Henkō

Statistical Analysis Plan

Research Institute of Hospital '12 de Octubre' ('imas12'), Spain
Trial registration number: NCT06300515
Version: 1
Date: 17/02/2026

Statistical analysis plan (SAP) revision history: no revisions

Roles and responsibility:

Name and address of principal investigators and trial coordinator:
Dr. Carmen Fiuza-Luces (principal investigator)
Physical Exercise and Pediatric Cancer Research Group, Research Institute of Hospital '12 de Octubre' ('imas12'), Spain
Dr. Alejandro Lucia (principal investigator)
Department of Sport Sciences, Faculty of Medicine, Health and Sports, Universidad Europea de Madrid, Spain
Mrs. Eva Garrido Rodríguez (trial coordinator)
Physical Exercise and Pediatric Cancer Research Group, Research Institute of Hospital '12 de Octubre' ('imas12'), Spain
Mrs. Elena Higes Núñez (trial coordinator)
Physical Exercise and Pediatric Cancer Research Group, Research Institute of Hospital '12 de Octubre' ('imas12'), Spain

Name of persons writing SAP:

Principal investigator e-mail: cfiuza.imas12@h12o.es
Physical Exercise and Pediatric Cancer Research Group, Research Institute of Hospital '12 de Octubre' ('imas12'), Spain
Principal investigator e-mail: alejandro.lucia@universidadeuropea.es
Department of Sport Sciences, Faculty of Medicine, Health and Sports, Universidad Europea de Madrid, Spain
Trial coordinator e-mail: evagarrido.imas12@h12o.es
Physical Exercise and Pediatric Cancer Research Group, Research Institute of Hospital '12 de Octubre' ('imas12'), Spain
Trial coordinator e-mail: elena.higes.imas12@h12o.es
Physical Exercise and Pediatric Cancer Research Group, Research Institute of Hospital '12 de Octubre' ('imas12'), Spain

Signatures:

Dr. Carmen Fiuza-Luces (principal investigator) Date: 17/02/2026
Dr. Alejandro Lucia (principal investigator) Date: 17/02/2026
Mrs. Eva Garrido Rodríguez (trial coordinator) Date: 17/02/2026
Mrs. Elena Higes Núñez (trial coordinator) Date: 17/02/2026

Table of contents

1. General information of Henkō study

1.1. Background and rationale

Despite survival improvements after hematopoietic stem cell transplantation (HSCT), it is often associated with severe toxicities which, among other deleterious consequences, negatively impact physical function[1]. Notably, total body irradiation applied in the conditioning regimens preceding the transplant can negatively affect muscle mass as well as important vital sign,[2] cardiorespiratory fitness (CRF), with meta-analytical evidence showing muscle, and especially CRF impairments in childhood HSCT survivors persisting years after treatment.[3] Functional decline is further aggravated by isolation and bedrest during neutropenic phases. As survival rates increase, there is therefore a growing need to develop strategies aiming at minimizing the functional decline and morbidity burden associated with HSCT. One such strategy is physical exercise.[4–8] A meta-analysis of only 3 randomized controlled trials (RCT, total *n*=91, age 3 to 18 years)[9] suggested that, in children/adolescents with cancer undergoing HSCT, concurrent exercise interventions (3 or more weekly sessions during 6-8 weeks) improves functional capacity, as assessed with the stair climbing test.[9] However, no significant improvements were found for muscle strength or CRF,[9] underlining a gap in the current state of knowledge. Furthermore, actual *implementability* needs to be addressed for this type of intervention.

1.2. Objective

The main aim of the present RCT (the *Henkō*, in Japanese *transformation*) is to investigate the effects of a supervised concurrent exercise intervention, lasting from hospitalization for childhood/adolescent HSCT until 8-9 weeks post-discharge, on muscle strength (primary outcome) and several fitness, clinical and biological variables (secondary outcomes) in children and adolescents with cancer. We also assess the *implementability* of the proposed intervention within the Reach, Effectiveness, Adoption, Implementation, and Maintenance (RE-AIM) evaluation framework.

The primary aim is to determine the effects of a supervised concurrent exercise intervention, lasting from hospitalization for childhood/adolescent HSCT until 8-9 weeks post-discharge, on unilateral knee-extension muscle strength assessed using a 5-repetition maximum (5-RM) test (primary outcome) in children and adolescents with cancer.

Secondary aims are to:

i. investigate the effects of a supervised concurrent exercise intervention, lasting from hospitalization for childhood/adolescent HSCT until 8-9 weeks post-discharge, on other muscle strength variables (unilateral maximal voluntary isometric contraction of the elbow flexor and knee extensor muscles (at 90° angle), handgrip strength, maximum inspiratory muscle strength) in children and adolescents with cancer.
ii. investigate the effects of a supervised concurrent exercise intervention, lasting from hospitalization for childhood/adolescent HSCT until 8-9 weeks post-discharge, on cardiopulmonary fitness in children and adolescents with cancer.

iii. investigate the effects of a supervised concurrent exercise intervention, lasting from hospitalization for childhood/adolescent HSCT until 8-9 weeks post-discharge, on functional mobility variables in children and adolescents with cancer.

iv. investigate the effects of a supervised concurrent exercise intervention, lasting from hospitalization for childhood/adolescent HSCT until 8-9 weeks post-discharge, on physical activity in children and adolescents with cancer.

v. investigate the effects of a supervised concurrent exercise intervention, lasting from hospitalization for childhood/adolescent HSCT until 8-9 weeks post-discharge, on DXA-determined body composition in children and adolescents with cancer.

