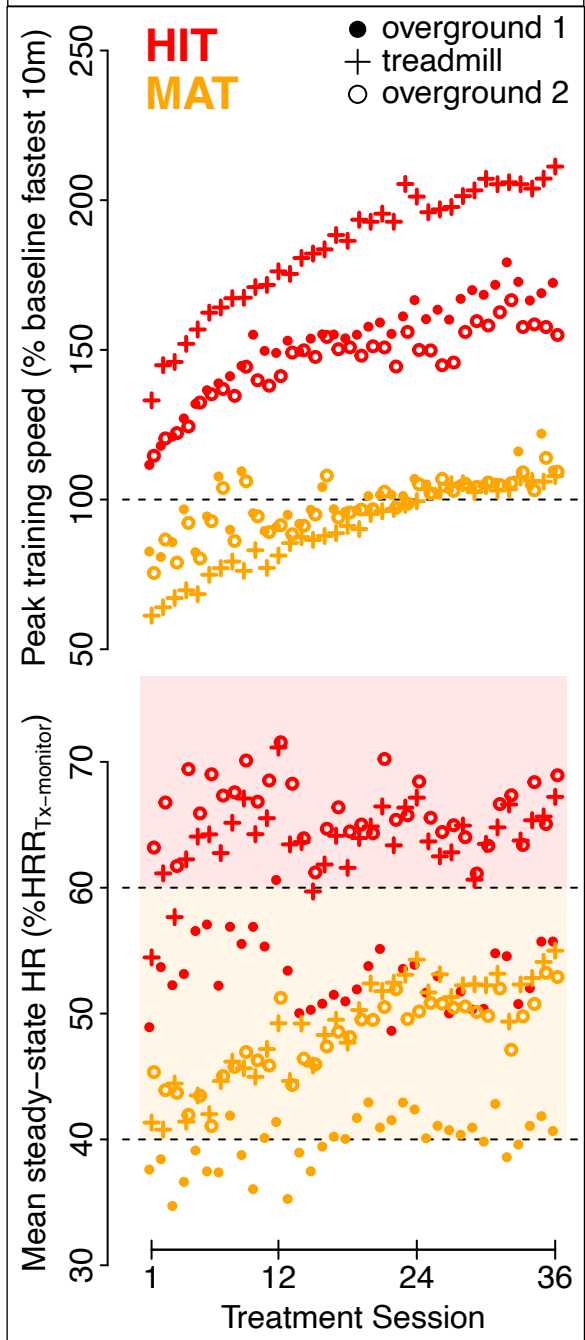


Table 2. Treatment Intensity. Values are model estimates [95% CI]. Speed data are peak values within a bout. Heart rate data are presented as steady-state mean (excluding first 3 minutes) or peak values within a bout. Lactate, step counts and perceived exertion data are singular session values. All data are averaged across sessions and participants.

	HIT (N=27)	MAT (N=28)	HIT - MAT	p- value
Peak bout training speed, m/s				
All bouts	1.23 [1.07, 1.39]	0.76 [0.60, 0.91]	0.47 [0.25, 0.70]	0.0001
Overground bouts	1.20 [1.04, 1.36]	0.78 [0.63, 0.94]	0.41 [0.19, 0.64]	0.0005
Overground 1	1.21 [1.06, 1.37]	0.79 [0.64, 0.95]	0.42 [0.20, 0.64]	0.0004
Overground 2	1.18 [1.02, 1.34]	0.78 [0.62, 0.93]	0.41 [0.18, 0.63]	0.0006
Treadmill bout	1.30 [1.14, 1.46]	0.71 [0.55, 0.86]	0.59 [0.37, 0.82]	<0.0001
Treadmill - overground	0.10 [0.08, 0.12]	-0.08 [-0.09, -0.07]	0.18 [0.16, 0.20]	<0.0001
Mean change per session	0.01 [0.01, 0.01]	0.01 [0.01, 0.01]	0.00 [-0.00, 0.00]	0.4427
Block 1 (sessions 1-12)	1.14 [0.98, 1.29]	0.66 [0.51, 0.82]	0.47 [0.25, 0.69]	<0.0001
Block 2 (sessions 13-24)	1.25 [1.09, 1.41]	0.77 [0.62, 0.93]	0.48 [0.26, 0.70]	<0.0001
Block 3 (sessions 25-36)	1.31 [1.15, 1.47]	0.84 [0.68, 0.99]	0.47 [0.25, 0.70]	<0.0001
Peak bout training speed, % baseline fastest 10-m gait speed				
All bouts	161 [140, 182]	96 [75, 116]	65 [36, 95]	<0.0001
Overground bouts	150 [128, 171]	99 [78, 119]	51 [21, 81]	0.0011
Overground 1	153 [132, 175]	99 [79, 120]	54 [24, 84]	0.0006
Overground 2	146 [125, 167]	98 [77, 119]	48 [19, 78]	0.0020
Treadmill bout	184 [163, 205]	90 [69, 111]	94 [65, 124]	<0.0001
Treadmill - overground	35 [30, 39]	-9 [-11, -6]	43 [38, 48]	<0.0001
Mean change per session	1.5 [1.2, 1.7]	0.9 [0.7, 1.1]	0.6 [0.3, 0.8]	<0.0001
Block 1 (sessions 1-12)	142 [121, 164]	85 [64, 106]	57 [27, 87]	0.0003
Block 2 (sessions 13-24)	165 [144, 187]	96 [75, 117]	69 [39, 99]	<0.0001
Block 3 (sessions 25-36)	176 [155, 197]	106 [85, 127]	70 [40, 100]	<0.0001
Mean bout steady-state HR, %HRR_{Tx-monitor} (relative to standing HR_{rest} and most current HR_{peak})				
All bouts	61 [57, 65]	46 [42, 49]	15 [10, 21]	<0.0001
Overground bouts	60 [56, 64]	44 [40, 48]	16 [10, 21]	<0.0001
Overground 1	53 [49, 57]	40 [36, 44]	13 [8, 19]	<0.0001
Overground 2	66 [62, 70]	48 [44, 52]	18 [13, 23]	<0.0001
Treadmill bout	64 [60, 68]	49 [45, 53]	15 [10, 20]	<0.0001
Treadmill - overground	4.2 [3.6, 4.8]	5.0 [4.4, 5.5]	-0.8 [-1.6, 0.0]	0.0639
Mean change per session	0.01 [-0.06, 0.07]	0.24 [0.19, 0.30]	-0.24 [-0.33, -0.15]	<0.0001
Block 1 (sessions 1-12)	62 [58, 66]	43 [39, 46]	19 [14, 25]	<0.0001
Block 2 (sessions 13-24)	61 [57, 64]	46 [42, 50]	14 [9, 20]	<0.0001
Block 3 (sessions 25-36)	61 [57, 65]	48 [44, 52]	13 [7, 18]	<0.0001
Peak bout HR, %HRR_{Tx-monitor} (relative to standing HR_{rest} and most current HR_{peak})				
All bouts	80 [75, 86]	60 [55, 65]	21 [13, 28]	<0.0001
Overground bouts	77 [72, 83]	57 [52, 62]	20 [13, 28]	<0.0001
Overground 1	72 [66, 77]	52 [47, 58]	19 [12, 27]	<0.0001
Overground 2	83 [78, 88]	61 [56, 67]	22 [14, 29]	<0.0001
Treadmill bout	86 [81, 92]	66 [60, 71]	21 [13, 28]	<0.0001
Treadmill - overground	9.0 [8.1, 9.9]	8.6 [7.9, 9.4]	0.4 [-0.8, 1.6]	0.5050
Mean change per session	-0.24 [-0.33, -0.15]	0.10 [0.03, 0.17]	-0.34 [-0.45, -0.23]	<0.0001
Block 1 (sessions 1-12)	85 [79, 90]	59 [54, 64]	26 [18, 34]	<0.0001
Block 2 (sessions 13-24)	79 [73, 84]	60 [55, 65]	19 [11, 26]	<0.0001
Block 3 (sessions 25-36)	77 [72, 83]	61 [55, 66]	17 [9, 24]	0.0001
Mean bout steady-state HR, %HRR (relative to true HR_{rest} and baseline HR_{peak})				
All bouts	75 [70, 81]	59 [53, 64]	17 [10, 24]	<0.0001
Overground bouts	74 [69, 79]	57 [52, 62]	17 [10, 24]	<0.0001
Overground 1	67 [62, 73]	53 [48, 58]	14 [7, 22]	0.0002
Overground 2	81 [75, 86]	61 [56, 66]	20 [13, 27]	<0.0001
Treadmill bout	78 [73, 84]	62 [57, 67]	16 [9, 24]	<0.0001
Treadmill - overground	4.4 [3.8, 5.0]	5.3 [4.7, 5.9]	-0.8 [-1.7, 0.0]	0.0571

