STATISTICAL ANALYSIS PLAN

Total Hip Arthroplasty versus Progressive Resistance Training for Improving Accelerometry assessed Physical Activity Outcomes in Persons with Severe Hip Osteoarthritis

Statistical Analysis Plan for a Secondary Analysis of the Randomized Controlled PROHIP Trial

Trial ID: Progressive Resistance Training versus Total Hip Arthroplasty (PROHIP)

Trial registration: ClinicalTrials.gov, NCT04070027

SAP version: Version 1.0, 2025-04-23.

Date: April 23, 2025

Based on Protocol Version: Version 1.3, 2020-08-21

Authors/Collaborators

Thomas Frydendal, PT, PhD; Rune Amstrup, BSc; Lisa Urup Tønning, PT, MSc; Kim Gordon Ingwersen, PT, PhD; Søren Overgaard, MD, DMSc; Lone Ramer Mikkelsen, PT, PhD; Robin Christensen, MSc, PhD; Inger Mechlenburg, PT, DMSc.

Authors writing the SAP

Thomas Frydendal, Lisa Urup Tønning; Inger Mechlenburg; and Robin Christensen.

Statistical analyst

Lisa Urup Tønning, PT, MSc

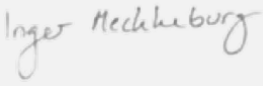
Senior biostatistician responsible

Robin Christensen, BSc, MSc, PhD; Professor of Biostatistics and Clinical Epidemiology.

Chief investigator/clinical lead

Inger Mechlenburg, PT, PhD, DMSc; Professor of Orthopedic Rehabilitation.

SIGNATURES

Date	Name	Roles and Responsibility	Signature
May 13, 2025	Thomas Frydendal	SAP first author	
May 8, 2025	Robin Christensen	Senior Biostatistician	
May 8, 2025	Inger Mechlenburg	Chief Investigator	

INTRODUCTION

Background

Hip osteoarthritis is a main contributor to disability,[1-5] and the leading cause for undergoing total hip arthroplasty (THA).[6-8] This surgical procedure is considered an effective treatment for reducing hip pain and functional impairments and improving quality-of-life in patients with severe hip osteoarthritis.[8] Even though more than one million THAs are performed annually worldwide,[8] randomized controlled trials comparing this procedure effectiveness to first-line treatment have been lacking.[9,10] Across six international clinical guidelines exercise such as progressive resistance training (PRT) is consistently recommended as a first-line treatment for hip osteoarthritis, with two guidelines recommending this regardless of pain, age and/or severity of disease.[11]

We recently reported the results from the Progressive Resistance Training versus Total Hip Arthroplasty randomized trial (PROHIP), which showed that THA followed by standard care resulted in clinically important greater relief in patient-reported hip pain and improvement in patient-reported function at 6 months, as compared with resistance training.[12] Our results align with those of a recent propensity-matched analysis using data from two prospective registries, which also demonstrated greater benefits of THA when compared to patient education and neuromuscular exercise.[13]

Rationale for this study

Although we found that THA was superior to PRT in improving patient-reported outcomes for persons with severe hip osteoarthritis who had an indication for surgery, our results did not show additional benefits of THA over PRT on performance-based gait and sit-to-stand function.[12] Among persons undergoing THA, objectively measured physical activity levels appear to be unchanged compared to before surgery and is considerably lower than healthy age- and sex-matched controls 6 to 12 months after surgery.[14,15] This lack of increase in physical activity level could be due to sedentary behavior adopted by the persons prior to surgery,[15] experience of pain in other joints, limitations related to comorbidities, or uncertainty.[16,17]

However, it remains unclear whether THA leads to greater improvements in objectively measured physical activity outcomes compared to PRT. Tracking number of daily steps helps assess how well persons with severe hip osteoarthritis are integrating back into their daily routines and activities. As such, a higher step count may indicate better recovery and return to normal life, and is associated with lower mortality and risk of cardiovascular disease.[18] Steps taken also provides a measure of a person's mobility and comparing step counts between different treatment groups can highlight the effectiveness of specific treatments in improving mobility.

The primary null hypothesis is that there is no between-group difference on changes in the number steps per day. The primary alternative hypothesis is that persons randomly assigned to receive THA will improve more in number of steps per day than those randomly assigned to receive PRT baseline to 6 months.

Objectives

Primary objective: To compare the effectiveness of THA, relative to PRT, on changes in the number of steps per day from baseline to 6 months after initiating the treatment, in persons 50 years of age or older with severe hip osteoarthritis and an indication for surgery.

Secondary objectives: To compare the effectiveness of THA versus PRT from baseline to 6 months after initiating the treatment on the following two outcomes: number of sit-to-stand transfers and number of short walking bouts per day.