vi. investigate the effects of a supervised concurrent exercise intervention, lasting from hospitalization for childhood/adolescent HSCT until 8-9 weeks post-discharge, on Mediterranea diet adherence in children and adolescents with cancer.

vii. investigate the effects of a supervised concurrent exercise intervention, lasting from hospitalization for childhood/adolescent HSCT until 8-9 weeks post-discharge, on psychological status in children and adolescents with cancer.

viii. investigate the effects of a supervised concurrent exercise intervention, lasting from hospitalization for childhood/adolescent HSCT until 8-9 weeks post-discharge, on disease burden in children and adolescents with cancer.

ix. investigate the effects of a supervised concurrent exercise intervention, lasting from hospitalization for childhood/adolescent HSCT until 8-9 weeks post-discharge, on cardiac dimensions and function in children and adolescents with cancer.

x. investigate the effects of a supervised concurrent exercise intervention, lasting from hospitalization for childhood/adolescent HSCT until 8-9 weeks post-discharge, on blood biomarkers (cardiac damage, cardiometabolic-related proteome, inflammation, serum lipid profile, glucose control, blood immune phenotype) in children and adolescents with cancer.

xi. investigate the effects of a supervised concurrent exercise intervention, lasting from hospitalization for childhood/adolescent HSCT until 8-9 weeks post-discharge, on gut microbiome in children and adolescents with cancer.

xii. investigate the implementability of the proposed intervention within the RE-AIM evaluation framework.

2. Study methods

The Henkō RCT follows a hybrid design type 1 unifying the efficacy and effectiveness of research (EER) with implementation of research (IR) (**Figure 1**).

2.1. Efficacy and effectiveness of research (EER)

2.1.1. Trial design

The EER uses a two-arm RCT (ClinicalTrials.gov, NCT06300515, submission date: March 1st, 2024) carried out in children/adolescents undergoing HSCT due to any type of cancer, from the conditioning phase until 8-9 weeks after discharge. At least 60 participants (male/female, 4-18 years, receiving HSCT due to any cancer, showing adequate baseline health status, understanding Spanish language and

providing written informed consent) are being recruited from 3 hospitals (Madrid, Spain) and randomly allocated to an intervention or control group. The intervention consists of supervised concurrent (aerobic and resistance) exercise sessions (3-5 days/week) performed inside isolation wards from transplantation until neutrophil engraftment, and thereafter in the hospital gymnasium, and/or at home. It also includes specific inspiratory muscle training (5 days/week). The *Henkō* research team will perform the data curation and statistical analyses.