Mean change per session	0.59 [0.52, 0.66]	0.50 [0.45, 0.55]	0.09 [0.01, 0.18]	0.0369
Block 1 (sessions 1-12)	67 [62, 72]	52 [47, 57]	15 [8, 22]	0.0002
Block 2 (sessions 13-24)	77 [72, 82]	59 [54, 64]	18 [11, 26]	<0.0001
Block 3 (sessions 25-36)	82 [77, 87]	65 [59, 70]	17 [10, 25]	<0.0001
Peak bout HR, %HRR (relative to true HR_{rest} and baseline HR_{peak})				
All bouts	94 [87, 100]	71 [65, 78]	22 [13, 32]	<0.0001
Overground bouts	91 [84, 97]	68 [62, 75]	22 [13, 32]	<0.0001
Overground 1	85 [78, 92]	64 [58, 71]	21 [12, 30]	<0.0001
Overground 2	96 [90, 103]	73 [66, 79]	24 [14, 33]	<0.0001
Treadmill bout	100 [93, 106]	77 [71, 84]	23 [13, 32]	<0.0001
Treadmill - overground	9.1 [8.2, 9.9]	8.7 [7.9, 9.6]	0.3 [-0.9, 1.5]	0.5849
Mean change per session	0.50 [0.43, 0.58]	0.49 [0.43, 0.55]	0.02 [-0.08, 0.11]	0.7473
Block 1 (sessions 1-12)	87 [80, 93]	65 [59, 72]	22 [12, 31]	<0.0001
Block 2 (sessions 13-24)	95 [88, 102]	71 [65, 78]	23 [14, 33]	<0.0001
Block 3 (sessions 25-36)	99 [93, 106]	77 [71, 84]	22 [13, 32]	<0.0001
Blood lactate after treadmill bout, mmol/L				
After treadmill bout	3.1 [2.6, 3.6]	2.3 [1.9, 2.7]	0.8 [0.2, 1.5]	0.0141
Mean change per session	0.04 [-0.01, 0.10]	0.01 [-0.02, 0.03]	0.04 [-0.02, 0.10]	0.2285
Block 1 (sessions 1-12)	2.9 [2.3, 3.6]	2.1 [1.7, 2.6]	0.8 [0.0, 1.5]	0.0485
Block 2 (sessions 13-24)	3.0 [2.5, 3.5]	2.4 [1.9, 2.9]	0.6 [-0.1, 1.3]	0.0767
Block 3 (sessions 25-36)	3.4 [2.8, 3.9]	2.3 [1.9, 2.7]	1.1 [0.4, 1.8]	0.0014
Session step count				
After treadmill bout	2,847 [2,529, 3,165]	3,532 [3,070, 3,993]	-685 [-1,244, -125]	0.0166
Mean change per session	13 [9, 16]	24 [21, 27]	-11 [-16, -7]	<0.0001
Block 1 (sessions 1-12)	2,656 [2,364, 2,948]	3,227 [2,737, 3,716]	-570 [-1,139, -2]	0.0494
Block 2 (sessions 13-24)	2,915 [2,596, 3,235]	3,576 [3,124, 4,028]	-661 [-1,213, -108]	0.0192
Block 3 (sessions 25-36)	2,970 [2,620, 3,319]	3,792 [3,344, 4,240]	-822 [-1,389, -255]	0.0046
Session rating of perceived exertion, 6-20				
After treadmill bout	14.4 [13.8, 15.0]	13.4 [12.8, 14.1]	1.0 [0.1, 1.9]	0.0340
Mean change per session	0.03 [0.02, 0.04]	0.02 [0.01, 0.03]	0.01 [-0.01, 0.02]	0.3413
Block 1 (sessions 1-12)	14.0 [13.5, 14.6]	13.1 [12.5, 13.7]	0.9 [0.1, 1.7]	0.0255
Block 2 (sessions 13-24)	14.4 [13.8, 15.0]	13.7 [13.0, 14.3]	0.8 [-0.1, 1.6]	0.0787
Block 3 (sessions 25-36)	14.7 [14.0, 15.4]	13.5 [12.7, 14.3]	1.2 [0.2, 2.3]	0.0262
<i>HR, heart rate; HRR, heart rate reserve; HR_{rest}, resting HR; HR_{peak}, highest HR achieved during exercise testing</i>				

Adverse events

There were no serious adverse events related to study procedures and no significant between-group differences in any adverse event categories (Table 3). Four participants experienced serious adverse events during study participation that were all determined to be unrelated to study procedures, including: a seizure leading to temporary hospitalization (n=1, MAT group), a fall with hip fracture (n=1, MAT group), an episode of delirium leading to temporary hospitalization (n=1, MAT group) and a recurrent stroke (n=1, HIT group).

Table 3. Adverse Events (AEs). Values are N participants with AE (total N AEs).

	HIT (N=27) 800 sessions	MAT (N=28) 875 sessions	HIT/MAT odds ratio [95% CI]*
Any post-randomization AE	25 (109)	23 (84)	3.2 [0.5, 20.5]
-Grade 1 (mild)	22 (80)	22 (62)	1.2 [0.3, 5.2]
-Grade 2 (moderate)	14 (28)	9 (19)	2.7 [0.8, 9.1]
-Grade 3 (severe)	1 (1)	3 (3)	0.3 [0.0, 3.4]
-Grade 4-5 (life threat-death)	0 (0)	0 (0)	1 [NA]
-Pain/Soreness	12 (25)	11 (32)	1.3 [0.4, 4.0]
-Fatigue	4 (4)	5 (6)	0.8 [0.2, 3.6]
-Lightheadedness	6 (17)	7 (15)	0.9 [0.2, 3.3]
-Fall	9 (36)	11 (14)	0.7 [0.2, 2.3]
Definitely treatment-related	13 (39)	10 (29)	1.7 [0.6, 5.2]
-Grade 1 (mild)	11 (26)	9 (24)	1.5 [0.5, 4.7]
-Grade 2 (moderate)	7 (13)	4 (5)	2.1 [0.5, 8.4]
-Grade 3-5 (severe-death)	0 (0)	0 (0)	1 [NA]
-Pain/Soreness	8 (15)	7 (18)	1.3 [0.4, 4.4]
-Fatigue	2 (2)	2 (2)	1.1 [0.1, 8.7]
-Lightheadedness	3 (10)	4 (8)	0.8 [0.1, 4.0]
-Fall	3 (4)	0 (0)	8.2 [0.4, 167] [†]
Possibly/not treatment-related	22 (70)	19 (55)	2.3 [0.6, 8.5]
-Grade 1 (mild)	19 (54)	17 (38)	1.6 [0.5, 5.2]
-Grade 2 (moderate)	10 (15)	7 (14)	2.4 [0.6, 9.7]
-Grade 3 (severe)	1 (1)	3 (3)	0.3 [0.0, 3.4]
-Grade 4-5 (life threat-death)	0 (0)	0 (0)	1 [NA]
-Pain/Soreness	7 (10)	8 (14)	0.9 [0.3, 3.2]
-Fatigue	2 (2)	3 (4)	0.7 [0.1, 5.9]
-Lightheadedness	3 (7)	5 (7)	0.6 [0.1, 2.9]
-Fall	9 (32)	11 (14)	0.7 [0.2, 2.3]

*Relative odds of a participant having an AE, adjusted for site and baseline gait dysfunction. [†]When AEs were only present in one group, continuity correction added 0.5 AEs to each group to permit (unadjusted) ratio calculation.

Primary outcome (6-minute walk test) changes

Within the HIT group, 6MWT distance significantly increased between each testing time point (Figure 5). Within the MAT group, the 6MWT did not significantly increase between any two consecutive time points but was significantly higher than PRE at the 8WK and 12WK time points.

At the 4WK time point, there was no significant difference in 6MWT gains between groups (Table 4). At the 8WK and 12WK time points, the HIT group had improved significantly more than the MAT group by 29 m [95% CI: 5, 54 m] and 44 [14, 74] m, respectively.

Secondary outcome changes

Both groups exhibited significant increases in self-selected gait speed, fastest gait speed and ventilatory threshold VO₂ at various time points relative to PRE (Table 4), with the HIT group showing significantly greater increases than MAT in gait speeds. Only HIT significantly decreased PROMIS-Fatigue scores and only at the 8WK time point, which was significantly different between groups.

Other exercise test measures to facilitate interpretation and exploratory outcomes

Other exercise capacity outcomes showed significant increases in both groups at various time points (Table 5), with some measures increasing significantly more for HIT versus MAT at 4WK (GXT duration, peak GXT speed and peak GXT HR) or 12WK (peak GXT HR and HR_{peak}). The verification test increased HR_{peak} by 2.7 [1.0, 4.4] bpm at baseline and 2.0 [1.3, 2.7] bpm across all time points. Baseline HR_{peak} was 87.0 ± 17.2% age-predicted maximal HR (unadjusted mean ± SD). Peak respiratory

exchange ratio and rating of perceived exertion during the baseline GXT were 1.01 ± 0.11 and 15.2 ± 2.6, respectively, and did not significantly change over time in either group. Meanwhile, there were significant improvements in both groups for ventilatory threshold treadmill speed, metabolic cost of gait in the last 3 minutes of the GXT, and functional ambulation category score, with no significant differences between groups. Only the HIT group reported significant improvement in overall quality of life (lower EuroQoL misery score) and performance of usual activities (lower EuroQoL score for that item), and this was only at 4WK, with no significant between-group differences after FDR correction.

Figure 5. Primary outcome (6MWT) changes

Values are model estimates and error bars are 95% CI.

‡p<0.05 change (Δ) between time points

†p<0.05 Δ from baseline (PRE)

*pFDR<0.05 between groups for Δ from PRE

FDR, false discovery rate corrected across time points; 6MWT, 6-minute walk test; CID, clinically important difference (20-50m); WK, week

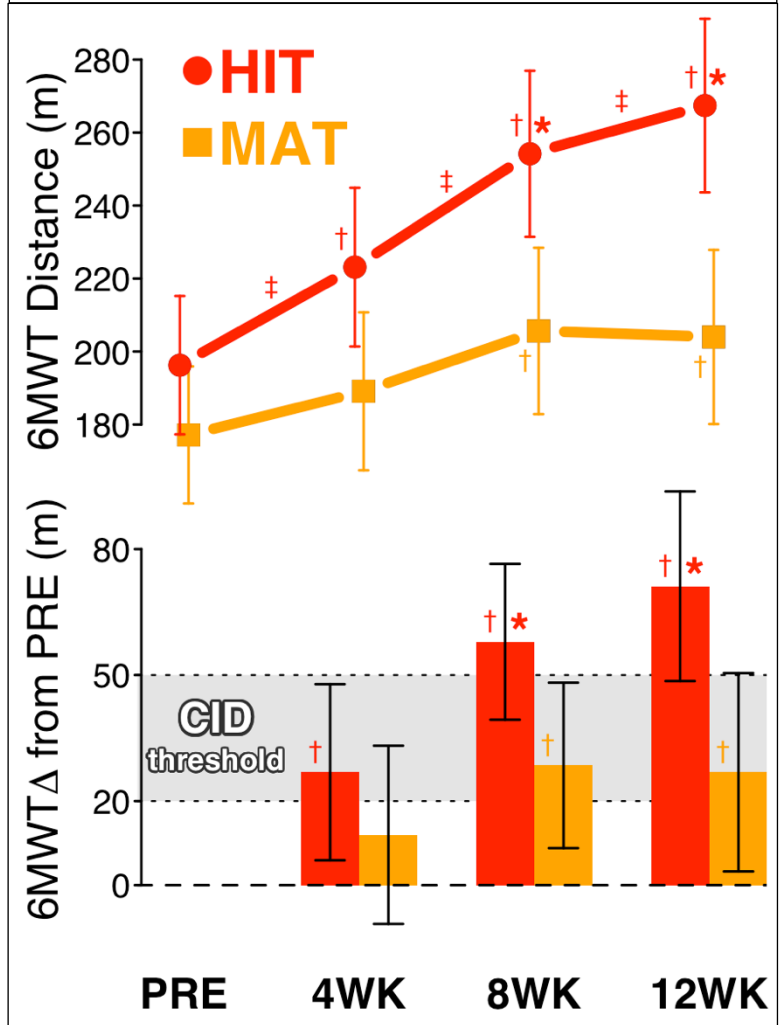


Table 4. Primary and secondary outcome changes

	HIT (N=27)	MAT (N=28)	HIT - MAT	p-value
Walking Capacity: 6-Minute Walk Distance, m (primary outcome)				
Baseline (PRE)	196 (98)	177 (99)	19 [-31, 69]	0.4438
4-Week Change	27 [6, 48]	12 [-9, 33]	15 [-13, 42]	0.2824
8-Week Change	58 [39, 76]	29 [9, 48]	29 [5, 54]	0.0215*
12-Week Change	71 [49, 94]	27 [3, 50]	44 [14, 74]	0.0048*

