Critical care cycling to improve lower extremity strength (CYCLE) RCT:

An international, multi-center, randomized clinical trial of early in-bed cycling for mechanically ventilated patients

Statistical Analysis Plan

Version 1.0

January 9, 2024

Statistical Analysis Plan Approval Sheet

Study: CYCLE

Title: CYCLE RCT: An international, multi-center, randomized clinical trial of early in-bed cycling for mechanically ventilated patients

Trial registration numbers: NCT03471247 (Full RCT); NCT02377830 (46 patient internal pilot)

Principal Authors of Analysis Plan: Michelle Kho, Diane Heels-Ansdell
Version: 1.0
Version date: January 9, 2024

SAP contributors

Name	Affiliation	Role in SAP development
Michelle Kho	Associate Professor, McMaster University	Principal Investigator
Diane Heels-Ansdell	Assistant Professor, McMaster University	Trial Statistician
Deborah Cook	Professor, McMaster University	Senior Trialist
Lehana Thabane	Professor, McMaster University	Senior Biostatistician

The undersigned have reviewed this plan and find it to be consistent with the requirements of the protocol as it applies to their respective areas.

[redacted]
Author: Associate Professor Michelle Kho, PT, PhD
McMaster University Principal Investigator
January 23, 2024
Date

[redacted]
Author: Assistant Professor Diane Heels-Ansdell, MSc
McMaster University, Trial Statistician
January 23, 2024
Date

[redacted]
Author: Professor Deborah Cook, MD, MSc
McMaster University, Senior Trialist
Jan 23/24
Date

[redacted]
Author: Professor Lehana Thabane, PhD
McMaster University, Study Biostatistician
January 23, 2024
Date

Table of Contents

1.0 SAP Revision History

Protocol Version	Updated SAP version number	Section number changed	Description of and reason for change	Date changed	Timing of revisions relative to interim analysis
4	Not applicable				Post

2.0 INTRODUCTION

2.1 Background and rationale

Survivors of critical illness are at risk of developing physical dysfunction lasting from 5 to 8 years post-intensive care unit (ICU) discharge.[1, 2] Physical rehabilitation started in the ICU can improve patients' functional outcomes at hospital discharge.[3, 4] In a randomized trial, critically ill patients randomized to in-bed cycling started 2 weeks after ICU admission had farther 6-minute walk distance at hospital discharge compared to those receiving routine physiotherapy alone.[3] While in-bed cycling started earlier in a patient's ICU stay is safe [5-7] and feasible [8], the efficacy of early cycling on patients' function is unknown. A randomized trial comparing early in-bed cycling in addition to routine physiotherapy versus routine physiotherapy alone was therefore needed. We report this document according to the guidelines for the content of statistical analysis plans in clinical trials.[9]

2.2 Objectives

The primary objective of this trial is to determine the efficacy of early in-bed cycling (started within 4 days of mechanical ventilation initiation) and routine physiotherapy versus routine physiotherapy alone on patients' physical function 3 days post-ICU discharge.

3.0 STUDY METHODS

3.1 Trial design

CYCLE (Critical care cycling to improve lower extremity strength) is a 360-patient, international, multi-center, open-label, parallel group randomized trial (1:1 ratio) with blinded primary outcome assessment at 3 days post ICU discharge. Assessors were blinded to treatment group allocation.

3.2 Randomization

Randomization occurred after informed consent was obtained. We concealed allocation and used a central randomization process. We used a web-based, secure randomization service (http://www.randomize.net/)**.** After informed consent, the site research coordinator logged in to the website, registered the patient, and received the randomized assignment, ensuring allocation concealment. Patients were stratified by center and age (≥ 65 vs. < 65 years old).

Interventions and comparator

Intervention (Cycling): Patients were randomized to receive 30 minutes/ day of in-bed cycling in addition to routine physiotherapy interventions, 5 days per week, during their ICU stay. Cycling occurred for a maximum of 28 days or stopped when the patient could march on the spot for 2 consecutive days, whichever occurred first.

Comparator (Routine): Patients received routine physiotherapy interventions per current institutional practice. Routine PT included, based on the patient's alertness and medical stability, activities to maintain or increase limb range of motion and strength, in- and out-of-bed mobility, ambulation, and assistance with optimizing airway clearance and respiratory function.[4, 10-12]

3.3 Sample size

Sample size and power

Our sample size of 360 patients was based on identifying a 1.0 point mean difference[13] between the Cycling and Routine groups for the Physical Function Test for ICU-scored (PFIT-s) measured at 3 days after ICU discharge.[14, 15] Psychometric studies of the PFIT-s identified the minimal clinically important difference was 1.0 points.[14, 16] Logistic regression analysis of patients enrolled in TryCYCLE [7] and the CYCLE pilot randomized study [8] identified that each 1.0 point increase in PFIT-s at ICU discharge (representing better function) was associated with a 40% reduction in the composite outcome of death, readmission to ICU, or requiring paid assistance for activities of daily living at hospital discharge.[13] Based on a standard deviation of 2.5 points at ICU discharge,[7, 17] a 1.0 point difference between groups,[13, 14, 16] and 90% power (0.05 alpha), we need to randomize and analyze 266 patients (133 per group). Based on 66 patients enrolled in the CYCLE Pilot RCT, we anticipated approximately 35% total attrition (25% ICU mortality, 1% mortality in the first 3 days post-ICU discharge, 5% missed primary outcome assessments at 3 days post-ICU, and 5% unblinded). Therefore, we will recruit 360 patients overall.

3.4 Framework

This trial was based on a superiority trial hypothesis that patients receiving in-bed cycling and routine physiotherapy early in their ICU stay will have better physical function at 3 days post-ICU discharge than those receiving routine physiotherapy alone.

3.5 Statistical interim analysis and stopping guidance

We conducted one blinded interim analysis that included the first 180 patients enrolled (half of our sample size) to assess for benefit and harm (serious adverse events). We used conservative statistical guidelines for data monitoring based on the modified Haybittle-Peto rule.[18] To maintain the overall type-I error rate (i.e., α), we evaluated the primary endpoint using a fixed simple conservative $\alpha=0.001$ for the interim analyses and $\alpha=0.05$ for the final analysis. The Data Monitoring Committee (DMC) recommended continuation of the trial on September 29, 2020.

3.6 Timing of final analysis

The first publication of the trial results will be prepared for the Cycling vs. Routine groups when every patient has reached 90 days post-randomization, and data for vital status at hospital discharge have been received. Longer-term endpoints for the economic evaluation will be reported in a separate publication.

In this document, we will outline only the analyses that will be included in the primary CYCLE manuscript.

3.7 Timing of outcome assessments

Supplement Table 1 outlines the schedule of study procedures with five timepoints for outcome assessments. The ICU Awakening timepoint was based on the physiotherapist's assessment of the patient's ability to consistently follow 5 verbal commands[19]. The ICU discharge timepoint occurred when the patient was discharged from the ICU or when a discharge order was written for the patient, whichever occurred first. The 3-day post-ICU timepoint occurred 3 days following the patient's physical discharge from the ICU. The hospital discharge timepoint occurred when a discharge order was written for the patient for the index admission (including alternative level of care). The 90-day timepoint occurred at 90 days following randomization.