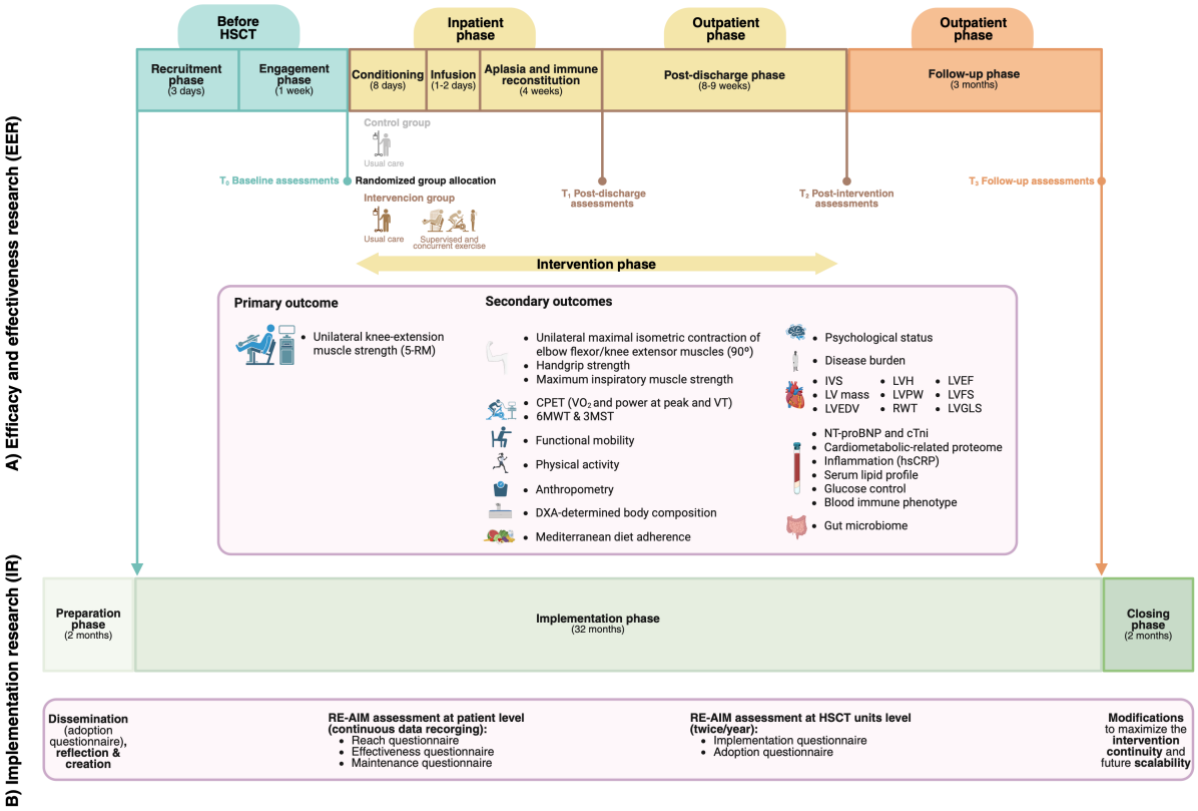

Figure 1. Overview of the *Henkō* Study. Abbreviations: 3MS5, 3-minute step test; 5-RM, 5-repetition maximum; 6MWT, 6-minute walking test; CPET, cardiopulmonary exercise testing; cTni, high-sensitivity cardiac troponin-I; DXA, dual-energy X-ray absorptiometry; hsCRP, high-sensitivity C-reactive protein; HSCT, hematopoietic stem cell transplantation; IVS, interventricular septum thickness; LV, left-ventricular; LVEDV, left-ventricular end-diastolic volume; LVEF, left-ventricular ejection fraction; LVFS, left-ventricular fractional shortening; LVGLS, left-ventricular global longitudinal strain; LVH, left-ventricular hypertrophy; LVPW, left-ventricular posterior wall thickness; NT-proBNP, N-terminal prohormone of brain natriuretic peptide; RE-AIM, Reach, Effectiveness, Adoption, Implementation, and Maintenance; RWT, relative wall thickness; VO_2, oxygen uptake; VT, ventilatory threshold.

2.1.2. Randomization

Participants are randomized using a parallel design with a 1:1 allocation ratio using a computer-generated random allocation sequence with a block on age and sex, after completing the baseline measurements.

2.1.3. Sample size

Considering prior evidence on the effects of exercise on muscle strength in children/adolescents with cancer,[5,10] we expect to find a mean between-group difference ± standard deviation ≥5.6 ± 6 kg for the primary study outcome (corresponding to a large standardized effect size (Cohen's f = 0.45)), knee-extension strength assessed using a 5-repetition maximum (5RM) test. Setting a confidence level of 95% with a statistical power of 90%, we need to recruit a minimum of 48 participants (n≥24 per group; expected dropout rate, 20%)—sample size calculation performed using the statistical software G*Power version 3.1.9.7.

2.1.4. Framework

We will use a superiority hypothesis testing framework. In the main analyses, we will compare whether a supervised concurrent exercise intervention, lasting from hospitalization for childhood/adolescent HSCT until 8-9 weeks post-discharge combined with usual care is superior to usual care alone.

2.1.5. Statistical interim analyses and stopping guidance

No pre-specified interim analyses are performed and hence, a stopping guidance will not be applicable.

2.1.6. Timing of the final analyses

The final analyses will be conducted after the data collection completion and processing of the primary and secondary outcomes.

2.1.7. Timing of outcome assessment

The primary (i.e., muscle strength) and secondary outcomes (i.e., other muscle strength variables, cardiopulmonary fitness, functional mobility variables, physical activity, DXA-determined body composition, Mediterranea diet adherence, psychological status, disease burden, cardiac dimensions and function, blood biomarkers, gut microbiome) will be assessed at four timepoints (T_0, T_1, T_2, T_3) in both groups: T_0, baseline (before hospital admission); T_1, discharge; T_2, 8-9 weeks post-discharge; and T_3, 3-month follow-up by the same trained staff to ensure consistency and reliability. The time frame of each outcome is explained in section *'5.1 Outcome definitions'*.

2.2. Implementation research (IR)

The evaluation of the external validity of the intervention, implementation feasibility and possible transferability of the results using the RE-AIM framework will comprise the same cohort of patients included in the abovementioned EER, on the one hand, and healthcare professionals from the pediatric HSCT units with contracts of ≥6 months and signed informed consents, on the other. Professionals who do not meet the study protocol are excluded. The RE-AIM framework is adopted in three phases, as explained below.

2.2.1. Preparation phase (2 months)

Dissemination. This includes presentation of the trial at the pediatric HSCT units to oncologists, nurses, and physiotherapists; detection of clinical staff members willing to participate; evaluation of the reasons to participate using the adoption questionnaire part 1; and reflection and characterization of the

implementation framework—specifying implementation strategies, underlying principles of the implementation process, clinical staff, spaces, and resources needed.

Reflection. This includes individual meetings with the clinical staff members who expressed interest in assessing their preferences and determining the specific role that they would ultimately assume. Each session lasts 15-30 minutes and is conducted online or in person. The actions to make the physical exercise intervention implementable are specified, as well as the barriers/facilitators that may influence their effectiveness.

Creation. Here, the initial plan, objectives, schedule and milestones, and start of the training of the team are adjusted to develop the physical exercise intervention as well as the outcome assessments. Specific training comprising theoretical and practical sessions is carried out with the finally included clinical staff members.

2.2.2. Implementation phase (32 months)

This includes development and execution of the physical exercise intervention; implementation of strategies to strengthen implementation (*i.e.,* one visit/week of the research team to each pediatric HSCT unit, continuous communication between the clinical and research teams, collaborative learning, access to educational resources, audits); RE-AIM patient level assessment—*i.e.,* implementation is accompanied by continuous data recording to assess scope (reach questionnaire), effectiveness (conducting EER assessment tests and with the effectiveness questionnaire) and maintenance (maintenance questionnaire); RE-AIM institutional level assessment—*i.e.,* the clinical and research team evaluates the implementation component using the implementation questionnaire and the results of the reach, effectiveness, and maintenance evaluation; and dissemination—*i.e.,* detection of participating clinical staff members and training and evaluation of adoption.

2.2.3. Closing phase (2 months)

This consists of RE-AIM framework analysis including any modification of the initial plan to maximize the continuity of physical exercise intervention within the pediatric HSCT units involved and its future scalability.