Self-Selected Gait Speed: 10-meter Walk Test, m/s				
Baseline	0.52 (0.20)	0.49 (0.21)	0.03 [-0.07, 0.13]	0.5607
4-Week Change	0.11 [0.06, 0.15]	0.02 [-0.02, 0.07]	0.08 [0.02, 0.15]	0.0091*
8-Week Change	0.14 [0.08, 0.20]	0.06 [0.00, 0.12]	0.08 [0.00, 0.15]	0.0426
12-Week Change	0.19 [0.13, 0.25]	0.06 [0.00, 0.12]	0.13 [0.05, 0.20]	0.0026*
Fastest Gait Speed: 10-meter Walk Test, m/s				
Baseline	0.70 (0.32)	0.62 (0.32)	0.08 [-0.08, 0.24]	0.3257
4-Week Change	0.22 [0.16, 0.28]	0.01 [-0.05, 0.07]	0.21 [0.13, 0.29]	<0.0001*
8-Week Change	0.24 [0.17, 0.32]	0.09 [0.01, 0.17]	0.15 [0.05, 0.25]	0.0029*
12-Week Change	0.28 [0.19, 0.37]	0.09 [-0.01, 0.18]	0.20 [0.08, 0.32]	0.0016*
Fatigue: PROMIS-Fatigue Scale, T score (lower score = less fatigue)				
Baseline	52.4 (8.6)	50.8 (8.7)	1.7 [-2.7, 6.0]	0.4473
4-Week Change	-1.7 [-4.0, 0.6]	0.0 [-2.3, 2.3]	-1.7 [-4.8, 1.3]	0.2590
8-Week Change	-3.0 [-5.5, -0.5]	1.0 [-1.6, 3.6]	-4.0 [-7.3, -0.7]	0.0185*
12-Week Change	-1.1 [-3.7, 1.5]	-0.1 [-2.7, 2.5]	-1.0 [-4.4, 2.5]	0.5678
Ventilatory Threshold Oxygen Consumption Rate (VO₂), mL/kg/min				
Baseline	11.5 (4.0)	10.9 (4.0)	0.6 [-1.4, 2.6]	0.5522
4-Week Change	1.3 [0.4, 2.2]	0.4 [-0.5, 1.3]	0.9 [-0.4, 2.1]	0.1609
8-Week Change	2.0 [0.7, 3.3]	1.7 [0.3, 3.1]	0.3 [-1.4, 2.1]	0.7209
12-Week Change	1.9 [0.4, 3.4]	1.4 [-0.2, 3.0]	0.5 [-1.5, 2.6]	0.6018
Baseline data presented as (covariate-adjusted) mean (SD) Change and difference data presented as mean difference [95% CI] *p<0.05 after false discovery rate-correction across time points & measures (primary & secondary) PROMIS, Patient Reported Outcomes Measurement Information System				

Table 5. Other exercise test measures to facilitate interpretation and exploratory outcomes				
	HIT (N=27)	MAT (N=28)	HIT - MAT	p-value
Graded Exercise Test (GXT) Duration, minutes				
Baseline (PRE)	9.6 (4.2)	8.9 (4.2)	0.7 [-1.4, 2.8]	0.5212
4-Week Change	1.8 [1.1, 2.4]	0.6 [-0.1, 1.2]	1.2 [0.3, 2.1]	0.0078*
8-Week Change	2.5 [1.5, 3.4]	1.6 [0.6, 2.5]	0.9 [-0.3, 2.1]	0.1537
12-Week Change	3.3 [2.2, 4.4]	2.2 [1.1, 3.4]	1.1 [-0.4, 2.6]	0.1350
Peak Treadmill Speed During GXT, m/s				
Baseline	0.69 (0.36)	0.64 (0.37)	0.05 [-0.14, 0.23]	0.6081
4-Week Change	0.15 [0.09, 0.22]	0.04 [-0.03, 0.11]	0.11 [0.03, 0.20]	0.0093*
8-Week Change	0.17 [0.09, 0.26]	0.13 [0.04, 0.22]	0.04 [-0.07, 0.16]	0.4391
12-Week Change	0.24 [0.14, 0.33]	0.20 [0.10, 0.30]	0.04 [-0.09, 0.16]	0.5570
Peak Oxygen Consumption Rate During GXT (VO_{2-peak}), mL/kg/min				
Baseline	13.4 (4.4)	13.0 (4.4)	0.4 [-1.8, 2.6]	0.7160
4-Week Change	1.5 [0.5, 2.5]	0.2 [-0.8, 1.2]	1.3 [0.0, 2.6]	0.0478
8-Week Change	2.1 [0.8, 3.5]	1.5 [0.0, 2.9]	0.7 [-1.2, 2.5]	0.4778
12-Week Change	2.5 [0.8, 4.2]	1.5 [-0.3, 3.3]	1.1 [-1.2, 3.4]	0.3562
Peak Heart Rate During GXT, bpm				
Baseline	123.4 (21.1)	128.2 (21.3)	-4.9 [-15.5, 5.8]	0.3663
4-Week Change	5.4 [0.9, 9.9]	-3.7 [-8.2, 0.7]	9.1 [3.3, 15.0]	0.0030*
8-Week Change	4.2 [-1.5, 9.9]	-2.2 [-8.2, 3.9]	6.4 [-1.3, 14.0]	0.0994
12-Week Change	12.6 [6.5, 18.6]	-1.0 [-7.2, 5.2]	13.6 [5.6, 21.6]	0.0013*
Highest Heart Rate During GXT or Verification Test (HR_{peak}), bpm				
Baseline	125.0 (22.1)	131.4 (22.3)	-6.4 [-17.6, 4.8]	0.2562
4-Week Change	5.7 [0.2, 11.1]	-3.7 [-9.2, 1.7]	9.4 [2.2, 16.5]	0.0111
8-Week Change	4.1 [-1.7, 10.0]	-3.4 [-9.6, 2.8]	7.5 [-0.3, 15.4]	0.0605
12-Week Change	11.2 [4.7, 17.8]	-2.6 [-9.3, 4.0]	13.9 [5.2, 22.5]	0.0023*

HR_{peak}, % Age-Predicted Maximal HR				
Baseline	83.3 (18.4)	90.0 (18.5)	-6.7 [-16.0, 2.6]	0.1528
4-Week Change	3.8 [0.2, 7.4]	-2.4 [-6.0, 1.2]	6.2 [1.4, 11.0]	0.0119
8-Week Change	3.4 [-0.7, 7.4]	-2.1 [-6.4, 2.2]	5.4 [0.0, 10.9]	0.0497
12-Week Change	7.8 [3.1, 12.6]	-2.2 [-7.0, 2.7]	10.0 [3.7, 16.3]	0.0026*
Peak Respiratory Exchange Ratio During GXT (VCO₂/VO₂)				
Baseline	1.01 (0.12)	1.02 (0.12)	-0.01 [-0.07, 0.05]	0.7949
4-Week Change	-0.01 [-0.04, 0.02]	-0.03 [-0.06, 0.00]	0.01 [-0.03, 0.05]	0.5181
8-Week Change	0.03 [-0.02, 0.07]	-0.02 [-0.06, 0.03]	0.04 [-0.01, 0.10]	0.1432
12-Week Change	0.00 [-0.04, 0.04]	-0.01 [-0.05, 0.03]	0.00 [-0.05, 0.06]	0.8689
Rating of Perceived Exertion (RPE) from GXT, 6-20				
Baseline	14.2 (2.5)	15.1 (2.5)	-0.9 [-2.1, 0.4]	0.1655
4-Week Change	0.4 [-1.0, 1.8]	-0.5 [-1.9, 0.8]	1.0 [-0.8, 2.7]	0.2796
8-Week Change	0.7 [-0.5, 1.9]	0.1 [-1.1, 1.3]	0.6 [-1.0, 2.1]	0.4559
12-Week Change	0.9 [-0.2, 2.0]	0.6 [-0.5, 1.6]	0.3 [-1.0, 1.7]	0.6185
Ventilatory Threshold Treadmill Speed, m/s				
Baseline	0.60 (0.32)	0.53 (0.33)	0.07 [-0.09, 0.23]	0.3972
4-Week Change	0.13 [0.07, 0.18]	0.06 [0.00, 0.11]	0.07 [-0.01, 0.14]	0.0691
8-Week Change	0.16 [0.09, 0.23]	0.12 [0.05, 0.19]	0.04 [-0.05, 0.13]	0.3615
12-Week Change	0.21 [0.13, 0.28]	0.19 [0.11, 0.27]	0.02 [-0.08, 0.12]	0.7148
Metabolic Cost of Treadmill Gait During Last 3 Minutes of GXT, mL O₂/kg/m				
Baseline	0.44 (0.19)	0.47 (0.19)	-0.03 [-0.12, 0.07]	0.5372
4-Week Change	-0.08 [-0.13, -0.03]	-0.05 [-0.10, 0.00]	-0.03 [-0.10, 0.03]	0.3293
8-Week Change	-0.10 [-0.14, -0.05]	-0.08 [-0.13, -0.03]	-0.01 [-0.08, 0.05]	0.6751
12-Week Change	-0.12 [-0.17, -0.06]	-0.11 [-0.17, -0.06]	0.00 [-0.07, 0.06]	0.9025
Metabolic Cost of Treadmill Gait During Last 3 Minutes of Shortest GXT, mL O₂/kg/m				
Baseline	0.45 (0.18)	0.48 (0.18)	-0.03 [-0.12, 0.06]	0.5095
4-Week Change	-0.01 [-0.04, 0.02]	-0.01 [-0.04, 0.02]	0.00 [-0.04, 0.04]	0.9400
8-Week Change	-0.02 [-0.06, 0.02]	-0.04 [-0.08, 0.00]	0.02 [-0.03, 0.08]	0.4346
12-Week Change	-0.01 [-0.06, 0.03]	-0.04 [-0.08, 0.01]	0.02 [-0.04, 0.08]	0.4298
Gait Stability: Functional Ambulation Category, 0-4				
Baseline	3.16 (0.58)	3.19 (0.58)	-0.03 [-0.32, 0.26]	0.8305
4-Week Change	-0.02 [-0.24, 0.20]	0.17 [-0.05, 0.39]	-0.19 [-0.48, 0.10]	0.1950
8-Week Change	0.22 [-0.03, 0.46]	0.31 [0.06, 0.56]	-0.09 [-0.42, 0.24]	0.5727
12-Week Change	0.26 [0.02, 0.50]	0.27 [0.03, 0.51]	-0.01 [-0.33, 0.31]	0.9475
Activities-Specific Balance Confidence (ABC) Scale Score, 0-100				
Baseline	59.6 (18.9)	63.0 (19.1)	-3.4 [-12.9, 6.2]	0.4815
4-Week Change	-2.7 [-7.2, 1.7]	-4.6 [-9.1, -0.1]	1.9 [-4.0, 7.7]	0.5266
8-Week Change	1.6 [-3.1, 6.4]	-1.8 [-6.7, 3.1]	3.5 [-2.8, 9.7]	0.2757
12-Week Change	2.5 [-2.7, 7.7]	0.6 [-4.7, 5.9]	1.9 [-5.0, 8.8]	0.5758
EuroQOL-5D-5L Misery Score, 5-25 (higher score = worse self-reported health problems)				
Baseline	11.0 (2.7)	10.7 (2.7)	0.3 [-1.1, 1.7]	0.6669
4-Week Change	-0.9 [-1.8, -0.1]	-0.2 [-1.1, 0.7]	-0.7 [-1.9, 0.4]	0.2158
8-Week Change	-0.6 [-1.5, 0.3]	-0.4 [-1.3, 0.5]	-0.2 [-1.3, 1.0]	0.7859
12-Week Change	-0.1 [-1.0, 0.8]	-0.6 [-1.5, 0.3]	0.5 [-0.7, 1.7]	0.3777
EuroQOL-5D-5L Mobility Score, 1-5 (higher score = worse self-reported problems)				
Baseline	2.56 (0.86)	2.51 (0.87)	0.05 [-0.38, 0.49]	0.8081
4-Week Change	-0.17 [-0.58, 0.24]	-0.15 [-0.56, 0.26]	-0.01 [-0.55, 0.52]	0.9573
8-Week Change	-0.17 [-0.60, 0.27]	-0.13 [-0.58, 0.32]	-0.04 [-0.62, 0.54]	0.8921
12-Week Change	-0.23 [-0.64, 0.18]	-0.24 [-0.65, 0.17]	0.01 [-0.53, 0.54]	0.9824
EuroQOL-5D-5L Usual Activities Score, 1-5 (higher score = worse self-reported problems)				
Baseline	2.53 (0.88)	2.23 (0.89)	0.30 [-0.15, 0.74]	0.1822
4-Week Change	-0.41 [-0.71, -0.11]	0.02 [-0.28, 0.31]	-0.43 [-0.82, -0.04]	0.0320
8-Week Change	-0.25 [-0.63, 0.14]	0.11 [-0.29, 0.52]	-0.36 [-0.88, 0.15]	0.1615
12-Week Change	-0.16 [-0.57, 0.24]	-0.07 [-0.48, 0.34]	-0.09 [-0.63, 0.44]	0.7262