4.0 STATISTICAL PRINCIPLES

4.1 Confidence intervals and P values

All statistical tests will be 2-sided and will be performed using a 5% significance level. We will report the two-sided 95% confidence intervals.

4.2 Adherence and protocol deviations

4.2.1 Adherence

Definitions:

"*Study days*" included all days in ICU from day of randomization up to 28 days post-randomization.

We did not plan for the randomized intervention to occur in the following circumstances:

- On days when a patient was randomized after normal physiotherapist working hours
- On days when a patient was transferred out of ICU before 12:00 pm
- On weekend days or statutory holidays
- For those randomized to in-bed cycling, patients who had marched on the spot for 2 consecutive days (and continued marching or higher mobility for the remainder of their ICU stay)

The remaining days were "*Planned intervention days*". On weekdays (i.e., Monday through Friday), physiotherapists reviewed study patients for one or more of the following "Temporary exemptions" before offering the randomized intervention:

1. Increase in vasopressor/ inotrope within last 2 hours
2. Active myocardial ischemia, or unstable/ uncontrolled arrhythmia per ICU team
3. Mean arterial pressure <60 or >110 mmHg or per treating team within the last 2 hours
4. Heart Rate <40 or >140 bpm within the last 2 hours
5. Persistent SpO_2 <88% or per treating team within the last 2 hours

6. Neuromuscular blocker within last 4 hours
7. Severe agitation (Richmond Agitation and Sedation Scale >2 [or equivalent][20]) within last 2 hours
8. Uncontrolled pain
9. Change in goals to palliative care
10. Team perception that in-bed cycling or therapy is not appropriate for other new reasons (e.g., acute peritonitis, new incision/wound, known/suspected rhabdomyolysis)

If the patient had no "*Temporary exemptions*", we offered the randomized intervention.

Each "Planned intervention day" without a "Temporary exemption" was an "*Eligible day*". An eligible day where a patient did not receive the randomized intervention, was a "*Missed opportunity*". Missed opportunities may have occurred due to:

1. Patient factors (e.g., patient not available due to a test, or declined)
2. Therapist factors (e.g., therapist not available due to vacation or sickness)
3. Equipment factors for patients in the Cycling arm (e.g., cycle ergometer malfunction)

Therefore, we define percent adherence whereby the numerator includes days in which patients received the randomized intervention or had a Temporary exemption and the denominator includes all planned intervention days (patients received the randomized intervention, Temporary exemptions, and Missed opportunities). We will report descriptive statistics on the percent protocol fidelity for the cohort by randomization group.

4.2.2 Major protocol deviation

If a patient who was randomized to routine physiotherapy alone received cycling, this was considered a major protocol deviation.

4.3 Analysis populations

We will include all eligible randomized patients (i.e., excluding post-randomization exclusions representing non-eligible patients), according to the treatment they were randomized to receive. The analyses of our primary outcome will only include patients who survived to 3 days post-ICU discharge, as specified in our original protocol and sample size calculation.[21] The analysis of the PFIT-s at other timepoints (ICU awakening, ICU discharge, and hospital discharge) and all performance-based (strength and function) and patient-reported (e.g., quality-of-life) outcomes will only include patients who survive to the given timepoint. The analysis of hospital discharge location will only include those discharged from the hospital alive. The analysis of ICU and hospital mortality and length of stay, as well as duration of mechanical ventilation, will include all enrolled patients.

5.0 TRIAL POPULATION

5.1 Screening data

We will report the total number of patients screened (i.e., meeting all inclusion criteria), and those with exclusion criteria based on screening logs from participating sites. For eligible patients, we will report reasons for non-enrollment.

5.2 Eligibility

Table 1 outlines trial inclusion and exclusion criteria.[21]

5.3 Recruitment

We will report the CONSORT diagram for all participants throughout the study (Figure 1).

5.4 Withdrawal/ follow-up

We will document patient withdrawals and losses to follow-up in our CONSORT diagram.

5.5 Baseline patient characteristics

We will summarize categorical data as counts and percentages. We will summarize continuous data as means, standard deviations, or median and interquartile range, if data are non-normally distributed. We will not conduct tests of statistical significance between randomized groups; rather, we will note the clinical importance of any imbalance between groups.

5.6 Consent and randomization

Informed consent and randomization procedures have been previously described in our protocol [21]. Briefly, research coordinators screened the participating ICUs for potentially eligible patients, identified eligible patients, and sought *a priori* consent from the patient, their substitute decision-maker or legally authorized representative.

6.0 ANALYSIS

6.1 Outcome definitions

Tables 2 and 3 describe our primary and secondary outcomes, measurement time points, and analysis methods.

Outcomes

Primary Outcome

The primary outcome is the PFIT-s measured at 3 days post-ICU discharge.[14, 15] The PFIT-s includes 4 items (arm strength, leg strength, ability to stand, and step cadence), each scored from 0 to 3, summed to a maximum of 12 points, and transformed to a total score of 10 (Table 4).[14] Higher scores represent better function. It was developed in an ICU population, includes functional items, and, unlike the 6-minute walk test (6MWT), can be measured serially over time (as few patients can walk at ICU awakening).[22] We chose the PFIT-s because we expect all

ICU patients will be able to complete part of the assessment even if they cannot stand (e.g., arm or leg strength), limiting floor effects. The PFIT-s is reliable and valid, with strong psychometric properties (reliability range = 0.996 to 1.00 [15]; convergent validity with the 6MWT and muscle strength [14]). We selected 3 days post-ICU discharge because it is proximal to the intervention and prior studies documented variable delivery of rehabilitation post-ICU[23] that may contaminate later evaluations.

Secondary Outcomes
Secondary outcomes include performance-based, patient-reported, and those collected by chart review. Performance-based measures included muscle strength (Medical Research Council Sum Score)[24, 25] and function (30-second sit to stand,[26, 27] 2-minute walk test).[28] The 30-second sit-to-stand and 2-minute walk test have age- and sex- based reference values, and good reliability in critically ill or frail elderly populations.[27, 28] Patient-reported measures included the Patient-Reported Functional Scale for the ICU (PRFS-ICU),[29, 30] critical care-related psychological distress using the Intensive Care Psychological Assessment Tool (IPAT),[31, 32] health-related quality-of-life using the EuroQuol (EQ-5D-5L™),[33-35] and the Hospital Anxiety and Depression Scale (HADS).[36] We also collected frailty (Clinical Frailty Scale),[37] Katz activities of daily living (ADL) scale,[38] duration of mechanical ventilation, length of stay (ICU and hospital), mortality (ICU, hospital, 90-day post-randomization), and change in living location at hospital discharge from baseline. Due to funding limitations, we restricted 90-day post-randomization outcomes to those enrolled after March 7, 2018.