Other objectives: To compare the effectiveness of THA versus PRT from baseline to 6 months after initiating the treatment on the following outcomes: average walking cadence, percentage of time spent in very low-, low-, moderate- and high intensity activity, hours per day walking, hours standing, hours sedentary, hours cycling, and number of cycle rotations per day.

STUDY METHODS

Trial design

This is a secondary analysis of the PROHIP trial. This was designed as a multicenter (4 sites), stratified (by recruitment site), randomized, controlled, parallel-group superiority trial. Participants were randomly assigned to undergo THA followed either by standard care or to participate in a PRT program, with a 1:1 allocation. Follow-up assessment of physical activity measures were conducted at 6 months after the treatment was initiated. The PROHIP trial was approved by The Regional Committees on Health Research Ethics for Southern Denmark (Project-ID: S-20180158) and the Danish Data Protection Agency (Journal No 19/20337).

We prespecified and registered the PROHIP trial at ClinicalTrials.gov (NCT04070027), and we have previously published the trial protocol[19] and the primary results.[12] This statistical analysis plan (SAP) is reported in line with the '*Guidelines for the Content of Statistical Analysis Plans in Clinical Trials'*,[20] while the reporting of this secondary analysis will follow the *"Consolidated Standards of Reporting Trials"* (CONSORT) statement.[21] All participants provided written informed consent.

Randomization and Blinding

After baseline measures were completed, participants were randomly assigned in a 1:1 ratio to either undergo THA or to participate in PRT, according to a computer-generated randomization list developed by and independent data manager in an electronic management system.[22] The list was based on permuted random blocks of variable size (2 to 6), and randomization was stratified by recruitment hospital site (4 levels). Treatment assignment was concealed until the randomization procedure was completed in the in an electronic management system[22] by a local trial coordinator. Participants and clinicians (orthopedic surgeons, nurses and physiotherapists) were aware of the treatment assignments in the trial. Outcome assessors, investigators, and data analysts will remain unaware to treatment assignments until all analyses, as described in this SAP, are completed.

Sample size and power considerations

The PROHIP trial was designed with an effective power of 92% to detect a minimal important difference on the primary endpoint (i.e., the Oxford Hip Score). This power calculation was based on a t-test on the mean difference (two-sided alpha level of 0.05) assuming equal variances, displaying that a sample size of 120 participants (i.e., 60 per group) were needed. The detailed sample size and power considerations are provided in the trial protocol.[19] For this secondary analysis, we did not perform a formal power or sample size calculation, but used data obtained from the 109 participants who accepted enrollment being randomly assigned to either THA or PRT in the PROHIP trial.[12]

Framework

The PROHIP trial was designed as a trial with the primary objective of showing that the response to the investigational treatment is superior to the comparative treatment (i.e., THA is superior to PRT).

Statistical interim analyses and stopping guidance

No interim analyses were performed during the PROHIP trial. The trial was not stopped early for benefit or harm, but we ended trial enrollment on June, 30 2021 in line with the pre-specified recruitment deadline before reaching the intended sample size of 120.[19] As such, no interim statistical methods, analyses, or specific timelines for interim assessments were planned or carried out. We monitored enrollment and attrition rates throughout the trial.

Timing of final analysis and outcome assessments

Outcome assessments for this secondary analysis of the PROHIP trial were conducted at two time points: baseline and at 6 months after the assigned treatment was initiated. No additional outcome assessments were planned during the trial. The timing of the 6-month follow-up assessment was calculated by adding 26 weeks to the date the assigned treatment (i.e., THA or PRT) was initiated. The specific time points for all outcome measurements are elaborated on in the published trial protocol.[19] We will perform the final analysis after this SAP is publicly available at ClinicalTrials.gov, as we collected all baseline and 6 months follow-up data between September 3, 2019 to May 5, 2022.

STATISTICAL PRINCIPLES

Confidence intervals and P values

All 95% confidence intervals (95%CIs) and P values will be two sided; for the primary endpoint superiority will be defined as $P<0{\cdot}05$. We will not apply explicit adjustments for multiplicity. Instead, key secondary outcomes will be analyzed in a prioritized order (as listed in the Table 2 Mockup) using a gatekeeping strategy, and the interpretation of these findings will account for the risk of spurious results. For all primary and secondary endpoints, the 95% confidence intervals will not be adjusted for multiplicity and should not be used as a substitute for hypothesis testing.

The analyses of the two key secondary endpoints will be performed in sequence until one of the analyses fails to show the statistically significant difference, or until all analyses have been completed at a statistical significance level of 0.05. The two secondary statistical tests will be reported with P values for hypothesis tests and tentative claims of statistical significance.