3. Statistical principles

3.1. P values

To address multiplicity due to the large number of outcomes, a hierarchical and domain-based inference strategy will be applied. For the primary outcome domain (i.e., muscle strength), *p-values* will not be adjusted, and a two-tailed level of statistical significance will be established at $p \leq .05$. For secondary outcomes, statistical significance will be set at *p-values* <0.01. For biological interpretation, intervention-responsive proteins will be mapped to their corresponding genes, and protein–protein interaction networks and functional enrichment analyses will be explored using the STRING database to identify overrepresented biological pathways and processes, as previously reported.[11,12]

3.2. Adherence and protocol deviations

Adherence will be defined as the percentage of exercise sessions attended by adolescents, recorded by the trainers, divided by the number of exercise sessions offered (i.e., 3-5 sessions/week during ~14 weeks in total).

3.3. Analysis populations

All primary analyses will be conducted under an intention-to-treat framework, in which all randomized participants will be analyzed in their originally assigned groups regardless of adherence, protocol deviations after randomization, or withdrawal, assuming that missing data are missing at random. Therefore, some participants will have valid data at all time points, while others may have missing data at any timepoint.

4. Trial population

4.1. Screening data

Screening data will be based on patients who are defined as eligible by the medical and research team through 3 hospitals located in Madrid city (Spain): *Hospital General Universitario Gregorio Marañón* (HGUGM, #PI23/00396), *Hospital Universitario Infantil Niño Jesús* (HIUNJ, #R-0065/23), and *Hospital Universitario La Paz* (HULP, #6593). The medical team initially screens participants for eligibility and introduces them to the study. Thereafter, researchers contact eligible participants to conduct interviews and confirm that they meet the inclusion criteria. Those fulfilling these criteria and willing to enter the study are given both verbal and written information about the study.

4.2. Eligibility

We will include both males and females, aged 4 to 21 years, undergoing HSCT for any cancer, understanding Spanish language, showing adequate baseline health status (Karnofsky $\geq$50, Eastern Cooperative Oncology Group scale score $\leq$2), and providing written informed consent. Exclusion criteria are any type of severe neuro-locomotor, cardiopulmonary condition contraindicating exercise practice.

4.3. Recruitment

The information necessary for the CONSORT flow diagram will be collected.[13] For the enrolment phase, we will consider the number of patients that are assessed for eligibility by the medical team through the 3 included hospitals and the researchers, the number of excluded patients (plus reason for exclusion), and the number of randomized patients. For the allocation, the number of patients allocated to the intervention, and the number of patients who received or did not receive the intervention (plus reasons) will be noted. For follow-up, the number of patients lost to follow-up and the number who discontinued the intervention (plus reasons) will be considered. Finally, the number of patients included in the analyses using the intention-to-treat principle will be described.

4.4. Withdrawal/follow-up

The number, timing, and reasons (i.e., lost motivation, adverse events, withdrew consent, could not or did not want to attend the post-intervention evaluations) for withdrawal will be counted.

4.5. Baseline patient characteristics

A table will be created to describe the characteristics of the study cohort at baseline. The characteristics include general sociodemographic outcomes (i.e., age, sex, hospital, race), anthropometry (i.e., body mass index) and disease burden (i.e., diagnosis, HSCT type/number). The characteristics of the total study cohort and each study arm will be summarized using mean (SD) or median (interquartile range) for normally and not normally distributed continuous variables, respectively, and as number (percentage) for categorical variables.

5. Analysis

5.1. Outcome definitions

Primary outcome

Muscle strength

1. Change in unilateral knee-extension muscle strength (baseline and 8-9 weeks post-discharge). Unilateral knee-extension muscle strength will be assessed using a 5-RM test.

Secondary outcomes

Muscle strength

2. Change in unilateral knee-extension muscle strength (baseline and 3-month follow-up). Unilateral knee-extension muscle strength will be assessed using a 5-RM test.
3. Change in unilateral maximal voluntary isometric contraction of the elbow flexor and knee extensor muscles (at 90º angle) (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Unilateral maximal voluntary isometric contraction of the elbow flexor and knee extensor muscles (at 90º angle) will be assessed using a portable digital dynamometer tissue.
4. Change in handgrip strength (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Handgrip strength will be assessed using a handheld digital Smedley dynamometer (TKK 5401, Takei Scientific Instruments Co., Ltd., Niigata, Japan).
5. Change in maximum inspiratory muscle strength (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Maximum inspiratory muscle strength will be assessed using a mouth pressure meter (CareFusion MicroRPM Respiratory Pressure Meter; Kent, UK).

Cardiopulmonary exercise testing (direct assessments):

6. Change in VO_2 ($mL \cdot kg^{-1} \cdot min^{-1}$) at peak (baseline, 8-9 weeks post-discharge, and 3-month follow-up). VO_2 ($mL \cdot kg^{-1} \cdot min^{-1}$) at peak will be assessed using a ramp-like bicycle ergometer.
7. Change in VO_2 ($mL \cdot kg^{-1} \cdot min^{-1}$) at the ventilatory threshold (baseline, 8-9 weeks post-discharge, and 3-month follow-up). VO_2 ($mL \cdot kg^{-1} \cdot min^{-1}$) at the ventilatory threshold will be assessed using a ramp-like bicycle ergometer.
8. Change in power output (watts) at peak (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Power output (watts) at peak will be assessed using a ramp-like bicycle ergometer.