Adverse Events
We will report the following adverse events if they occurred during or immediately after in-bed cycling or routine physiotherapy interventions, if they were attributable by the clinical team to the randomized intervention, and resulted in a clinical deterioration of the patient's status [3, 7, 39-41]:

Severe adverse events: unplanned extubation, cardiac arrest, or fall to knees during routine PT/ rehabilitation activities.

Serious adverse events: Concern for myocardial ischemia or suspected new unstable/ uncontrolled arrhythmia; Sustained symptomatic bradycardia (<40 bpm) or tachycardia (>140 bpm) and clinical deterioration attributed to in-bed cycling or routine PT/ rehabilitation activities; Sustained hypertension (mean arterial pressure >120 mmHg) and clinical deterioration attributed to in-bed cycling or routine PT/ rehabilitation activities; Sustained O_2 desaturation below baseline and clinical deterioration attributed to in-bed cycling or routine PT/ rehabilitation activities; Marked ventilator dysynchrony; Bleeding at femoral catheter site attributed to in-bed cycling or routine PT/ rehabilitation activities; New bruising at femoral catheter site attributed to in-bed cycling or routine PT/ rehabilitation activities.

6.2 Analysis methods

Statistical analysis
We will analyze patients according to the group to which they were randomized for all outcomes and will use multiple imputation to handle missing data (see section 6.3). We will summarize

baseline characteristics by group reported as mean (SD) or median (first quartile, third quartile) for continuous variables and count (percent) for categorical variables. The criterion for statistical significance will be set at $\alpha = 0.05$ for the Primary Outcome. P-values will be reported to 3 decimal places, with those less than 0.001 reported as $p<0.001$.

Analysis of the Primary outcome

To determine if there is a difference in PFIT-s score at 3 days after ICU discharge between the Cycling and Routine groups, we will conduct a linear regression, including randomization (Cycling vs. Routine) as an independent variable. We will adjust for age (≥ 65 years versus <65 years) and clinical site as these were used as stratification variables in the randomization. We will report the results of the regression as mean difference in PFIT-s with corresponding 95% confidence intervals (CIs) and p-values. We consider a 1-point difference in score clinically important.[14, 16] Although the goal was to have all outcome assessors remain blinded to treatment allocation, this was not always feasible. To maximize use of available data, we will include all PFIT-s measures at 3 days post-randomization, regardless of blinding status of the outcomes assessor and report the proportion of assessments performed by blinded assessors.

To account for incomplete component data in the PFIT-s at 3 days post-ICU, we will concurrently consider data from the PFIT-s, 30s sit-to-stand, and 2-minute walk tests. We will evaluate all PFIT-s data components at 3-days post-ICU. We will identify all patients with any incomplete physical function data and review the scored values for all 4 components of the PFIT-s (i.e., shoulder flexion, knee extension, level of assistance required for the sit-to-stand, and step cadence). In the PFIT-s, a score of "0" represents a lack of physical ability to complete the outcome measure. Thus, if a patient attempted the item and was unsuccessful, the item receives a score of "0", which is a true 0 (Table 4). See supplementary appendix for further details. Table 2 describes the primary outcome analysis.

Analysis of Secondary outcomes

For each continuous secondary outcome, we will conduct a linear regression. We will conduct secondary outcome analyses adjusting for age (≥ 65 years versus <65 years) only. To avoid the risk of overfitting in our analyses, we will not adjust for center when analysing our secondary outcomes. We will report the results of the linear regressions as mean differences with corresponding 95% CIs. If needed to normalize the data, we will perform the linear regression on the log-transformed outcome. If the data are still skewed, we will perform nonparametric analyses. Because secondary analyses are underpowered and hypothesis-generating, we will not present p-values. In the Supplementary appendix, we describe the scoring algorithm to account for incomplete data in the 30 Second Sit to Stand and 2-minute walk tests based on a patient's observed function.

Time to ICU, hospital, and 90-day mortality will be analyzed using Cox proportional hazards regression analysis. We will report hazard ratios (HRs) and corresponding 95% CIs. All other binary outcomes will be analyzed using logistic regression analysis, reporting odds ratios (ORs) with corresponding 95% CIs. We will check the assumptions of the different regression analyses by examining residuals and using other relevant methods. Table 3 describes secondary outcomes analyses. Analysis of Harms is described in Section 6.5.

6.3 Missing data

To account for missing data in our performance-based and patient-reported outcomes, we will use multiple imputation for our analyses.[42-44] In the Cox proportional hazards analyses for ICU and hospital mortality outcomes, patients with incomplete follow-up will be censored at the time of last contact.

6.4 Additional analyses

Subgroup analyses
We will perform three exploratory *a priori* subgroup analyses to investigate potential treatment effect modification for the primary outcome. All subgroup analyses will be adjusted for age and center.

i. Age ≥65 versus <65 years.
ii. Baseline Clinical Frailty ≥5 versus <5.
iii. Male versus female.

In separate linear regression models for each of the three subgroup analyses, we will include randomized treatment allocation, the subgroup variable, and the interaction between the subgroup variable and randomized treatment allocation as independent variables. These analyses will be adjusted for age and center. We hypothesize that the treatment effect will be greater for older compared to younger patients [45], greater in patients with frailty compared to those without [45], and greater in males compared to females [46, 47]. For statistical significance in the subgroup analyses, we will use an alpha of 0.10 for the interaction term. We will assess the credibility of any statistically significant subgroup effect using the method of Schandelmaier *et al*.[48] We will report these data in a forest plot.

Sensitivity analyses
To assess the robustness of our findings, we will conduct five sensitivity analyses for our primary outcome. All sensitivity analyses will be adjusted for age and center, unless specified.

i. To account for ICU mortality on the primary outcome, we will conduct an analysis where we include all patients who died prior to 3-days post-ICU discharge and will assign a PFIT-s score of 0 for those patients.
ii. We will conduct a linear regression analysis that includes only PFIT-s assessments performed by assessors blinded to treatment allocation.
iii. We will conduct an analysis where we only include patients with adherence to the protocol on ≥80% of planned ICU days. Adherence is defined as either received the randomized intervention or had a temporary exemption.
iv. We will investigate the effect of missing data by conducting a complete case analysis, including only those patients with a total PFIT-s score at 3 days post ICU discharge.
v. To determine if the cycling effect is affected by centre, we will conduct analysis adjusting for age only.

See Table 2 for further details.

6.5 Harms
For the safety analysis, we will only include the days on which the patients received the randomized intervention (i.e., days at risk of a safety event associated with rehabilitation

activities). We will report the frequency and percentage of patients with severe and serious adverse events, by group. We will also report the frequency and percentage of randomized intervention days with severe and serious adverse events, by group.

6.6 Statistical software

All analyses will be performed using the most current version of SAS 9.4 (Cary, NC).