Adherence and protocol deviations

We reported adherence to the two treatments previously in the primary manuscript as number and percentage of participants who were adherent, along with the number and percentage (with reasons) for participants who were non-adherent, such as treatment crossover (i.e., declining surgery in the THA group or receiving THA in the PRT group) or early discontinuation.[12] Protocol deviations for this secondary analysis will be defined as participants having less than three days of physical activity measurements containing at least 10 hours of data each day.[23]

Analysis populations

The primary analyses will be based on the Intention-to-Treat (ITT) population and includes participants with valid accelerometer measurement at baseline (defined as at least 3 days with each 8 hours of measurement). The ITT principle asserts that the effect of a treatment policy is best assessed by evaluating participants based on the treatment they were originally assigned to receive, rather than the treatment they actually received. This means that participants will be followed, assessed, and analyzed according to their assigned treatment group, regardless of adherence to the planned treatment or initiation of other interventions.[24,25]

Treatment policy estimand: The main analysis population (i.e. the ITT population) represents the treatment policy estimand, corresponding to the average treatment effect among all randomly assigned participants. This will be used for the main inferential measures.

Hypothetical estimand: Secondary analyses will report on the trial product estimand, representing the hypothetical estimand. That is like asking an additional clinical question of interest is what is the average treatment effect of THA relative to PRT, in the subgroup of randomized patients, measured by change from baseline to 6 months after initiating the treatment in number of steps per day, number of sit-to-stand transfers per day, and number of short walking bouts per day, had they remained on their randomized treatment for the entire planned duration of the trial, not initiated other competing interventions: declined surgery and initiated exercise on their own in the THA group and THA in the PRT group, and had they additionally complied with the protocol.

TRIAL POPULATION

Screening data/Eligibility/Recruitment

Screening and recruitment processes were conducted in accordance with the PROHIP trial's eligibility criteria to ensure that participants met all required conditions for inclusion.[12,19,26] Screening data were collected for all potential participants, including demographic, clinical, and radiographic imaging information, to assess eligibility. Participants were deemed eligible if they fulfilled the criteria outlined

in the trial protocol, including specific inclusion and exclusion factors. The detailed in and exclusion criteria are listed in the trial protocol.[19]

Recruitment proceeded according to the pre-established strategy, with participants being invited to enroll once their eligibility had been confirmed.[12,19,26] The recruitment period was monitored to ensure an adequate sample size was achieved within the specified timeline. Screening and recruitment data were summarized descriptively, with the number of individuals screened, the number eligible, and the final number enrolled in the study. Any reasons for non-eligibility or non-participation were documented and reported. We will report the participant flow in CONSORT Flow Diagram (see **Figure 1 Mockup**), in which the total number of participants screened and the number of eligible participants will be reported, as well as the number of participants enrolled in the study will be shown. Reasons for non-eligibility or non-participation (e.g., failure to meet inclusion criteria, withdrawal of consent) will be indicated as part of the flowchart.

Baseline patient characteristics

The baseline characteristics of the randomized participants (both groups) will be presented in a Table 1 (see **Table 1 Mockup**). These characteristics will be reported as descriptive statistics only, including means and standard deviations for continuous variables that are apparently normally distributed, and medians with interquartile ranges for those that are not normally distributed. Counts and percentages are provided for categorical variables. No statistical testing will be performed to compare the groups at baseline, as the purpose of this table is solely to describe the demographic and clinical profile of the participants in each group. This descriptive summary should provide an overview of the participants' characteristics before exposure to the allocated intervention and ensures transparency regarding the initial comparability of the groups.

STATISTICAL ANALYSIS

Outcome definitions and Endpoints

The primary outcome is the change from baseline to 6 months in number of steps per day, assessed using tri-axial accelerometry (AX3 Axivity Ltd., Newcastle UK), with data being post-processed using a custom designed algorithm (MATLAB R2019b, Mathworks, Natick, MA, USA).[27,28] The primary endpoint is the between-group difference in change from baseline to 6 months in the number of steps per day; the estimated difference between groups will be reported with corresponding 95% confidence intervals.