Baseline data presented as (covariate-adjusted) mean ± SD
Change and difference data presented as mean difference [95% CI]
*p<0.05 after false discovery rate-correction across timepoints & all measures
PROMIS, Patient Reported Outcomes Measurement Information System

Primary outcome changes stratified by baseline gait dysfunction subgroup

Among the 41 participants (20 HIT, 21 MAT) with mild/moderate baseline gait dysfunction (self-selected speed 0.4-1.0 m/s), 6MWT distance significantly increased in both the HIT and MAT groups, with significantly greater improvement in HIT at 12WK (Figure 6, Table 6). Among the 14 participants (7 HIT, 7 MAT) with severe baseline gait dysfunction (self-selected speed <0.4 m/s), 6MWT distance significantly increased in the HIT group only, but with no significant difference between groups.

Table 6. Primary outcome (6-Minute Walk Distance) changes stratified by baseline gait dysfunction

Mild/Moderate Baseline Dysfunction (gait speed 0.4-1.0 m/s; N=41)				
	HIT (N=20)	MAT (N=21)	HIT - MAT	p-value
Baseline (PRE)	305 (93)	285 (93)	21 [-38, 79]	0.4816
4-Week Change	31 [8, 55]	20 [-2, 42]	11 [-21, 43]	0.4960
8-Week Change	72 [50, 93]	43 [24, 62]	29 [0, 57]	0.0490
12-Week Change	88 [62, 114]	44 [20, 67]	44 [9, 79]	0.0142*
Severe Baseline Dysfunction (gait speed <0.4 m/s; N=14)				
	HIT (N=7)	MAT (N=7)	HIT - MAT	p-value
Baseline (PRE)	86 (93)	71 (93)	15 [-85, 115]	0.7656
4-Week Change	26 [-12, 64]	-1 [-42, 40]	27 [-29, 83]	0.3348
8-Week Change	45 [12, 77]	14 [-29, 56]	31 [-22, 85]	0.2464
12-Week Change	55 [15, 95]	10 [-40, 60]	45 [-19, 109]	0.1656

Baseline data presented as (covariate-adjusted) mean ± SD
 Change and difference data presented as mean difference [95% CI]
 *p<0.05 after false discovery rate-correction across time points

Sensitivity of primary results to missing data assumptions

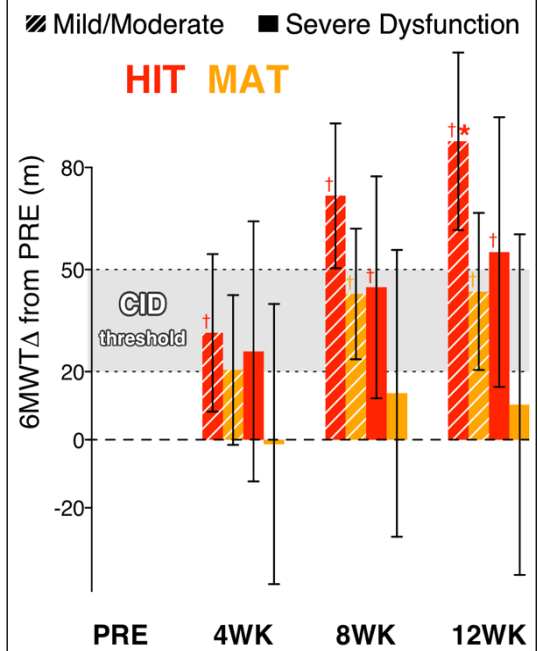
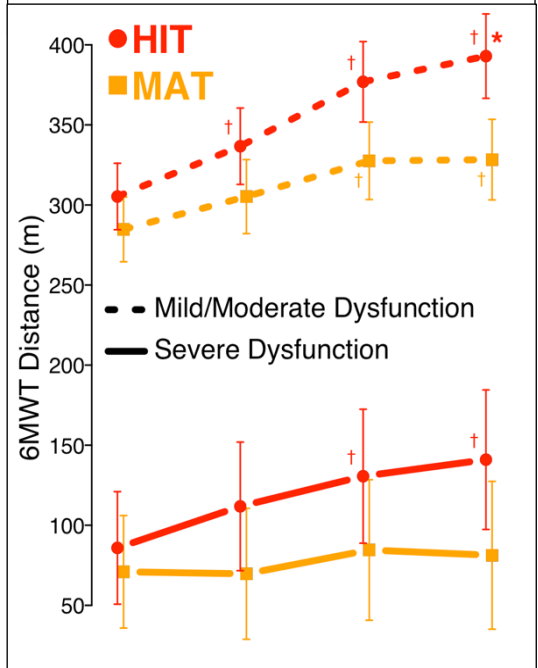
When assuming that participants who dropped out due to an adverse event had no mean improvement (rather than the common “missing at random” assumption), both groups still showed significant improvements in 6MWT distance and the direction of the between-group differences remained unchanged, but the between-group effect sizes were smaller and no longer statistically significant (Table 7).

Table 7. Estimated 6-minute walk test changes when assuming no mean improvement in cases of adverse-event-related dropout

	HIT (N=27)	MAT (N=28)	HIT - MAT	p-value
Walking Capacity: 6-Minute Walk Distance, m (primary outcome)				
Baseline (PRE)	196 (98)	177 (99)	19 [-31, 69]	0.4438
4-Week Change	26 [7, 46]	13 [-7, 32]	14 [-12, 40]	0.2992
8-Week Change	47 [28, 66]	28 [9, 47]	19 [-6, 44]	0.1325
12-Week Change	52 [30, 73]	28 [6, 49]	24 [-4, 53]	0.0985

Figure 6. Primary analysis stratified by baseline gait dysfunction.

Values are model estimates and error bars are 95% CI. Mild/moderate dysfunction, gait speed 0.4-1.0 m/s; severe dysfunction, <0.4 m/s.
 †p<0.05 change (Δ) from baseline (PRE)
 *pFDR<0.05 between groups for Δ from PRE
 FDR, false discovery rate corrected across time points; 6MWT, 6-minute walk test; CID, clinically important difference; WK, week



DISCUSSION

Results of this multi-center trial gave proof of concept that the vigorous training intensity of HIT is a crucial dosing parameter for locomotor exercise in chronic stroke. Findings also indicated that locomotor HIT can produce significant and meaningful gains in walking capacity in 4 weeks, but that a training duration of at least 12 weeks is needed to maximize immediate gains in walking capacity with this intervention. Our primary hypothesis that HIT would improve walking capacity significantly more than MAT after 4 weeks of training was not supported (mean 6MWT change in HIT vs MAT: +15 meters [95% CI: -13-42]). However, the gap between groups continued to widen over time, and HIT elicited significantly greater improvement in walking capacity than MAT after 8 weeks (+29 [5-54] meters) and 12 weeks of training (+44 [14-74] meters) in the prespecified primary analysis, supporting the secondary hypotheses. Among the secondary and exploratory outcomes, HIT also elicited significantly greater improvements than MAT in gait speed (self-selected and fastest), fatigue, and exercise capacity (GXT duration, peak GXT speed and peak HR) at various time points after adjustment for multiple comparisons (Tables 4 & 5). In addition, intensive safety monitoring was successful at identifying expected AEs and results suggest that post-stroke HIT continues to be sufficiently safe for further study.