6.7 References

1. Herridge MS, Tansey CM, Matte A, Tomlinson G, Diaz-Granados N, Cooper A, Guest CB, Mazer CD, Mehta S, Stewart TE *et al*: **Functional disability 5 years after acute respiratory distress syndrome**. *N Engl J Med* 2011, **364**(14):1293-1304.
2. Iwashyna TJ, Ely EW, Smith DM, Langa KM: **Long-term cognitive impairment and functional disability among survivors of severe sepsis**. *JAMA* 2010, **304**(16):1787-1794.
3. Burtin C, Clerckx B, Robbeets C, Ferdinande P, Langer D, Troosters T, Hermans G, Decramer M, Gosselink R: **Early exercise in critically ill patients enhances short-term functional recovery**. *Crit Care Med* 2009, **37**(9):2499-2505.
4. Schweickert WD, Pohlman MC, Pohlman AS, Nigos C, Pawlik AJ, Esbrook CL, Spears L, Miller M, Franczyk M, Deprizio D *et al*: **Early physical and occupational therapy in mechanically ventilated, critically ill patients: a randomised controlled trial**. *Lancet* 2009, **373**(9678):1874-1882.
5. Camargo Pires-Neto R, Fogaca Kawaguchi YM, Sayuri Hirota A, Fu C, Tanaka C, Caruso P, Park M, Ribeiro Carvalho CR: **Very early passive cycling exercise in mechanically ventilated critically ill patients: physiological and safety aspects--a case series**. *PLoS One* 2013, **8**(9):e74182.
6. Kho ME, Martin RA, Toonstra AL, Zanni JM, Mantheiy EC, Nelliot A, Needham DM: **Feasibility and safety of in-bed cycling for physical rehabilitation in the intensive care unit (ICU)**. *J Crit Care* 2015, **30**(6):e1-5.
7. Kho ME, Molloy AJ, Clarke FJ, Ajami D, McCaughan M, Obrovac K, Murphy C, Camposilvan L, Herridge MS, Koo KK *et al*: **TryCYCLE: A Prospective Study of the Safety and Feasibility of Early In-Bed Cycling in Mechanically Ventilated Patients**. *PLoS One* 2016, **11**(12):e0167561.
8. Kho ME, Molloy AJ, Clarke FJ, Reid JC, Herridge MS, Karachi T, Rochwerg B, Fox-Robichaud AE, Seely AJ, Mathur S *et al*: **Multicentre pilot randomised clinical trial of early in-bed cycle ergometry with ventilated patients**. *BMJ Open Respir Res* 2019, **6**(1):e000383.
9. Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Dore C, Williamson PR, Altman DG, Montgomery A, Lim P *et al*: **Guidelines for the Content of Statistical Analysis Plans in Clinical Trials**. *JAMA* 2017, **318**(23):2337-2343.
10. Morris PE, Goad A, Thompson C, Taylor K, Harry B, Passmore L, Ross A, Anderson L, Baker S, Sanchez M *et al*: **Early intensive care unit mobility therapy in the treatment of acute respiratory failure**. *Crit Care Med* 2008, **36**(8):2238-2243.
11. Koo K: **Survey of Acute Rehabilitation in Canadian Intensive Care Units**. Hamilton, ON: McMaster University; 2012.
12. Denehy L, Skinner EH, Edbrooke L, Haines K, Warrillow S, Hawthorne G, Gough K, Hoorn SV, Morris ME, Berney S: **Exercise rehabilitation for patients with critical illness: a randomized controlled trial with 12 months of follow-up**. *Crit Care* 2013, **17**(4):R156.
13. Kho ME, Heels-Ansdell D, Thabane L, Cook DJ, and the TryCYCLE and CYCLE Pilot RCT investigators: **ICU Discharge Function Predicts Poor Hospital Discharge Outcomes: Lessons from the CYCLE Research Program**. In: *Critical Care Canada Forum: 2017; Toronto, ON Canada*; 2017.

14. Denehy L, de Morton NA, Skinner EH, Edbrooke L, Haines K, Warrillow S, Berney S: **A physical function test for use in the intensive care unit: validity, responsiveness, and predictive utility of the physical function ICU test (scored)**. *Phys Ther* 2013, **93**(12):1636-1645.
15. Skinner EH, Berney S, Warrillow S, Denehy L: **Development of a physical function outcome measure (PFIT) and a pilot exercise training protocol for use in intensive care**. *Crit Care Resusc* 2009, **11**(2):110-115.
16. Parry SM, Denehy L, Beach LJ, Berney S, Williamson HC, Granger CL: **Functional outcomes in ICU - what should we be using? - an observational study**. *Crit Care* 2015, **19**(1):127.
17. Nordon-Craft A, Schenkman M, Edbrooke L, Malone DJ, Moss M, Denehy L: **The physical function intensive care test: implementation in survivors of critical illness**. *Phys Ther* 2014, **94**(10):1499-1507.
18. Pocock SJ: **When (not) to stop a clinical trial for benefit**. *JAMA* 2005, **294**(17):2228-2230.
19. De Jonghe B, Sharshar T, Lefaucheur JP, Authier FJ, Durand-Zaleski I, Boussarsar M, Cerf C, Renaud E, Mesrati F, Carlet J *et al*: **Paresis acquired in the intensive care unit: a prospective multicenter study**. *JAMA* 2002, **288**(22):2859-2867.
20. Ely EW, Truman B, Shintani A, Thomason JW, Wheeler AP, Gordon S, Francis J, Speroff T, Gautam S, Margolin R *et al*: **Monitoring sedation status over time in ICU patients: reliability and validity of the Richmond Agitation-Sedation Scale (RASS)**. *JAMA* 2003, **289**(22):2983-2991.
21. Kho ME, Reid J, Molloy AJ, Herridge MS, Seely AJ, Rudkowski JC, Buckingham L, Heels-Ansdell D, Karachi T, Fox-Robichaud A *et al*: **Critical Care Cycling to Improve Lower Extremity Strength (CYCLE): protocol for an international, multicentre randomised clinical trial of early in-bed cycling for mechanically ventilated patients**. *BMJ Open* 2023, **13**(6):e075685.
22. Zanni JM, Korupolu R, Fan E, Pradhan P, Janjua K, Palmer JB, Brower RG, Needham DM: **Rehabilitation therapy and outcomes in acute respiratory failure: an observational pilot project**. *J Crit Care* 2010, **25**(2):254-262.
23. Hopkins RO, Miller RR, 3rd, Rodriguez L, Spuhler V, Thomsen GE: **Physical therapy on the wards after early physical activity and mobility in the intensive care unit**. *Phys Ther* 2012, **92**(12):1518-1523.
24. Fan E, Ciesla ND, Truong AD, Bhoopathi V, Zeger SL, Needham DM: **Inter-rater reliability of manual muscle strength testing in ICU survivors and simulated patients**. *Intensive Care Med* 2010, **36**(6):1038-1043.
25. Hermans G, Clerckx B, Vanhullebusch T, Segers J, Vanpee G, Robbeets C, Casaer MP, Wouters P, Gosselink R, Van Den Berghe G: **Interobserver agreement of Medical Research Council sum-score and handgrip strength in the intensive care unit**. *Muscle Nerve* 2012, **45**(1):18-25.
26. Jones CJ, Rikli RE, Beam WC: **A 30-s chair-stand test as a measure of lower body strength in community-residing older adults**. *Res Q Exerc Sport* 1999, **70**(2):113-119.
27. Tveter AT, Dagfinrud H, Moseng T, Holm I: **Health-related physical fitness measures: reference values and reference equations for use in clinical practice**. *Arch Phys Med Rehabil* 2014, **95**(7):1366-1373.