9. Change in power output (watts) at the ventilatory threshold (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Power output (watts) at the ventilatory threshold will be assessed using a ramp-like bicycle ergometer.

Cardiorespiratory fitness (indirect assessments):

10. Change in walking distance covered (m) (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Walking distance will be assessed using the 6-minute walking test.
11. Change in heart rate recovery (beats/min) (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Heart rate recovery will be assessed using a heart rate monitor after completing the 3-minute step test.

Functional mobility:

12. Change in functional mobility (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Functional mobility will be assessed using the Timed Up and Go, Timed Up and Down Stairs, 30-second chair stand tests, 5-times Sit-To-Stand and Quick Test.
13. Change in range of motion of the ankle (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Range of motion of the ankle will be assessed using a goniometer (Baseline Evaluation Instruments, Fabrication Enterprises Inc.; Elmsford, NY).

Physical activity:

14. Change in physical activity (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Physical activity will be assessed using the Godin Leisure-Time Exercise Questionnaire (GLTEQ).

Anthropometry and body composition:

15. Change in body mass index (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Body mass index will be calculated dividing body weight in kilograms by the square of the height in meters (kg/m^2).
16. Change in waist-to-hip ratio (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Waist-to-hip ratio will be calculated dividing waist by hip circumference (Gulick II Tape Measure, Country Technology, Inc.; Gays Mills, WI) using the same units.
17. Change in arm circumference (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Arm circumference will be assessed under relaxed and contracted conditions (Gulick II Tape Measure, Country Technology, Inc.; Gays Mills, WI).
18. Change in lean mass (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Lean mass will be assessed using a dual-energy X-ray absorptiometry assessment (Hologic Serie Discovery QDR, Software Physician's Viewer, APEX System Software version 3.1.2.; Bedford, MA).
19. Change in fat mass (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Fat mass will be assessed using a dual-energy X-ray absorptiometry assessment (Hologic Serie Discovery QDR, Software Physician's Viewer, APEX System Software version 3.1.2.; Bedford, MA).

20. Change in visceral fat (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Visceral fat will be assessed using a dual-energy X-ray absorptiometry assessment (Hologic Serie Discovery QDR, Software Physician's Viewer, APEX System Software version 3.1.2.; Bedford, MA).
21. Change in fat percentage (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Fat percentage will be assessed using triceps skinfold (Harpenden caliper, Crymych, United Kingdom).
22. Change in bone mineral density of the total body (less head) (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Bone mineral density of the total body (less head) will be assessed using a dual-energy X-ray absorptiometry assessment (Hologic Serie Discovery QDR, Software Physician's Viewer, APEX System Software version 3.1.2.; Bedford, MA).
23. Change in bone mineral density of the femoral neck (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Bone mineral density of the femoral neck will be assessed using a dual-energy X-ray absorptiometry assessment (Hologic Serie Discovery QDR, Software Physician's Viewer, APEX System Software version 3.1.2.; Bedford, MA).
24. Change in phase angle (º) (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Phase angle (º) will be assessed using multi-frequency bioelectrical impedance analysis (BodyComposition software, version 9.0.21212-29; Dietosystem, Italy), and it is calculated as the arctangent of reactance to resistance, following standardized measurement procedures.

Diet:

25. Change in adherence to Mediterranean diet (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Adherence to Mediterranean diet will be assessed using the Mediterranean Diet Quality Index for children and adolescents (KIDMED) questionnaire.

Psychological status:

26. Change in health-related quality of life (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Health-related quality of life will be assessed using the Pediatric Quality of Life Inventory (PedsQL) 3.0 (Patients and Tutor's version, Cancer Module).
27. Change in cancer-related fatigue (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Cancer-related fatigue will be assessed using the Pediatric Quality of Life Inventory (PedsQL) 3.0 (Patients and Tutor's version, Multidimensional Fatigue Scale).

Disease burden:

28. Survival (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Survival will be assessed from diagnosis to the end of the study or death using medical records.
29. Treatment tolerability (baseline, and 8-9 weeks post-discharge). Treatment tolerability will be assessed as the number of days of treatment interruption/delay and hospitalization length (additional/prolonged hospitalization during treatment) using medical records.
30. Change in the number and duration of viral/bacterial/fungal infections (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Number and duration of viral/bacterial/fungal infections will be retrieved from medical records.

31. Change in toxicity grade (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Toxicity grade will be assessed using the formula from Langlais et al (2022).

Cardiac dimensions:

32. Change in left-ventricular (LV) mass (baseline, 8-9 weeks post-discharge, and 3-month follow-up). LV mass will be assessed using 2-D guided M-mode imaging.
33. Change in interventricular septum thickness (IVS) (baseline, 8-9 weeks post-discharge, and 3-month follow-up). IVS will be assessed using 2-D guided M-mode imaging.
34. Change in LV end-diastolic volume (LVEDV) (baseline, 8-9 weeks post-discharge, and 3-month follow-up). LVEDV will be assessed using 2-D guided M-mode imaging.
35. Change in LV posterior wall thickness (LVPW) (baseline, 8-9 weeks post-discharge, and 3-month follow-up). LVPW will be assessed using 2-D guided M-mode imaging.
36. Change in relative wall thickness (RWT) (baseline, 8-9 weeks post-discharge, and 3-month follow-up). RWT will be assessed using the following formula: RWT = (IVS + LVPW)/LVEDV.
37. Change in LV hypertrophy (baseline, 8-9 weeks post-discharge, and 3-month follow-up). LV hypertrophy will be assessed using the age-specific >95th percentile for LV mass indexed by height (in $g \cdot m^{-2.7}$).
38. Change in LV ejection fraction (baseline, 8-9 weeks post-discharge, and 3-month follow-up). LV ejection fraction will be assessed using color tissue Doppler echocardiography.
39. Change in LV fractional shortening (baseline, 8-9 weeks post-discharge, and 3-month follow-up). LV fractional shortening will be assessed using color tissue Doppler echocardiography.
40. Change in LV global longitudinal strain (baseline, 8-9 weeks post-discharge, and 3-month follow-up). LV global longitudinal strain will be assessed using 2D-speckle tracking echocardiography.