Within the MAT group, the mean 6MWT gain of +27 meters [95% CI: 3-50] after 12 weeks of training was comparable to many similar previous studies in chronic stroke,^{25,26,29-31,33,115} and exceeded the minimally important difference of +20 meters.⁸⁶ This confirms that MAT was successfully implemented in this study and has some value for improving walking capacity in this population. Surprisingly, our results suggest that most outcomes reached an apparent plateau after just 8 weeks of MAT, which may not support the longer 3-6 month training durations used in most previous stroke trials focused on MAT. However, there are prior indications that outcomes may continue to slowly improve between 3-6 months of MAT.^{28,33} In addition, it is plausible that longer training durations could result in more sustained gains in walking capacity or greater long-term improvements in cardiovascular-metabolic health, but this has not been thoroughly tested.

Within the HIT group, the mean 6MWT gain after 4 weeks of training was comparable to that found in prior single-site 4-week pilot testing,⁶³ confirming successful multi-center implementation. As hypothesized, gains in walking capacity and other outcomes were even larger after 8 and 12 weeks of HIT, with no apparent plateau. These findings suggest that most prior post-stroke HIT studies have used insufficient training durations, and that HIT durations of at least 12 weeks should be considered in the future. Outpatient therapy durations longer than 12 weeks (36 sessions) are not likely to be routinely feasible in the current model for clinical rehabilitation due to reimbursement constraints¹¹⁶⁻¹¹⁸ and adherence issues.^{119,120} Thus, a 12-week duration for clinician-led HIT program may be an optimal standard to target.

Our prespecified subgroup analysis suggested that the vigorous training intensity of HIT may have been particularly important for participants with severe baseline gait dysfunction (self-selected gait speed <0.4 m/s^{62,63}). For these individuals, MAT elicited mean 6MWT gains less than the minimally important difference of +20 meters⁸⁶ (+10 [-40-60] meters). Previous research also found minimal response to MAT among stroke survivors with severe gait dysfunction,^{43,62} leading to prior suggestions that these patients may not be good candidates for gait training.⁶² However, in the current study, participants with severe gait dysfunction increased mean 6MWT distance by +55 [15-95] meters with HIT, which is above the +50-meter threshold for a large clinically-meaningful gain.⁸⁶ This study was not powered to detect significant differences between HIT and MAT within the severe dysfunction subgroup (N=14, p=0.17) or to assess within-group changes with sufficient precision to rule meaningful benefits in or out with confidence, so these findings should be taken as preliminary. However, HIT did show significantly greater improvement than MAT in the larger mild/moderate dysfunction subgroup (+88 [62-114] vs +44 [20-67], N=41, p=0.01), suggesting that vigorous training intensity could be optimal regardless of baseline dysfunction.

The intensity-dependent improvement (decrease) in PROMIS-Fatigue T scores after 8 weeks of training (HIT -3.0 [-5.5, -0.5], MAT 1.0 [-1.6, 3.6], p=0.02) is especially noteworthy because identifying a treatment for post-stroke fatigue has been highlighted as a research priority.¹²¹ This finding is consistent with prior neurophysiologic evidence showing that post-stroke fatigue is related to lower corticomotor excitability¹²² and that HIT elicits significantly greater acute increases in corticomotor excitability than MAT.¹²³ Interestingly, participants randomized to HIT reported this significant decrease in fatigue during a time period when they were coming to the research sites multiple times per week for training sessions that commonly *induced* temporary fatigue. One possible explanation for this apparent paradox is that the acute central motor activation effects of post-stroke HIT appear to counterbalance or exceed the effects of fatigue.¹²³ It is also possible that the greater longitudinal increases in walking capacity and exercise tolerance with HIT may have reduced fatigue with everyday activities, especially when coupled with significant reductions in the metabolic cost of

gait. Oddly, HIT group improvements and between-group differences in fatigue were only statistically significant after 8 weeks of training and not after 4 or 12 weeks of training. Yet all these effect estimates were in the same direction, suggesting that the lack of significance at the 4 and 12-week timepoints may have been due to insufficient statistical power, rather than an exclusive fatigue benefit at only the 8-week time point. Conversely, it is always possible that the significant improvement in fatigue observed after 8 weeks of HIT was due to chance, but such improvement was also observed in a prior 4-week pilot study,⁶³ suggesting otherwise. Confirming these findings in a larger trial with follow up testing is warranted.

Both groups showed significant improvement in aerobic capacity (ventilatory threshold VO_2) and VO_2 -peak. The lack of significant between-group differences in aerobic capacity gains suggests that the greater improvement in walking capacity with HIT vs MAT may have been primarily driven by neuromotor rather than cardiopulmonary adaptations. Although, it is also possible that this lack of significant between-group differences was due to greater measurement error for VO_2 -based outcomes (ventilatory threshold and peak) or confounding effects of gait efficiency changes on VO_2 during exercise testing (improved gait efficiency decreases measured VO_2).^{89,90,93,94} As expected, the symptom-limited GXT was primarily terminated by non-cardiorespiratory factors on average (respiratory exchange ratio < 1.05, RPE < 17, HR \leq 90% age-predicted max), so GXT duration and peak GXT measures cannot be interpreted as specific measures of cardiopulmonary (aerobic) capacity, but can still be interpreted as measures of *exercise capacity*, which can be influenced by neuromotor function, cardiopulmonary capacity and other factors (e.g. motivation).

Importantly, significant improvements in the metabolic cost of gait and gait stability for both groups suggest that compensatory movements and compromised stability were not the primary mechanisms by which individuals achieved gains in walking capacity with training. However, it did appear that gains in efficiency were primarily related to gains in gait speed, since efficiency at matched speeds did not significantly change.

Strengths and limitations

Most recent post-stroke HIT studies have focused on improving aerobic fitness among persons with mild stroke or transient ischemic attack who had unimpaired, near-normal or unreported baseline walking function.^{41,42,44,46,47,124,125} Thus, one important strength of this trial is its relevance to clinical rehabilitation, since we focused on improving walking capacity as the primary outcome and recruited a sample with impaired gait (self-selected speed, 0.63 ± 0.31 m/s, which was $49 \pm 24\%$ of normative predicted values) for therapist-led training. Another major strength of this trial is our systematic AE data collection with blinded adjudication, which is rare in the rehabilitation literature, but critical for understanding the risk/benefit ratio of an intervention.

In terms of rigor, this study used randomization, concealed allocation, assessor blinding and intent-to-treat analysis to minimize risk of bias¹¹³ and only had 10% missing outcome data. We also added the rigor of multi-center implementation earlier than usual in the rehabilitation research continuum,¹²⁶ increasing confidence in reproducibility and generalizability of the results. In addition, we included more rigorous assessment than many clinical trials about how missing data might influence the results.¹¹³ Further, we successfully implemented several novel technological methods to enhance protocol fidelity, including automated real-time prompts, calculations and feedback from the direct electronic data entry system during all visits, plus automated bout timing, recording & intensity guidance from a fully wireless HR monitoring system during treatment visits.

The primary limitation to this trial was its relatively modest sample size (N=55), since it was designed to be an early proof of concept trial with a secondary aim of optimizing training duration dosing. This meant that we were likely underpowered to detect all but very large differences in AEs between HIT and MAT. The small sample size is also probably why the primary findings were not fully robust to different plausible assumptions about the true values of the missing data (Table 7). This means that the findings should not be viewed as definitive and that they require verification in a larger trial. This sensitivity analysis also validates the general intuition that larger samples are needed to have more confidence in the results of clinical trials.

Another important limitation was the lack of follow up testing to assess sustained effects after treatment ended. This also makes it more challenging to interpret the results of the questionnaires (PROMIS-Fatigue, ABC scale and EuroQOL-5D-5L) since they were asking participants to report about time periods in which they were engaged in the study treatment at 4WK, 8WK and 12WK, but asking about a time period when participants were largely just participating in normal daily activities at PRE. We attempted to mitigate this issue by specifically asking participants to answer questions based on their usual activities outside of the study but are skeptical that this strategy was completely successful. Thus, a larger trial with post-treatment follow-up testing is needed to confirm and expand on these promising results.

Conclusions

For locomotor exercise among stroke survivors with chronic gait dysfunction, vigorous training intensity appears to be significantly better than moderate intensity for eliciting immediate improvements in walking capacity and other outcomes. With this vigorous intensity locomotor exercise, 12 weeks of training produces greater gains than shorter durations. Preliminary results among patients with severe baseline gait dysfunction (self-selected speed <0.4 m/s) suggest that a vigorous training intensity could possibly be necessary for this subgroup to have meaningful benefits.