28. Pin TW: **Psychometric properties of 2-minute walk test: a systematic review**. *Arch Phys Med Rehabil* 2014, **95**(9):1759-1775.
29. Reid JC, Clarke F, Cook DJ, Molloy AJ, Rudkowski J, Kho ME: **Interrater reliability and responsiveness of the icu patient-reported functional scale**. *Crit Care Med* 2016, **44**(12):284.
30. Reid JC, Clarke F, Cook DJ, Molloy AJ, Rudkowski JC, Stratford PW, Kho ME: **Reliability, responsiveness, and validity of the Patient-Reported Functional Scale for the Intensive Care Unit (PRFS-ICU): A pilot study**. *Annals of Intensive Care* 2017.
31. Wade DM, Hankins M, Smyth DA, Rhone EE, Mythen MG, Howell D, Weinman JA: **Detecting acute distress and risk of future psychological morbidity in critically ill patients: Validation of the Intensive care psychological assessment tool**. *Crit Care* 2014, **18**(5):519.
32. Wade DM, Howell DC, Weinman JA, Hardy RJ, Mythen MG, Brewin CR, Borja-Boluda S, Matejowsky CF, Raine RA: **Investigating risk factors for psychological morbidity three months after intensive care: a prospective cohort study**. *Crit Care* 2012, **16**(5):R192.
33. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, Bonsel G, Badia X: **Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L)**. *Qual Life Res* 2011, **20**(10):1727-1736.
34. van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, Lloyd A, Scalone L, Kind P, Pickard AS: **Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets**. *Value Health* 2012, **15**(5):708-715.
35. Xie F, Pullenayegum E, Gaebel K, Bansback N, Bryan S, Ohinmaa A, Poissant L, Johnson JA, Canadian EQDLVSG: **A Time Trade-off-derived Value Set of the EQ-5D-5L for Canada**. *Med Care* 2016, **54**(1):98-105.
36. Zigmond AS, Snaith RP: **The hospital anxiety and depression scale**. *Acta Psychiatr Scand* 1983, **67**(6):361-370.
37. Bagshaw SM, Stelfox HT, McDermid RC, Rolfson DB, Tsuyuki RT, Baig N, Artiuch B, Ibrahim Q, Stollery DE, Rokosh E, Majumdar SR: **Association between frailty and short- and long-term outcomes among critically ill patients: a multicentre prospective cohort study**. *CMAJ* 2014, **186**(2):E95-102.
38. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW: **Studies of Illness in the Aged. The Index of Adl: A Standardized Measure of Biological and Psychosocial Function**. *JAMA* 1963, **185**:914-919.
39. Nydahl P, Sricharoenchai T, Chandra S, Kundt FS, Huang M, Fischill M, Needham DM: **Safety of Patient Mobilization and Rehabilitation in the Intensive Care Unit. Systematic Review with Meta-Analysis**. *Annals of the American Thoracic Society* 2017, **14**(5):766-777.
40. Adler J, Malone D: **Early mobilization in the intensive care unit: a systematic review**. *Cardiopulm Phys Ther J* 2012, **23**(1):5-13.
41. Sricharoenchai T, Parker AM, Zanni JM, Nelliot A, Dinglas VD, Needham DM: **Safety of physical therapy interventions in critically ill patients: A single-center prospective evaluation of 1110 intensive care unit admissions**. *J Crit Care* 2014, **29**:395-400.
42. Zhang Y, Alyass A, Vanniyasingam T, Sadeghirad B, Florez ID, Pichika SC, Kennedy SA, Abdulkarimova U, Zhang Y, Iljon T *et al*: **A systematic survey of the methods literature on the reporting quality and optimal methods of handling participants**

with missing outcome data for continuous outcomes in randomized controlled trials. *J Clin Epidemiol* 2017, **88**:67-80.
43. Little RJA, Rubin DB: **Statistical Analysis with Missing Data**. Newy York, UNITED STATES: John Wiley & Sons, Incorporated; 2002.
44. Rubin DB: **Multiple imputation for nonresponse in surveys**, vol. 81: John Wiley & Sons; 2004.
45. Jones JRA, Karahalios A, Puthucheary ZA, Berry MJ, Files DC, Griffith DM, McDonald LA, Morris PE, Moss M, Nordon-Craft A *et al*: **Responsiveness of Critically Ill Adults With Multimorbidity to Rehabilitation Interventions: A Patient-Level Meta-Analysis Using Individual Pooled Data From Four Randomized Trials**. *Crit Care Med* 2023, **51**(10):1373-1385.
46. Chi NF, Huang YC, Chiu HY, Chang HJ, Huang HC: **Systematic Review and Meta-Analysis of Home-Based Rehabilitation on Improving Physical Function Among Home-Dwelling Patients With a Stroke**. *Arch Phys Med Rehabil* 2020, **101**(2):359-373.
47. Oikawa SY, Callahan DM, McGlory C, Toth MJ, Phillips SM: **Maintenance of skeletal muscle function following reduced daily physical activity in healthy older adults: a pilot trial**. *Appl Physiol Nutr Metab* 2019, **44**(10):1052-1056.
48. Schandelmaier S, Briel M, Varadhan R, Schmid CH, Devasenapathy N, Hayward RA, Gagnier J, Borenstein M, van der Heijden G, Dahabreh IJ *et al*: **Development of the Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in randomized controlled trials and meta-analyses**. *CMAJ* 2020, **192**(32):E901-E906.
49. Kleyweg RP, van der Meche FG, Schmitz PI: **Interobserver agreement in the assessment of muscle strength and functional abilities in Guillain-Barre syndrome**. *Muscle Nerve* 1991, **14**(11):1103-1109.
50. O'Grady HK, Edbrooke L, Farley C, Berney S, Denehy L, Puthucheary Z, Kho ME, International MCCDG: **The sit-to-stand test as a patient-centered functional outcome for critical care research: a pooled analysis of five international rehabilitation studies**. *Crit Care* 2022, **26**(1):175.
51. Bohannon RW, Wang YC, Gershon RC: **Two-Minute Walk Test Performance by Adults 18 to 85 Years: Normative Values, Reliability, and Responsiveness**. *Arch Phys Med Rehabil* 2015, **96**(3):472-477.
52. Reid JC, Clarke F, Cook DJ, Molloy A, Rudkowski JC, Stratford P, Kho ME: **Feasibility, Reliability, Responsiveness, and Validity of the Patient-Reported Functional Scale for the Intensive Care Unit: A Pilot Study**. *J Intensive Care Med* 2020, **35**(12):1396-1404.
53. Knaus WA, Draper EA, Wagner DP, Zimmerman JE: **APACHE II: a severity of disease classification system**. *Crit Care Med* 1985, **13**(10):818-829.
54. Charlson ME, Pompei P, Ales KL, MacKenzie CR: **A new method of classifying prognostic comorbidity in longitudinal studies: development and validation**. *J Chronic Dis* 1987, **40**(5):373-383.
55. Groll DL, To T, Bombardier C, Wright JG: **The development of a comorbidity index with physical function as the outcome**. *J Clin Epidemiol* 2005, **58**(6):595-602.
56. Reid JC, Clarke F, Cook DJ, Molloy A, Rudkowski JC, Stratford P, Kho ME: **Feasibility, Reliability, Responsiveness, and Validity of the Patient-Reported Functional Scale for the Intensive Care Unit: A Pilot Study**. *J Intensive Care Med* 2019:885066618824534.