Cardiac damage (blood biomarkers):

41. Change in NT-proBNP (baseline, 8-9 weeks post-discharge, and 3-month follow-up). NT-proBNP will be determined with the relevant immunoassay kits on an automated biochemistry analyzer (Cobas C701, Roche Diagnostics; Madrid, Spain).
42. Change in high-sensitivity cardiac troponin-I (baseline, 8-9 weeks post-discharge, and 3-month follow-up). High-sensitivity cardiac troponin-I will be determined with the relevant immunoassay kits on an automated biochemistry analyzer (Cobas C701, Roche Diagnostics; Madrid, Spain).

Cardiometabolic-related proteome (blood biomarkers):

43. Change in cardiometabolic-related proteome (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Cardiometabolic-related proteome will be assessed using the Olink® Cardiovascular panel (Olink Target 96 Cardiovascular II Protein assay list), which measures 92 cardiometabolic-related human proteins simultaneously in plasma (expressed in Normalized Protein eXpression (NPX) values, an arbitrary unit presented in Log2 scale) (https://olink.com/products-services/target/cardiometabolic-panel/).

Inflammation (blood biomarkers):

44. Change in high-sensitivity C-reactive protein levels (basel baseline, 8-9 weeks post-discharge, and 3-month follow-up). High-sensitivity C-reactive protein levels will be assessed using a chemistry analyzer (Cobas C701, Roche Diagnostics; Madrid, Spain).

Serum lipid profile and glucose control (blood biomarkers):

45. Change in total cholesterol (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Fasting blood samples will be used to assess total cholesterol.
46. Change in high-density lipoprotein cholesterol (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Fasting blood samples will be used to assess high-density lipoprotein cholesterol.
47. Change in low-density lipoprotein cholesterol (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Fasting blood samples will be used to assess low-density lipoprotein cholesterol.
48. Change in triglycerides (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Fasting blood samples will be used to assess triglycerides.
49. Change in apolipoprotein B (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Fasting blood samples will be used to assess apolipoprotein B.
50. Change in fasting glycaemia (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Fasting blood samples will be used to assess glycaemia.
51. Change in glycated hemoglobin (baseline, 8-9 weeks post-discharge, and 3-month follow-up p). Fasting blood samples will be used to assess glycated hemoglobin.
52. Change in insulin (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Fasting blood samples will be used to assess insulin.
53. Change in homeostasis model assessment-insulin resistance index (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Fasting blood samples will be used to assess glucose and insulin, and homeostasis model assessment-insulin resistance index will be computed.

Blood immune phenotype (blood biomarkers):

54. Change in total leukocyte and monocyte count (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Total leukocyte and monocyte count will be assessed using a hematology analyzer (Advia 120 Hematology System, Bayer Corporation; Tarrytown, NY).
55. Change in main lymphocyte subpopulations (%) (baseline, 8-9 weeks post-discharge, and 3-month follow-up). Main lymphocyte subpopulations will be assessed on fresh blood samples using a multiparametric flow cytometer (FACSCantoTM II, Becton Dickinson and Company BD Biosciences; San Jose, CA) together with BD FACSDiva™ software version 8 (Becton Dickinson and Company BD Biosciences).

Gut microbiome:

56. Change in gut microbiome diversity (baseline, 8-9 weeks post-discharge, and 3-month follow-up). DNA sequencing to determine gut microbiome diversity (i.e., alpha and beta).
57. Change in specific bacteria abundance (baseline, 8-9 weeks post-discharge, and 3-month follow-up). DNA sequencing to determine specific bacteria abundance.

RE-AIM framework (This analysis will be carried out combining components at the patient (Reach, Effectiveness, Maintenance) and institutional (Adoption, Implementation) level):

58. Reach. This component will evaluate the scope of the intervention within the patients and the representativeness of participants vs non-participants. It will involve prospective data collection during the implementation phase using the reach questionnaire and measuring the following items: exclusion rate and reasons for exclusion; non-participation and recruitment rates; reasons for participating or not; medical, sociocultural, economic, demographic, and motivational conditions of patients/parents (or caregivers); feasibility rate of the EER assessments; acceptability rate of the intervention sessions and reasons for non-execution; dropout rate and reasons; and degree of satisfaction with the intervention using a specific questionnaire.
59. Effectiveness. This outcome will assess the impact of the proposed exercise intervention, vs usual care alone on the outcomes measured at T_0, T_1 and T_2, including potential adverse effects. Prospective data will be collected during the implementation phase using the effectiveness questionnaire.
60. Adoption. This will evaluate the clinical team and the pediatric HSCT units during the preparatory and implementation phases, using the adoption questionnaire. Clinical team. Here, the following items will be considered: participation and non-participation rates; recording of reasons for participation or non-participation (sociocultural, economic, demographic, and motivational characteristics, and dropout rate and reasons); and environment—participation rate, identification of facilitators and barriers to implement the intervention, analysis of the suitability of the implementation strategies, identification of differences between the EER and IR, and intervention costs (i.e., time spent by the clinical and research team participating in the trial associated with the cost per hour of work, and cost of the patient's hospitalization duration).
61. Implementation. This will measure the extent to which the components of the intervention are implemented in their initial version. Using information obtained from the implementation questionnaire during the implementation phase, we will calculate the feasibility rate related to patient detection and recruitment (data recorded with the reach questionnaire), communication between families, patients, research team, and clinician (satisfaction questionnaire), execution of the assessment test battery (measured using the feasibility rate of the EER test battery), development of the intervention sessions (session acceptability rate), and the degree of implementation of the clinical team's IR functions within their usual practice.
62. Maintenance. This will describe the level at which the intervention is integrated into the patient's routine, as assessed using the results of the post-discharge evaluation of EER (8-9