REFERENCES

1. Mayo NE, Wood-Dauphinee S, Cote R, Durcan L, Carlton J. Activity, participation, and quality of life 6 months poststroke. *Arch Phys Med Rehabil.* 2002;83(8):1035-1042.
2. Jorgensen HS, Nakayama H, Raaschou HO, Vive-Larsen J, Stoier M, Olsen TS. Outcome and time course of recovery in stroke. part II: Time course of recovery. the copenhagen stroke study. *Arch Phys Med Rehabil.* 1995;76(5):406-412.
3. Hill K, Ellis P, Bernhardt J, Maggs P, Hull S. Balance and mobility outcomes for stroke patients: A comprehensive review. *Aust J Physiother.* 1997;43:173-80.
4. Lohse K, Bland MD, Lang CE. Quantifying change during outpatient stroke rehabilitation: A retrospective regression analysis. *Arch Phys Med Rehabil.* 2016;97(9):1423-1430.e1. PMID: PMC5003650.
5. Rudberg AS, Berge E, Laska AC, Jutterstrom S, Nasman P, Sunnerhagen KS, Lundstrom E. Stroke survivors' priorities for research related to life after stroke. *Top Stroke Rehabil.* 2020:1-6.
6. Billinger SA, Arena R, Bernhardt J, Eng JJ, Franklin BA, Johnson CM, MacKay-Lyons M, Macko RF, Mead GE, Roth EJ, Shaughnessy M, Tang A, on behalf of the American Heart Association Stroke Council, Council on Cardiovascular and Stroke Nursing, Council on Lifestyle and Cardiometabolic Health, Council on Epidemiology and Prevention, and Council on Clinical Cardiology. Physical activity and exercise recommendations for stroke survivors: A statement for healthcare professionals from the american heart association/american stroke association. *Stroke.* 2014;45(8):2532-53.
7. Winstein CJ, Stein J, Arena R, Bates B, Chorney LR, Cramer SC, Deruyter F, Eng JJ, Fisher B, Harvey RL, Lang CE, MacKay-Lyons M, Ottenbacher KJ, Pugh S, Reeves MJ, Richards LG, Stiers W, Zorowitz RD, American Heart Association Stroke Council, Council on Cardiovascular and Stroke Nursing, Council on Clinical Cardiology, and Council on Quality of Care and Outcomes Research. Guidelines for adult stroke rehabilitation and recovery: A guideline for healthcare professionals from the american heart association/american stroke association. *Stroke.* 2016;47(6):e98-e169.
8. Hornby TG, Reisman DS, Ward IG, Scheets PL, Miller A, Haddad D, Fox EJ, Fritz NE, Hawkins K, Henderson CE, Hendron KL, Holleran CL, Lynskey JE, Walter A, and the Locomotor CPG Appraisal Team. Clinical practice guideline to improve locomotor function following chronic stroke, incomplete spinal cord injury, and brain injury. *J Neurol Phys Ther.* 2020;44(1):49-100.
9. Boyne P, Dunning K, Carl D, Gerson M, Khoury J, Kissela B. High-intensity interval training in stroke rehabilitation. *Top Stroke Rehabil.* 2013;20(4):317-330.
10. Lam JM, Globas C, Cerny J, Hertler B, Uludag K, Forrester LW, Macko RF, Hanley DF, Becker C, Luft AR. Predictors of response to treadmill exercise in stroke survivors. *Neurorehabil Neural Repair.* 2010;24(6):567-574. PMID: PMC3024239.
11. Boyne P, Welge J, Kissela B, Dunning K. Factors influencing the efficacy of aerobic exercise for improving fitness and walking capacity after stroke: A meta-analysis with meta-regression. *Arch Phys Med Rehabil.* 2017;98(3):581-595. PMID: PMC5868957.
12. Hornby TG, Straube DS, Kinnaird CR, Holleran CL, Echaz AJ, Rodriguez KS, Wagner EJ, Narducci EA. Importance of specificity, amount, and intensity of locomotor training to improve ambulatory function in patients poststroke. *Top Stroke Rehabil.* 2011;18(4):293-307.
13. Outermans JC, van Peppen RP, Wittink H, Takken T, Kwakkel G. Effects of a high-intensity task-oriented training on gait performance early after stroke: A pilot study. *Clin Rehabil.* 2010;24(11):979-987.
14. Rose D, Paris T, Crews E, Wu SS, Sun A, Behrman AL, Duncan P. Feasibility and effectiveness of circuit training in acute stroke rehabilitation. *Neurorehabil Neural Repair.* 2011;25(2):140-148.
15. Pohl M, Mehrholz J, Ritschel C, Ruckriem S. Speed-dependent treadmill training in ambulatory hemiparetic stroke patients: A randomized controlled trial. *Stroke.* 2002;33(2):553-558.
16. Lau KW, Mak MK. Speed-dependent treadmill training is effective to improve gait and balance performance in patients with sub-acute stroke. *J Rehabil Med.* 2011;43(8):709-713.

17. Sullivan KJ, Knowlton BJ, Dobkin BH. Step training with body weight support: Effect of treadmill speed and practice paradigms on poststroke locomotor recovery. *Arch Phys Med Rehabil.* 2002;83(5):683-691.
18. Moore JL, Roth EJ, Killian C, Hornby TG. Locomotor training improves daily stepping activity and gait efficiency in individuals poststroke who have reached a "plateau" in recovery. *Stroke.* 2010;41(1):129-135.
19. Kwakkel G, Wagenaar RC, Twisk JW, Lankhorst GJ, Koetsier JC. Intensity of leg and arm training after primary middle-cerebral-artery stroke: A randomised trial. *Lancet.* 1999;354(9174):191-196.
20. Lisman J, Spruston N. Postsynaptic depolarization requirements for LTP and LTD: A critique of spike timing-dependent plasticity. *Nat Neurosci.* 2005;8(7):839-841.
21. Peinemann A, Reimer B, Loer C, Quartarone A, Munchau A, Conrad B, Siebner HR. Long-lasting increase in corticospinal excitability after 1800 pulses of subthreshold 5 hz repetitive TMS to the primary motor cortex. *Clin Neurophysiol.* 2004;115(7):1519-1526.
22. Ada L, Dean CM, Hall JM, Bampton J, Crompton S. A treadmill and overground walking program improves walking in persons residing in the community after stroke: A placebo-controlled, randomized trial. *Arch Phys Med Rehabil.* 2003;84(10):1486-1491.
23. Eich HJ, Mach H, Werner C, Hesse S. Aerobic treadmill plus bobath walking training improves walking in subacute stroke: A randomized controlled trial. *Clin Rehabil.* 2004;18(6):640-651.
24. Ivey FM, Hafer-Macko CE, Ryan AS, Macko RF. Impaired leg vasodilatory function after stroke: Adaptations with treadmill exercise training. *Stroke.* 2010;41(12):2913-2917.
25. Ivey FM, Ryan AS, Hafer-Macko CE, Macko RF. Improved cerebral vasomotor reactivity after exercise training in hemiparetic stroke survivors. *Stroke.* 2011;42(7):1994-2000.
26. Luft AR, Macko RF, Forrester LW, Villagra F, Ivey F, Sorkin JD, Whittall J, McCombe-Waller S, Katzel L, Goldberg AP, Hanley DF. Treadmill exercise activates subcortical neural networks and improves walking after stroke. A randomized controlled trial. *Stroke.* 2008;39(12):3341-50.
27. Kuys SS, Brauer SG, Ada L. Higher-intensity treadmill walking during rehabilitation after stroke is feasible and not detrimental to walking pattern or quality: A pilot randomized trial. *Clin Rehabil.* 2011;25(4):316-326.
28. Macko RF, Ivey FM, Forrester LW, Hanley D, Sorkin JD, Katzel LI, Silver KH, Goldberg AP. Treadmill exercise rehabilitation improves ambulatory function and cardiovascular fitness in patients with chronic stroke: A randomized, controlled trial. *Stroke.* 2005;36(10):2206-2211.
29. Olawale OA, Jaja SI, Anigbogu CN, Appiah-Kubi KO, Jones-Okai D. Exercise training improves walking function in an african group of stroke survivors: A randomized controlled trial. *Clin Rehabil.* 2011;25(5):442-450.
30. Lee J, Moon H, Lee S, Lee H, Park Y. The effects of a community-based walking program on walking ability and fall-related self-efficacy of chronic stroke patients. *Journal of exercise rehabilitation.* 2019;15(1):20-25.
31. Langhammer B, Stanghelle JK. Exercise on a treadmill or walking outdoors? A randomized controlled trial comparing effectiveness of two walking exercise programmes late after stroke. *Clin Rehabil.* 2010;24(1):46-54.
32. Klassen TD, Dukelow SP, Bayley MT, Benavente O, Hill MD, Krassioukov A, Liu-Ambrose T, Pooyania S, Poulin MJ, Schneeberg A, Yao J, Eng JJ. Higher doses improve walking recovery during stroke inpatient rehabilitation. *Stroke (1970).* 2020;51(9):2639-2648.
33. Ada L, Dean CM, Lindley R. Randomized trial of treadmill training to improve walking in community-dwelling people after stroke: The AMBULATE trial. *Int J Stroke.* 2013;8(6):436-444.
34. Billinger SA, Boyne P, Coughenour E, Dunning K, Matlaga A. Does aerobic exercise and the FITT principle fit into stroke recovery? *Curr Neurol Neurosci Rep.* 2015;15(2):519-8. PMID: PMC4560458.
35. Boyne P, Billinger S, MacKay-Lyons M, Barney B, Khoury J, Dunning K. Aerobic exercise prescription in stroke rehabilitation: A web-based survey of united states physical therapists. *J Neurol Phys Ther.* 2017;41(2):119-128. PMID: PMC5357184.

36. Reynolds H, Steinfors S, Tillyard J, Ellis S, Hayes A, Hanson ED, Wijeratne T, Skinner EH. Feasibility and adherence to moderate intensity cardiovascular fitness training following stroke: A pilot randomized controlled trial. *BMC Neurol*. 2021;21(1):132-8. PMID: PMC7983371.
37. Wiener J, McIntyre A, Janssen S, Chow JT, Batey C, Teasell R. Effectiveness of High-Intensity interval training for fitness and mobility post stroke: A systematic review. *PM & R*. 2019;11(8):868-878.
38. Crozier J, Roig M, Eng JJ, MacKay-Lyons M, Fung J, Ploughman M, Bailey DM, Sweet SN, Giacomantonio N, Thiel A, Trivino M, Tang A. High-intensity interval training after stroke: An opportunity to promote functional recovery, cardiovascular health, and neuroplasticity. *Neurorehabil Neural Repair*. 2018;32(6-7):543-556.
39. Dupont G, Blondel N, Lensele G, Berthoin S. Critical velocity and time spent at a high level of VO₂ for short intermittent runs at supramaximal velocities. *Can J Appl Physiol*. 2002;27(2):103-115.
40. Guiraud T, Nigam A, Juneau M, Meyer P, Gayda M, Bosquet L. Acute responses to high-intensity intermittent exercise in CHD patients. *Med Sci Sports Exerc*. 2011;43(2):211-217.
41. Askim T, Dahl AE, Aamot IL, Hokstad A, Helbostad J, Indredavik B. High-intensity aerobic interval training for patients 3-9 months after stroke. A feasibility study. *Physiother Res Int*. 2014;19(3):129-39.
42. Gjellesvik TI, Brurok B, Hoff J, Torhaug T, Helgerud J. Effect of high aerobic intensity interval treadmill walking in people with chronic stroke: A pilot study with one year follow-up. *Top Stroke Rehabil*. 2012;19(4):353-360.
43. Boyne P, Dunning K, Carl D, Gerson M, Khoury J, Rockwell B, Keeton G, Westover J, Williams A, McCarthy M, Kissela B. High-intensity interval training and moderate-intensity continuous training in ambulatory chronic stroke: A feasibility study. *Phys Ther*. 2016;96(10):1533-1544. PMID: PMC5046191.
44. Munari D, Pedrinolla A, Smania N, Picelli A, Gandolfi M, Saltuari L, Schena F. High-intensity treadmill training improves gait ability, VO₂peak and cost of walking in stroke survivors: Preliminary results of a pilot randomized controlled trial. *Eur J Phys Rehabil Med*. 2018;54(3):408-418.
45. Madhavan S, Cleland BT, Sivaramakrishnan A, Freels S, Lim H, Testai FD, Corcos DM. Cortical priming strategies for gait training after stroke: A controlled, stratified trial. *J NeuroEngineering Rehabil*. 2020;17(1):111.
46. Gjellesvik TI, Becker F, Tjønnå AE, Indredavik B, Lundgaard E, Solbakken H, Brurok B, Tørhaug T, Lydersen S, Askim T. Effects of high-intensity interval training after stroke (the HIIT stroke study) on physical and cognitive function: A multicenter randomized controlled trial. *Arch Phys Med Rehabil*. 2021;102(9):1683-1691.
47. Bosch PR, Holzapfel SD, Nordin K, Ojameruaye O, Zubriski M, Angadi SS. High-intensity interval training for adults with chronic stroke: A pilot feasibility study. *Cardiopulmonary physical therapy journal*. 2021;32(1):20-29.
48. Reisman D, Kesar T, Perumal R, Roos M, Rudolph K, Higginson J, Helm E, Binder-Macleod S. Time course of functional and biomechanical improvements during a gait training intervention in persons with chronic stroke. *J Neurol Phys Ther*. 2013;37(4):159-165. PMID: PMC3890376.
49. Miller A, Reisman DS, Billinger SA, Dunning K, Doren S, Ward J, Wright H, Wagner E, Carl D, Gerson M, Awosika O, Khoury J, Kissela B, Boyne P. Moderate-intensity exercise versus high-intensity interval training to recover walking post-stroke: Protocol for a randomized controlled trial. *Trials*. 2021;22(1):457.
50. Fugl-Meyer AR, Jaasko L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient. I. A method for evaluation of physical performance. *Scand J Rehab Med*. 1975;7:13-31.
51. Williams LS, Brizendine EJ, Plue L, Bakas T, Tu W, Hendrie H, Kroenke K. Performance of the PHQ-9 as a screening tool for depression after stroke. *Stroke*. 2005;36(3):635-638.
52. Macko RF, Ivey FM, Forrester LW. Task-oriented aerobic exercise in chronic hemiparetic stroke: training protocols and treatment effects. *Topics in Stroke Rehabilitation*. 2005;12(1):45-57.
53. Tilson JK, Sullivan KJ, Cen SY, Rose DK, Koradia CH, Azen SP, Duncan PW, Locomotor Experience Applied Post Stroke (LEAPS) Investigative Team. Meaningful gait speed improvement during the first 60