Figure 1: CONSORT Flow diagram

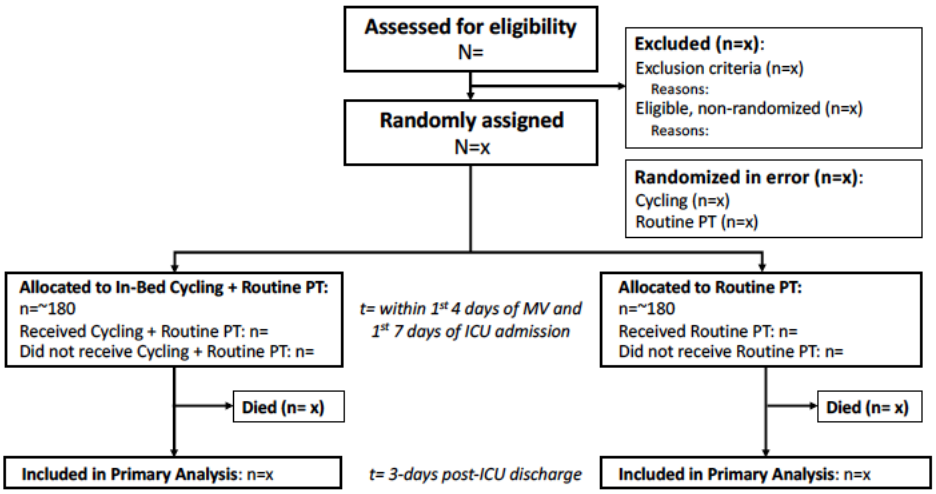

Table 1: Inclusion and Exclusion Criteria

Inclusion Criteria 1. Adults (>=18 years old) 2. Within the first 4 days of MV 3. Expected additional 2 days ICU stay 4. Within the first 7 days of ICU admission 5. Could ambulate independently before hospital admission (with or without a gait aid)
Exclusion Criteria 1. Acute condition impairing patients' ability to cycle (e.g., leg fracture) 2. Acute proven or suspected neuromuscular weakness affecting the legs (e.g., stroke or Guillain-Barré syndrome) 3. Traumatic brain injury 4. Inability to follow commands in local language pre-ICU 5. Severe cognitive impairment pre-ICU 6. Temporary pacemaker (internal or external) 7. Pregnant (or suspected pregnancy) 8. Expected hospital mortality >90% 9. Body habitus unable to fit the bike (e.g., leg amputation, morbid obesity) 10. Specific surgical exclusion as stipulated by surgical or ICU team 11. Palliative goals of care 12. Able to march on the spot at the time of screening 13. Persistent therapy exemptions in the first 4 days of mechanical ventilation: i. Increase in vasopressor/ inotrope within the last 2 hours ii. Active myocardial ischemia, or unstable/ uncontrolled arrhythmia per ICU team iii. Mean arterial pressure <60 or >110 mmHg or per treating team within the last 2 hours iv. Heart Rate <40 or >140 bpm within the last 2 hours v. Persistent SpO_2 <88% or per treating team within the last 2 hours vi. Neuromuscular blocker within the last 4 hours vii. Severe agitation (Richmond Agitation and Sedation Scale >2 [or equivalent] [20]) within last 2 hours viii. Uncontrolled pain ix. Change in goals to palliative care x. Team perception that in-bed cycling or physiotherapy is not appropriate for other new reasons (e.g., acute peritonitis, new incision/wound, known/suspected muscle inflammation (e.g., rhabdomyolysis))
Eligible, non-randomized exclusion criteria 1. Enrolled previously in CYCLE RCT or related study 2. Patient unable to give consent and no substitute decision maker (SDM) identified 3. Patient or SDM declines consent 4. ICU Physician declines patient or SDM to be approached 5. Other, specified by attending team

Table 2: Description of Primary outcome measure and analyses

Analysis of Primary Outcome	**Description of Outcome**	**Measurement Timing**	**Analysis**
Primary Outcome: Physical Function ICU Test-scored[1,2]	Based on 4 patient activities: arm strength, leg strength, ability to stand, and step cadences. Total scores range from 0 to 10 with higher scores meaning better function.	3 days after ICU discharge	Linear regression, adjusted for age and clinical site
Summary of subgroup and sensitivity analyses of the primary outcome			
Objective	**Hypothesis**		**Analysis**
Subgroup Analyses			
i. To determine if age modifies the effect of cycling plus routine PT versus routine PT alone on the primary outcome.	Cycling will be more effective in older patients than in younger patients.		Linear regression adjusted for age and clinical site, which also includes a term for the interaction between age (≥65 versus <65 years) and randomized allocation
ii. To determine if baseline clinical frailty modifies the effect of cycling plus routine PT versus routine PT alone on the primary outcome.	Cycling will be more effective in patients with baseline frailty than in patients without baseline frailty.		Linear regression adjusted for age and clinical site, which also includes the main effect of frailty (≥5 versus <5) and a term for the interaction between frailty and randomized allocation
iii. To determine if sex modifies the effect of cycling plus routine PT versus routine PT alone on the primary outcome.	Cycling will be more effective in male than in female patients.		Linear regression adjusted for age and clinical site, which also includes the main effect of sex and a term for the interaction between sex and randomized allocation
Sensitivity Analyses			
i. To account for ICU mortality on the primary outcome.	Accounting for mortality will not change the effect of cycling on the primary outcome.		Linear regression, adjusted for age and clinical site. All patients will be included. Those who died prior to 3 days post-ICU