weeks follow-up), the attrition rate, and the reasons for non-compliance (prospectively collected using the maintenance questionnaire during the implementation phase).

Data collection will be conducted through individual and semi-structured online interviews via videoconference at the end of the intervention. A shared Google Calendar will be created to schedule interviews, allowing participants to independently select the day and time that best suits their availability. All interviews will be audio-recorded, ensuring the confidentiality and anonymity of the participants. The audio files will be transcribed verbatim using the TurboScribe software. Subsequently, all transcriptions will be reviewed to verify their accuracy according to the original audio, and to remove any identifying information before analysis. The transcripts will then be imported into the qualitative analysis software NVivo, which will facilitate the management of the material. Following the mixed research principles of Braun and Clarke,[14] and the narrative integration proposals of Fetters,[15] quantitative and qualitative data will be integrated at an interpretive level, using the qualitative findings to contextualize the quantitative results, and understand the processes of adherence, perception and feasibility of the program from the experience of patients and family members.

5.2. Analysis methods

5.2.1. Efficacy and effectiveness of research (EER)

The effects of the exercise intervention on primary and secondary outcomes will be assessed separately using constrained baseline (meaning baseline adjusted) linear mixed models with random intercepts for participants. Repeated measurements, including baseline and follow-up assessments, will be modeled as a function of evaluation time, randomly assigned group, and their interaction, with group means constrained to be equal at baseline.[16] Intervention effects will be estimated from the group-by-time interaction contrasts at post-intervention. In addition, pre-specified adjusted models will include age, sex, and transplant type as covariates. Categorical outcomes will be compared between groups using chi-square or Fisher's exact test, as appropriate based on expected cell counts.

For the primary outcome domain (i.e., muscle strength), *p-values* will not be adjusted, and a two-tailed level of statistical significance will be established at $p \leq .05$. For secondary outcomes, statistical significance will be set at *p-values* <.01 to minimize risk of statistical error type I. For biological interpretation of proteome data, intervention-responsive (i.e., differentially expressed) proteins will be mapped to their corresponding genes, and protein–protein interaction networks and functional enrichment analyses will be explored using the Search Tool for the Retrieval of Interacting Genes/Proteins (commonly known as 'STRING') database to identify enriched biological pathways and processes, as previously reported.[11,12] For gut microbiome profiling based on 16S rRNA gene sequencing, raw paired-end reads will be processed using QIIME 2 (version 2018.6.0),[17] and quality filtering and denoising will be performed using DADA2, followed by taxonomic assignment against a reference database. Alpha-diversity (within-sample diversity) and beta-diversity (between-sample diversity) metrics, as well as the relative abundance of bacterial taxa, will be calculated and visualized to explore quantitative and qualitative differences between groups.[17] Overall survival will be described using Kaplan–Meier curves, defined as time from randomization to death from any cause, with censoring at

the end of follow-up. Given the expected low number of events, survival analyses will be considered exploratory and primarily descriptive, and any between-group comparisons will be interpreted with caution.

All primary analyses will be conducted under an intention-to-treat framework, in which all randomized participants will be analyzed in their originally assigned groups regardless of adherence, protocol deviations after randomization, or withdrawal, assuming that missing data are missing at random. No pre-specified interim analyses will be performed and hence, stopping guidance will not be applicable.

Strict access controls will be put in place to ensure that data are securely stored on hospital computer systems, preserving confidentiality and data integrity. To maintain high data quality, weekly quality assurance reviews will be performed to detect potential issues, including missing data or forms, values outside predefined ranges, incorrect data entries, implausible or inconsistent dates over time, discrepancies across study forms and visits, and incomplete fields without justified reasons. The main statistical analyses will be performed by an independent researcher who is not involved in the recruitment, evaluations, and interventions, and will be conducted blinded to the treatment allocation by coding the intervention arms (i.e., A, B).

5.2.2. Implementation of research (IR)

Exploratory factor analysis of each dimension of the questionnaires applied in the IR will be used to determine the variables with a common meaning and to analyze the structure of the interrelations between them. Principal component analysis will be conducted as an extraction method, and will be used to create a linear combination that explains the highest possible percentage (and at least 60%) of variance: the correlation matrix of the variables, the communalities, the Kaiser-Meyer-Olikin index and the Bartlett sphericity test will be determined; the number of factors to be kept in the analysis will be identified; a factor rotation will be performed following the Varimax rotation method with Kaiser normalization to improve the interpretation of the factor structure; and, once the variables are determined, each factor will be named with an identifying term. Qualitative analysis will follow a reflective thematic analysis approach, based on the phases proposed by Braun and Clarke.[14] Transcripts will be imported into the NVivo software for data management and analysis. The process will include data familiarization, initial inductive coding, code organization, construction and revision of themes and subthemes, and the preparation of analytical memos to support the interpretation of the findings.