- days poststroke: Minimal clinically important difference. *Phys Ther.* 2010;90(2):196-208. PMID: PMC2816032.
54. Macko RF, Katzel LI, Yataco A, Tretter LD, DeSouza CA, Dengel DR, Smith GV, Silver KH. Low-velocity graded treadmill stress testing in hemiparetic stroke patients. *Stroke.* 1997;28(5):988-992.
55. Bohannon RW, Williams Andrews A. Normal walking speed: A descriptive meta-analysis. *Physiotherapy.* 2011;97(3):182-189.
56. Ivey FM, Hafer-Macko CE, Macko RF. Task-oriented treadmill exercise training in chronic hemiparetic stroke. *J Rehabil Res Dev.* 2008;45(2):249-259. PMID: PMC2990953.
57. American College of Sports Medicine, (ACSM). *ACSM's guidelines for exercise testing and prescription.* 9th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2014.
58. Brott T, Adams HP, Jr, Olinger CP, Marler JR, Barsan WG, Biller J, Spilker J, Holleran R, Eberle R, Hertzberg V. Measurements of acute cerebral infarction: A clinical examination scale. *Stroke.* 1989;20(7):864-870.
59. Ashworth B. Preliminary trial of carisoprodol in multiple sclerosis. *Practitioner.* 1964;192:540-542.
60. Duncan PW, Sullivan KJ, Behrman AL, Azen SP, Wu SS, Nadeau SE, Dobkin BH, Rose DK, Tilson JK, LEAPS Investigative Team. Protocol for the locomotor experience applied post-stroke (LEAPS) trial: A randomized controlled trial. *BMC Neurol.* 2007;7:39. PMID: PMC2222229.
61. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform.* 2009;42(2):377-381. PMID: PMC2700030.
62. Dean CM, Ada L, Lindley RI. Treadmill training provides greater benefit to the subgroup of community-dwelling people after stroke who walk faster than 0.4m/s: A randomised trial. *J Physiother.* 2014;60(2):97-101.
63. Boyne P, Doren S, Scholl V, Staggs E, Whitesel D, Carl D, Shatz R, Sawyer R, Awosika OO, Reisman DS, Billinger SA, Kissela B, Vannest J, Dunning K. Preliminary outcomes of combined treadmill and overground high-intensity interval training in ambulatory chronic stroke. *Frontiers in Neurology.* 2022;13:812875.
64. Boyne P, Scholl V, Doren S, Carl D, Billinger SA, Reisman DS, Gerson M, Kissela B, Vannest J, Dunning K. Locomotor training intensity after stroke: Effects of interval type and mode. *Top Stroke Rehabil.* 2020;27(7):483-493. PMID: PMC7429314.
65. Boyne P, Dunning K, Carl D, Gerson M, Khoury J, Kissela B. Within-session responses to high-intensity interval training in chronic stroke. *Med Sci Sports Exerc.* 2015;47(3):476-484.
66. Macko RF, Haeuber E, Shaughnessy M, Coleman KL, Boone DA, Smith GV, Silver KH. Microprocessor-based ambulatory activity monitoring in stroke patients. *Med Sci Sports Exerc.* 2002;34(3):394-399.
67. Fletcher GF, Ades PA, Kligfield P, Arena R, Balady GJ, Bittner VA, Coke LA, Fleg JL, Forman DE, Gerber TC, Gulati M, Madan K, Rhodes J, Thompson PD, Williams MA, American Heart Association Exercise, Cardiac Rehabilitation, and Prevention Committee of the Council on Clinical Cardiology, Council on Nutrition, Physical Activity and Metabolism, Council on Cardiovascular and Stroke Nursing, and Council. Exercise standards for testing and training: A scientific statement from the American Heart Association. *Circulation.* 2013;128(8):873-934.
68. National Cancer Institute. Common Terminology Criteria for Adverse Events v4.0. *NIH publication # 09-7473.* 2009.
69. Kluding PM, Pasnoor M, Singh R, D'Silva LJ, Yoo M, Billinger SA, LeMaster JW, Dimachkie MM, Herbelin L, Wright DE. Safety of aerobic exercise in people with diabetic peripheral neuropathy: Single-group clinical trial. *Phys Ther.* 2015;95(2):223-234.
70. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: Guidelines for the six-minute walk test. *Am J Respir Crit Care Med.* 2002;166(1):111-117.
71. Mudge S, Stott NS. Timed walking tests correlate with daily step activity in persons with stroke. *Arch Phys Med Rehabil.* 2009;90(2):296-301.

72. Fulk GD, Reynolds C, Mondal S, Deutsch JE. Predicting home and community walking activity in people with stroke. *Arch Phys Med Rehabil.* 2010;91(10):1582-1586.
73. Donovan K, Lord SE, McNaughton HK, Weatherall M. Mobility beyond the clinic: The effect of environment on gait and its measurement in community-ambulant stroke survivors. *Clin Rehabil.* 2008;22(6):556-563.
74. Miller AE, Russell E, Reisman DS, Kim HE, Dinh V. A machine learning approach to identifying important features for achieving step thresholds in individuals with chronic stroke. *PLoS One.* 2022;17(6):e0270105.
75. Muren MA, Hutler M, Hooper J. Functional capacity and health-related quality of life in individuals post stroke. *Top Stroke Rehabil.* 2008;15(1):51-58.
76. Rich S, MD. The 6-minute walk test as a primary endpoint in clinical trials for pulmonary hypertension. *J Am Coll Cardiol.* 2012;60(13):1202-1203.
77. Muenzer J, Wraith JE, Beck M, Giugliani R, Harmatz P, Eng CM, Vellodi A, Martin R, Ramaswami U, Gucevas-Calikoglu M, Vijayaraghavan S, Wendt S, Puga A, Ulbrich B, Shinawi M, Cleary M, Piper D, Conway AM, Kimura A. A phase II/III clinical study of enzyme replacement therapy with idursulfase in mucopolysaccharidosis II (hunter syndrome). *Genetics in medicine.* 2006;8(8):465-473.
78. Wraith JE, Clarke LA, Beck M, Kolodny EH, Pastores GM, Muenzer J, Rapoport DM, Berger KI, Swiedler SJ, Kakkis ED, Braakman T, Chadbourne E, Walton-Bowen K, Cox GF. Enzyme replacement therapy for mucopolysaccharidosis I: A randomized, double-blinded, placebo-controlled, multinational study of recombinant human α -L-iduronidase (laronidase). *J Pediatr.* 2004;144(5):581-588.
79. van der Ploeg, Ans T., Clemens PR, Corzo D, Escolar DM, Florence J, Groeneveld GJ, Herson S, Kishnani PS, Laforet P, Lake SL, Lange DJ, Leshner RT, Mayhew JE, Morgan C, Nozaki K, Park DJ, Pestronk A, Rosenbloom B, Skrinar A, van Capelle CI, van der Beek, Nadine A., Wasserstein M, Zivkovic SA. A randomized study of alglucosidase alfa in late-onset pompe's disease. *N Engl J Med.* 2010;362(15):1396-1406.
80. Raval AN, Cook TD, Duckers HJ, Johnston PV, Traverse JH, Abraham WT, Altman PA, Pepine CJ. The CardiAMP heart failure trial: A randomized controlled pivotal trial of high-dose autologous bone marrow mononuclear cells using the CardiAMP cell therapy system in patients with post-myocardial infarction heart failure: Trial rationale and study design. *The American heart journal.* 2018;201:141-148.
81. Fulk GD, Echternach JL, Nof L, O'Sullivan S. Clinometric properties of the six-minute walk test in individuals undergoing rehabilitation poststroke. *Physiother Theory Pract.* 2008;24(3):195-204.
82. Eng JJ, Dawson AS, Chu KS. Submaximal exercise in persons with stroke: Test-retest reliability and concurrent validity with maximal oxygen consumption. *Arch Phys Med Rehabil.* 2004;85(1):113-118. PMID: PMC3167868.
83. Wevers LE, Kwakkel G, van de Port, I G. Is outdoor use of the six-minute walk test with a global positioning system in stroke patients' own neighbourhoods reproducible and valid? *J Rehabil Med.* 2011;43(11):1027-1031.
84. Flansbjerg U, Holmbäck AM, Downham D, Patten C, Lexell J. Reliability of gait performance tests in men and women with hemiparesis after stroke. *J Rehabil Med.* 2005;37(2):75-82.
85. Kosak M, Smith T. Comparison of the 2-, 6-, and 12-minute walk tests in patients with stroke. *J Rehabil Res Dev.* 2005;42(1):103-107.
86. Perera S, Mody SH, Woodman RC, Studenski SA. Meaningful change and responsiveness in common physical performance measures in older adults. *J Am Geriatr Soc.* 2006;54(5):743-749.
87. Cella D, Lai J, Jensen SE, Christodoulou C, Junghaenel DU, Reeve BB, Stone AA. PROMIS fatigue item bank had clinical validity across diverse chronic conditions. *J Clin Epidemiol.* 2016;73:128-134.
88. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, Amtmann D, Bode R, Buysse D, Choi S, Cook K, DeVellis R, DeWalt D, Fries JF, Gershon R, Hahn EA, Lai J, Pilkonis P, Revicki D, Rose M, Weinfurt K, Hays R, PROMIS Cooperative Group, PROMIS Cooperative Grp. The patient-reported outcomes measurement information system (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *J Clin Epidemiol.* 2010;63(11):1179-1194.