			discharge will be assigned a PFIT-s score of 0.
ii.	To determine the effect of cycling plus routine PT versus routine PT alone including only blinded assessment of primary outcome.	Including only patients with blinded assessment of the primary outcome will not change the effect of cycling on the primary outcome.	Linear regression, adjusted for age and clinical site. Only patients with blinded PFIT-s assessments will be included.
iii.	To determine the effect of cycling plus routine PT versus routine PT alone under maximal protocol conditions.	Cycling will more greatly be associated with increased function in patients with higher protocol adherence.	Linear regression, adjusted for age and clinical site. Only patients who received the randomized intervention or had a temporary exemption on ≥80% of planned intervention days will be included.
iv.	To determine the effect of cycling plus routine PT versus routine PT alone in those patients with completed assessment of the primary outcome.	Including only patients with complete assessment of the primary outcome will not change the effect of cycling on the primary outcome.	Linear regression adjusted for age and clinical site. Only patients with a total score for the PFIT-s at 3 days post-ICU discharge will be included.
v.	To determine if the cycling effect is affected by centre, we will conduct analysis adjusting for age only.	Excluding adjustment for clinical site will not change the estimated effect of cycling on the primary outcome.	Repeat the primary linear regression adjusted for age only (i.e., exclude clinical site)

Table 3: Description of Secondary outcome measures and analyses

Outcome	Description of Outcome	Measurement Timing	Analysis
Physical Function ICU Test [14, 15]	Patients complete 4 activities: arm strength, leg strength, ability to stand, and step cadences. Total scores range from 0 to 10 with higher scores representing better function.	ICU awakening, ICU discharge, hospital discharge	Includes survivors at each time point. Separate linear regressions for each timepoint, adjusted for age
Medical Research Council Sum Score [19, 49]	Standardized physical exam of 6 muscle groups (3 upper, 3 lower), using a 6-point scale (0=no contraction; 5= contraction sustained against maximal resistance), summed to a total score. Total scores range from 0 to 60 with higher scores representing more strength.	ICU awakening, ICU discharge, 3 days after ICU discharge, hospital discharge	Includes survivors at each time point. Separate linear regressions for each timepoint, adjusted for age
Medical Research Council Sum Score, categorized as <48 versus ≥48 [19, 49]	Standardized physical exam of 6 muscle groups (3 upper, 3 lower), using a 6-point scale (0=no contraction; 5= contraction sustained against maximal resistance), summed to a total score. Total scores range from 0 to 60 with higher scores representing more strength	ICU awakening, ICU discharge, 3 days after ICU discharge, hospital discharge	Includes survivors at each time point. Separate logistic regressions for each timepoint, adjusted for age
30 second sit to stand test [26, 50]	Patient completes as many full sit to stand repetitions as possible within 30 seconds. Higher scores represent better strength.	ICU awakening, ICU discharge, 3 days after ICU discharge, hospital discharge	Includes survivors at each time point. Separate linear regressions for each timepoint, adjusted for age
2-minute walk test [28, 51]	Patient walks as far as possible over 2 minutes. Higher scores represent better endurance.	ICU discharge, 3 days after ICU discharge, hospital discharge	Includes survivors at each time point. Separate linear regressions for each timepoint, adjusted for age

Outcome	Description of Outcome	Measurement Timing	Analysis
Intensive Care Psychological Assessment [31]	Patients answer 10 questions related to psychological distress in the ICU using a 3-point scale (0=no; 1=yes, a bit; 2=yes, a lot), summed to a total score. Total scores range from 0 to 20, with higher scores representing more distress.	Following ICU awakening assessment	Includes survivors. Linear regression, adjusted for age
Patient-reported functional score for ICU [52]	Patients answer 6 questions about their current perception of function, using an 11-point scale (0=unable to perform activity; 10=able to perform activity at same level as before ICU admission), summed to a total score. Total scores range from 0 to 60, with higher scores representing better function.	ICU discharge, hospital discharge, 90-days post-randomization	Includes survivors at each time point. Separate linear regressions for each timepoint, adjusted for age
Euro-QOL 5D 5L Index [35]	Patients answer 5 questions about their current perception of mobility, self care, usual activities, pain/discomfort, and anxiety/depression, scored according to a prescribed algorithm. Higher scores represent better perceptions of health.	ICU discharge, hospital discharge, 90 days post-randomization	Includes survivors at each time point. Separate linear regressions for each timepoint, adjusted for age
Euro-QOL Visual Analogue Scale [35]	Patients rate their overall health on a 100-point visual analogue scale (0= worst health; 100= best health).	ICU discharge, hospital discharge, 90 days post-randomization	Includes survivors at each time point. Separate linear regressions for each timepoint, adjusted for age
Katz Activities of Daily Living scale [38]	The patient's ability to complete 6 tasks: bathing, dressing, toileting, feeding, continence, and bed mobility. A rater assesses whether the patient is dependent or independent according to pre-specified criteria. Total scores range from 0 to 6,	ICU discharge, hospital discharge	Includes survivors at each time point. Separate linear regressions for each timepoint, adjusted for age

Outcome	Description of Outcome	Measurement Timing	Analysis
	with higher scores representing better function.		
Clinical Frailty Scale [37]	Frailty includes a reduction in physical reserve and loss of function across multiple body systems. The clinical frailty scale is a 9-point scale, with higher scores representing more frailty.	Hospital discharge, 90 days post-randomization	Includes survivors at each time point. Separate linear regressions for each timepoint, adjusted for age
ICU Length of Stay	Days in ICU	ICU discharge	Linear regression, adjusted for age
Hospital Length of Stay	Days in hospital	Hospital discharge	Linear regression, adjusted for age
Mortality	Death	ICU discharge, hospital discharge, 90 days post-randomization	Separate Cox proportional hazards models for each timepoint, adjusted for age
Hospital Discharge Location	Same or better living location at hospital discharge from baseline	Hospital discharge	Includes survivors. Logistic regression, adjusted for age

Table 4: PFIT-s Scoring (adapted from Denehy et al.)[14]

	PFIT-s Component value score			
	0	1	2	3
Shoulder strength	MRC grade 0, 1, or 2	MRC grade 3	MRC grade 4	MRC grade 5
Knee Strength	MRC grade 0, 1, or 2	MRC grade 3	MRC grade 4	MRC grade 5
Sit-to-Stand assistance	Unable	Assist of 2 people	Assist of 1 person	No assistance
Step Cadence	Unable	>0 to 49	50 to <80	>80

MRC = Medical Research Council strength grade (0 to 5)