5.3. Missing data

We will report and explore missing data and possible patterns. We expect that missing data will be assumed as missing at random. Therefore, the linear mixed model analyses will handle the missing data. Nevertheless, we will reconsider this assumption once the data processing is finalized. If we believe this assumption does not hold, we will take appropriate measures for the data analyses by using other multiple imputation techniques, for example.

5.4. Additional analyses

Changes in other outcomes will be analyzed using a similar protocol as described by '5.2 Analysis methods' unless other analyses would be more appropriate depending on the outcome.

5.5. Harms

Ascertained adverse events related to exercise testing or training sessions, including muscle pain, fatigue, and general aches and pains according to patients/parents/caregivers (self-reported) will be collected by each study arm. No formal statistical testing will be undertaken.

5.6. Statistical software

The analyses will be performed using R. For the main analyses we will use the 'lme4' or 'LMMstar' package. The use of packages will be reported in the manuscript.

6. References

1. Baker KS, Bresters D, Sande JE. The burden of cure: Long-term side effects following Hematopoietic Stem Cell Transplantation (HSCT) in children. *Pediatr Clin North Am*. 2010;57(1):323-342. doi:10.1016/j.pcl.2009.11.008
2. Ross R, Blair SN, Arena R, et al. Importance of Assessing Cardiorespiratory Fitness in Clinical Practice: A Case for Fitness as a Clinical Vital Sign: A Scientific Statement From the American Heart Association. *Circulation*. 2016;134(24):e653-e699. doi:10.1161/CIR.0000000000000461
3. Fridh MK, Simonsen C, Schmidt-Andersen P, et al. Cardiorespiratory fitness and physical performance after childhood hematopoietic stem cell transplantation: a systematic review and meta-analysis. *Bone Marrow Transplant*. 2021;56(9):2063-2078. doi:10.1038/S41409-021-01370-2
4. Smith C, Farhat R, Fern-Buneo A, et al. Effects of an exercise program during pediatric stem cell transplantation: A randomized controlled trial. *Pediatr Blood Cancer*. 2022;69(5). doi:10.1002/PBC.29618
5. Senn-Malashonak A, Wallek S, Schmidt K, et al. Psychophysical effects of an exercise therapy during pediatric stem cell transplantation: a randomized controlled trial. *Bone Marrow Transplant*. 2019;54(11):1827-1835. doi:10.1038/S41409-019-0535-Z
6. Wallek S, Senn-Malashonak A, Vogt L, Schmidt K, Bader P, Banzer W. Impact of the initial fitness level on the effects of a structured exercise therapy during pediatric stem cell transplantation. *Pediatr Blood Cancer*. 2018;65(2). doi:10.1002/PBC.26851
7. Chamorro-Vina C, Ruiz JR, Santana-Sosa E, et al. Exercise during hematopoietic stem cell transplant hospitalization in children. *Med Sci Sports Exerc*. 2010;42(6):1045-1053. doi:10.1249/MSS.0B013E3181C4DAC1
8. Chamorro-Viña C, Valentín J, Fernández L, et al. Influence of a Moderate-Intensity Exercise Program on Early NK Cell Immune Recovery in Pediatric Patients After Reduced-Intensity Hematopoietic Stem Cell Transplantation. *Integr Cancer Ther*. 2017;16(4):464-472. doi:10.1177/1534735416679515

9. Morales JS, Valenzuela PL, Rincón-Castanedo C, et al. Exercise training in childhood cancer: A systematic review and meta-analysis of randomized controlled trials. *Cancer Treat Rev*. 2018;70:154-167. doi:10.1016/j.ctrv.2018.08.012
10. Fiuza-Luces C, Padilla JR, Soares-Miranda L, et al. Exercise Intervention in Pediatric Patients with Solid Tumors: The Physical Activity in Pediatric Cancer Trial. *Med Sci Sports Exerc*. 2017;49(2):223-230. doi:10.1249/MSS.0000000000001094
11. Carrera-Bastos P, Plaza-Florido A, Santos-Lozano A, et al. The 'autoimmunome' of centenarians. *J Transl Autoimmun*. 2025;11. doi:10.1016/j.jtauto.2025.100295
12. Santos-Lozano A, Plaza-Florido A, Carrera-Bastos P, et al. Tumor Signatures of Physical Fitness: Insights from a Preclinical Model. *Med Sci Sports Exerc*. 2026;58(1). doi:10.1249/MSS.0000000000003836
13. Hopewell S, Chan AW, Collins GS, et al. CONSORT 2025 statement: updated guideline for reporting randomized trials. *Nat Med*. Published online April 15, 2025. doi:10.1038/S41591-025-03635-5
14. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. 2006;3(2):77-101. doi:10.1191/1478088706QP063OA
15. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. *Health Serv Res*. 2013;48(6 Pt 2):2134-2156. doi:10.1111/1475-6773.12117
16. Coffman CJ, Edelman D, Woolson RF. To condition or not condition? Analysing "change" in longitudinal randomised controlled trials. *BMJ Open*. 2016;6(12). doi:10.1136/BMJOPEN-2016-013096
17. Bárcena C, Valdés-Mas R, Mayoral P, et al. Healthspan and lifespan extension by fecal microbiota transplantation into progeroid mice. *Nat Med*. 2019;25(8):1234-1242. doi:10.1038/S41591-019-0504-5