89. Boyne P, Reisman D, Brian M, Barney B, Franke A, Carl D, Khoury J, Dunning K. Ventilatory threshold may be a more specific measure of aerobic capacity than peak oxygen consumption rate in persons with stroke. *Top Stroke Rehabil.* 2017;24(2):149-157. PMID: PMC5588902.
90. Marzolini S, Oh P, McIlroy W, Brooks D. The feasibility of cardiopulmonary exercise testing for prescribing exercise to people after stroke. *Stroke.* 2012;43(4):1075-1081.
91. American Thoracic Society, American College of Chest Physicians. ATS/ACCP statement on cardiopulmonary exercise testing. *Am J Respir Crit Care Med.* 2003;167(2):211-277.
92. Mezzani A, Hamm LF, Jones AM, McBride PE, Moholdt T, Stone JA, Urhausen A, Williams MA, European Association for Cardiovascular Prevention and Rehabilitation, American Association of Cardiovascular and Pulmonary Rehabilitation, Canadian Association of Cardiac Rehabilitation. Aerobic exercise intensity assessment and prescription in cardiac rehabilitation: A joint position statement of the european association for cardiovascular prevention and rehabilitation, the american association of cardiovascular and pulmonary rehabilitation, and the canadian association of cardiac rehabilitation. *J Cardiopulm Rehabil Prev.* 2012;32(6):327-350.
93. Marzolini S, Oh P, McIlroy W, Brooks D. The effects of an aerobic and resistance exercise training program on cognition following stroke. *Neurorehabil Neural Repair.* 2013;27(5):392-402.
94. Holleran CL, Rodriguez KS, Echaz A, Leech KA, Hornby TG. Potential contributions of training intensity on locomotor performance in individuals with chronic stroke. *J Neurol Phys Ther.* 2015;39(2):95-102.
95. Kim KJ, Rivas E, Prejean B, Frisco D, Young M, Downs M. Novel computerized method for automated determination of ventilatory threshold and respiratory compensation point. *Frontiers in physiology.* 2021;12:782167.
96. Ekkekakis P, Lind E, Hall EE, Petruzzello SJ. Do regression-based computer algorithms for determining the ventilatory threshold agree? *J Sports Sci.* 2008;26(9):967-976.
97. Gaskill SE, Ruby BC, Walker AJ, Sanchez OA, Serfass RC, Leon AS. Validity and reliability of combining three methods to determine ventilatory threshold. *Med Sci Sports Exerc.* 2001;33(11):1841-1848.
98. R Development Core Team. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria.* 2004.
99. Lindeløv JK. Mcp: An R package for regression with multiple change points. *OSF Preprints [preprint].* 2020:<https://doi-org.uc.idm.oclc.org/10.31219/osf.io/fzqxv>.
100. Brawner CA, Ehrman JK, Schairer JR, Cao JJ, Keteyian SJ. Predicting maximum heart rate among patients with coronary heart disease receiving beta-adrenergic blockade therapy. *Am Heart J.* 2004;148(5):910-914.
101. Gellish RL, Goslin BR, Olson RE, McDonald A, Russi GD, Moudgil VK. Longitudinal modeling of the relationship between age and maximal heart rate. *Med Sci Sports Exerc.* 2007;39(5):822-829.
102. Tyrell CM, Roos MA, Rudolph KS, Reisman DS. Influence of systematic increases in treadmill walking speed on gait kinematics after stroke. *Phys Ther.* 2011;91(3):392-403. PMID: PMC3048817.
103. Leddy AL, Connolly M, Holleran CL, Hennessy PW, Woodward J, Arena RA, Roth EJ, Hornby TG. Alterations in aerobic exercise performance and gait economy following high-intensity dynamic stepping training in persons with subacute stroke. *J Neurol Phys Ther.* 2016;40(4):239-248. PMID: PMC5026397.
104. Mehrholz J, Wagner K, Rutte K, Meissner D, Pohl M. Predictive validity and responsiveness of the functional ambulation category in hemiparetic patients after stroke. *Arch Phys Med Rehabil.* 2007;88(10):1314-1319.
105. Powell LE, Myers AM. The activities-specific balance confidence (ABC) scale. *J Gerontol A Biol Sci Med Sci.* 1995;50A(1):28.
106. Botner EM, Miller WC, Eng JJ. Measurement properties of the activities-specific balance confidence scale among individuals with stroke. *Disabil Rehabil.* 2005;27(4):156-163.
107. Golicki D, Niewada M, Buczek J, Karlinska A, Kobayashi A, Janssen MF, Pickard AS. Validity of EQ-5D-5L in stroke. *Quality of life research.* 2014;24(4):845-850.

108. Pickard AS, PhD, Law, Ernest H., PharmD, PhD, Jiang R, PharmD, Pullenayegum E, PhD, Shaw, James W., PharmD, MPH, PhD, Xie F, PhD, Oppe M, PhD, Boye KS, PhD, Chapman RH, PhD, Gong, Cynthia L., PharmD, PhD, Balch A, PhD, Busschbach JJV, PhD. United states valuation of EQ-5D-5L health states using an international protocol. *Value in health*. 2019;22(8):931-941.
109. Enright PL, Sherrill DL. Reference equations for the six-minute walk in healthy adults. *Am J Respir Crit Care Med*. 1998;158(5 Pt 1):1384-1387.
110. Guazzi M, Adams V, Conraads V, Halle M, Mezzani A, Vanhees L, Arena R, Fletcher GF, Forman DE, Kitzman DW, Lavie CJ, Myers J, European Association for Cardiovascular Prevention & Rehabilitation, American Heart Association. EACPR/AHA scientific statement. clinical recommendations for cardiopulmonary exercise testing data assessment in specific patient populations. *Circulation*. 2012;126(18):2261-2274.
111. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Statist Soc B*. 1995;57(1):289-300.
112. Fitzmaurice GM, Laird NM, Ware JH. *Applied longitudinal analysis*. 2nd ed. Hoboken, NJ: John Wiley & Sons, Inc.; 2011.
113. Higgins J, Savović J, Page MJ, Elbers RG, Sterne J. Chapter 8: Assessing risk of bias in a randomized trial. In: Higgins J, Thomas J, Chandler J, et al, eds. *Cochrane handbook for systematic reviews of interventions version 6.3*. www.training.cochrane.org/handbook; Cochrane; 2022.
114. Guo Y, Logan HL, Glueck DH, Muller KE. Selecting a sample size for studies with repeated measures. *BMC Med Res Methodol*. 2013;13:100-100. PMID: PMC3734029.
115. Ivey FM, Stookey AD, Hafer-Macko CE, Ryan AS, Macko RF. Higher treadmill training intensity to address functional aerobic impairment after stroke. *J Stroke Cerebrovasc Dis*. 2015;24(11):2539-2546.
116. Duncan P, Studenski S, Richards L, Gollub S, Lai SM, Reker D, Perera S, Yates J, Koch V, Rigler S, Johnson D. Randomized clinical trial of therapeutic exercise in subacute stroke. *Stroke*. 2003;34(9):2173-2180.
117. Kwakkel G. Impact of intensity of practice after stroke: Issues for consideration. *Disabil Rehabil*. 2006;28(13-14):823-830.
118. Duncan PW, Sullivan KJ, Behrman AL, Azen SP, Wu SS, Nadeau SE, Dobkin BH, Rose DK, Tilson JK, Cen S, Hayden SK, LEAPS Investigative Team. Body-weight-supported treadmill rehabilitation after stroke. *N Engl J Med*. 2011;364(21):2026-2036. PMID: PMC3175688.
119. Tiedemann A, Sherrington C, Dean CM, Rissel C, Lord SR, Kirkham C, O'Rourke SD. Predictors of adherence to a structured exercise program and physical activity participation in community dwellers after stroke. *Stroke Res Treat*. 2012;2012:136525. PMID: PMC3191741.
120. Jurkiewicz MT, Marzolini S, Oh P. Adherence to a home-based exercise program for individuals after stroke. *Top Stroke Rehabil*. 2011;18(3):277-284.
121. Saunders DH, Greig CA, Mead GE. Physical activity and exercise after stroke: Review of multiple meaningful benefits. *Stroke*. 2014;45(12):3742-3747.
122. Kuppaswamy A, Clark EV, Turner IF, Rothwell JC, Ward NS. Post-stroke fatigue: A deficit in corticomotor excitability? *Brain (London, England : 1878)*. 2014;138(Pt 1):136-148.
123. Boyne P, Meyrose C, Westover J, Whitesel D, Hatter K, Reisman DS, Cunningham D, Carl D, Jansen C, Khoury JC, Gerson M, Kissela B, Dunning K. Exercise intensity affects acute neurotrophic and neurophysiologic responses post-stroke. *Journal of Applied Physiology*. 2019;126(2):431-433. PMID: PMC6397406.
124. Hsu C, Fu T, Huang S, Chen CP, Wang J. Increased serum brain-derived neurotrophic factor with high-intensity interval training in stroke patients: A randomized controlled trial. *Annals of physical and rehabilitation medicine*. 2020;64(4):101385.

125. Steen Krawczyk R, Vinther A, Petersen NC, Faber J, Iversen HK, Christensen T, Lambertsen KL, Rehman S, Klausen TW, Rostrup E, Kruuse C. Effect of home-based high-intensity interval training in patients with lacunar stroke: A randomized controlled trial. *Front Neurol.* 2019;10:664. PMID: PMC6611174.
126. Dobkin BH. Progressive staging of pilot studies to improve phase III trials for motor interventions. *Neurorehabil Neural Repair.* 2009;23(3):197-206.