Supplement Table 1: Summary of Assessments

	Study Period							
	Enrolment	**Allocation**	**Post-allocation**					
Timepoint	ICU admission	0	In ICU	ICU awakening	ICU D/C	3 days post-ICU D/C	Hospital D/C	90 days post-randomization
Enrolment:								
Eligibility screening	X							
Informed consent	X							
Allocation		X						
Interventions:								
In-bed cycling + routine PT			X	X	X			
Routine PT			X	X	X			
Assessments:								
Severity of illness: APACHE II [53]	X							
Charlson Comorbidity Index[54]	X							
Functional Comorbidity Index[55]	X							
Clinical Frailty Scale [37]	X						X	X
Function: Katz Activities of Daily Living scale[38]	X				X		X	
Patient-reported functional scale for the ICU [56]					X		X	X
Physical Strength and Function:				X	X	X (blinded)	X (blinded)	
Physical Function ICU Test (scored)[14, 15]				X	X	X (blinded, Primary)	X (blinded)	
Medical Research Council Sum Score[49]				X	X	X (blinded)	X (blinded)	
30-second sit to stand[26]				X	X	X (blinded)	X (blinded)	

	Study Period							
	Enrolment	**Allocation**	**Post-allocation**					
Timepoint	ICU admission	0	In ICU	ICU awakening	ICU D/C	3 days post-ICU D/C	Hospital D/C	90 days post-randomization
2-minute walk test[28]					X	X (blinded)	X (blinded)	
Psychological distress: Intensive Care Psychological Assessment Tool [31]				X				
Health-related Quality of Life: Euro-QOL 5D-5L[33, 34]					X		X	X
Hospital Anxiety and Depression Scale [36]								X
Data Collection:								
Baseline demographics	X							
Co-interventions			Document daily on CRFs					
Study-related Serious Adverse and Adverse Events			Document daily on CRFs					
Duration of mechanical ventilation								
ICU and Hospital Length of Stay							X	
Mortality					X		X	X

Supplement 2: Scoring algorithm for incomplete data in the PFIT-s

To account for incomplete component data in the PFIT-s, we will concurrently consider data from the PFIT-s, 30s sit-to-stand, and 2-minute walk tests at 3 days post-ICU or relevant time point (i.e., ICU awakening, ICU discharge, hospital discharge). For each patient with incomplete PFIT-s component data, we will evaluate the item within the context of the other scored PFIT-s items. For example, if a patient had a knee strength PFIT-s score of 1 or less (reflecting a Medical Research Council score of 3 or less and representing an inability to complete movement against gravity with any resistance), and had incomplete data for the sit-to-stand assistance component, then we would code sit to stand a score of 0, as this patient would likely not have the strength to stand. Likewise, if a patient completed the sit-to-stand component, needed assistance from 2 or more people, and data for step cadence were incomplete, we would assign step cadence a score of 0, as this patient would likely not have the strength to take any steps. In contrast, if a patient required 1 person to stand, and data for step cadence were incomplete, we would code this as missing data, as we do not have enough information to assess this patient's ability to perform the assessment.

To reduce bias, two research personnel will be blinded to the patient's randomized allocation and clinical site. The research personnel will only evaluate patients with incomplete data. We will document the rationale for all decisions in the study analytic data set. The two research personnel will aim for consensus, however if there is disagreement, we will invite a third clinical research coordinator to make a final judgement. We will use the following scoring algorithm based on a patient's observed function (Table S2):

Table S2: PFIT-s scoring algorithm

		PFIT-s Component value scored if incomplete			
	Observed	Shoulder strength	Knee strength	Sit-to-Stand Assistance	Step Cadence
PFIT-s Components:					
Shoulder strength	Any	N/A	Missing	Missing	Missing
Knee Strength	0 (MRC grade 0, 1, or 2)	Missing	N/A	**0 (unable)**	**0 (unable)**
	1 (MRC grade 3)	Missing	N/A	**0 (unable)**	**0 (unable)**
	2 (MRC grade 4)	Missing	N/A	Missing	Missing
	3 (MRC grade 5)	Missing	N/A	Missing	Missing
Sit-to-Stand assistance	0 (unable)	Missing	Missing	N/A	**0 (unable)**
	1 (assist of 2 people)	Missing	Missing	N/A	**0 (unable)**
	2 (assist of 1 person)	Missing	Missing	N/A	Missing
	3 (no assistance)	Missing	Missing	N/A	Missing
Step Cadence	Unable	Missing	Missing	Missing	N/A
	Able to clear their foot more than once	Missing	Missing	Missing	N/A

		PFIT-s Component value scored if incomplete			
	Observed	Shoulder strength	Knee strength	Sit-to-Stand Assistance	Step Cadence
Other Performance-based outcomes:					
30s Sit to Stand	1 or more repetitions; assistance of 2 or more people	Missing	Missing	**1 (assist of 2 people)**	**0 (unable)**
2-minute walk	Any score >0 m	Missing	Missing	Missing	Missing
	Any score >0 m and information re: assistance	Missing	Missing	Missing	Missing

N/A = not applicable; MRC = Medical Research Council strength grade (0 to 5)

Supplement 3: Scoring algorithm for incomplete data in the 30 Second Sit to Stand and 2-minute walk tests

30 Second Sit to Stand (30STS): We will review the 30STS within the context of the knee strength and sit-to-stand assistance components of the PFIT-s at each corresponding time point. If a patient had a knee strength PFIT-s score of 1 or less (reflecting a Medical Research Council score of 3 or less and representing an inability to complete movement against gravity with any resistance) and had missing data for 30STS, then we would code a score of 0 repetitions, as this patient would likely not have the strength the stand. If a patient received a sit-to-stand PFIT-s item score of 0 (representing inability to stand even with assistance), we will code the 30STS a score of 0, representing 0 repetitions since this patient cannot stand.

2-minute walk test (2MWT): We will review the 2MWT within the context of the knee strength, sit-to-stand assistance, and step cadence components of the PFIT-s at each corresponding time point. If a patient had a knee strength PFIT-s score of 1 or less, then we would code the 2MWT a score of 0 m, as they likely do not have the strength to stand or walk. If a patient completed sit-to-stand and required assistance from 2 or more people, and data for 2MWT was missing, we could code the 2MWT a score of 0, as this patient would likely not have the strength to walk. If a patient received a PFIT-s step cadence score of 1 or higher (representing ability to take at least 1 step) and data for the 2MWT was missing, then we would code this as missing data, as the patient may have been able to walk some distance.

We will use the same procedures to reduce bias, as described in Supplement 2 for the PFIT-s. We will use the following scoring algorithm based on a patient's observed function (Table S3):

Table S3: 30s Sit to Stand and 2-minute walk scoring algorithm

		Score if incomplete	
PFIT-s Component	Observed	30s Sit to Stand repetitions	2-minute walk distance
Shoulder strength	Any	Missing	Missing
Knee Strength	0 (MRC grade 0, 1, or 2)	**0**	**0**
	1 (MRC grade 3)	**0**	**0**
	2 (MRC grade 4)	Missing	Missing
	3 (MRC grade 5)	Missing	Missing
Sit-to-Stand assistance	0 (unable)	**0**	**0**
	1 (assist of 2 people)	Missing	**0**
	2 (assist of 1 person)	Missing	Missing
	3 (no assistance)	Missing	Missing
Step cadence	0 (unable)	Missing	**0**
	Able to clear their foot more than once	Missing	Missing

Legend: MRC = Medical Research Council Score