Key secondary outcomes include the change from baseline to 6 months in number of sit-to-stand transfers per day and number of short walking bouts per day assessed using tri-axial accelerometry. Other secondary outcomes include the change from baseline to 6 months in average cadence (steps/min), percentage of time spent in very low-, low-, moderate- and high intensity activity, walking (hours per day), standing (hours per day), sedentary (hours per day), cycling (min per day), and cycle rotations (number per day) assessed using tri-axial accelerometry.[27,28]

Endpoint title	Time frame	Unit
Primary endpoint		
Change in number of steps	From baseline (week 0) to 6 months after treatment initiation (week 26)	Number per day
Secondary endpoints		
Change in number of sit-to-stand transfers	From baseline (week 0) to 6 months after treatment initiation (week 26)	Number per day
Change in number of short walking bouts	From baseline (week 0) to 6 months after treatment initiation (week 26)	Number per day
Other endpoints		
Change in average cadence	From baseline (week 0) to 6 months after treatment initiation (week 26)	Steps per minute
Change in time spent very low-intensity activity	From baseline (week 0) to 6 months after treatment initiation (week 26)	Percentage of per day
Change in time spent low-intensity activity	From baseline (week 0) to 6 months after treatment initiation (week 26)	Percentage per day
Change in time spent moderate-intensity activity	From baseline (week 0) to 6 months after treatment initiation (week 26)	Percentage per day
Change in time spent high-intensity activity	From baseline (week 0) to 6 months after treatment initiation (week 26)	Percentage per day
Change in time spent walking	From baseline (week 0) to 6 months after treatment initiation (week 26)	Hours per day
Change in time spent standing	From baseline (week 0) to 6 months after treatment initiation (week 26)	Hours per day
Change in time spent sedentary	From baseline (week 0) to 6 months after treatment initiation (week 26)	Hours per day
Change in time spent cycling	From baseline (week 0) to 6 months after treatment initiation (week 26)	Minutes per day
Change in number of cycling rotations	From baseline (week 0) to 6 months after treatment initiation (week 26)	Number per day

Analysis methods

All descriptive statistics and statistical analyses will be reported in accordance with the recommendations of the *'Enhancing the QUAlity and Transparency Of health Research'* (EQUATOR) network[29] and the CONSORT statement.[21] Visual inspection (QQ-plot, histograms, and scatterplots) of

the standardised residuals from the statistical model will be used to assess the assumption of normality and homogeneity of variances.

All results from statistical analyses will be accompanied by two-sided 95% CIs and corresponding P values (superiority defined as P<0·05 for the primary endpoint). [30] Between-group differences of continuous endpoints will be analyzed using an analysis of covariance model with randomized treatment (THA or PRT) and stratification groups (Vejle, Odense, Aarhus, or Næstved) as factors and the baseline endpoint value as a covariate to reduce random variation.[31] Missing data will be multiply imputed using data from participants in the same randomized treatment group, and results will be combined using Rubin's rules.[32]

Changes from baseline will be reported as least squares means with standard errors, and between-group differences will be presented as least-squares means with 95% confidence intervals (see **Table 2 Mockup**). A two-sided P value of less than 0.05 will be considered to indicate statistical significance for the key secondary endpoints. Trajectories in the number of steps per day from baseline to 6 months will be (see **Figure 2 Mockup**), as well as a scatter plot of individual changes in the number of steps per day from baseline to 6 months will be presented (see **Figure 3 Mockup**).

Finally, we will follow previously published procedures for blinded results interpretation of the statistical analysis.[33] In this regard, the results from the analysis will be presented to the author group with the two treatment groups coded as Group A and Group B followed by development of two written interpretations. The author group will sign a consensus statement comprising both interpretations prior to the unsealing of the randomization code.

Missing data

The intention-to-treat (ITT) principle mandates that all participants in a clinical trial be included in the analysis according to the groups to which they were randomized, regardless of deviations from the assigned treatment. This approach is crucial for minimizing bias, as excluding participants who deviate from their assigned treatment can introduce selection bias, potentially reintroducing confounders into the analysis. We will adhere to the strategy proposed by White et al. for handling incomplete activity measurement at follow-up in ITT analysis: (1) make every effort to follow up with all randomized participants, even if they withdraw from the allocated treatment; (2) conduct the primary analysis using all observed data, assuming a plausible missing data mechanism (i.e., Multiple Imputation technique, applicable assuming data is Missing At Random [MAR]); (3) perform sensitivity analyses to assess the robustness of results under alternative assumptions about the missing data (i.e., Complete Case analysis and the Hypothetical Estimand will be estimated for the primary and key secondary endpoints); and (4) ensure all randomized participants are accounted for, at least in the sensitivity analyses.[34] Missing data will be imputed 1000 times from retrieved patients of the same randomized treatment and the results will be combined using Rubin's rules.[32]

Additional analyses – incl. sensitivity analyses

As indicated above, sensitivity analyses will be performed to test the robustness of the main analysis, excluding patients who refused to undergo surgery despite being randomly assigned to total hip arthroplasty and patients who underwent total hip arthroplasty despite being randomly assigned to progressive resistance training.

Harms

Serious adverse events (SAEs) will not be reported, as they have been reported previously.[12]

Statistical software

All statistical analyses and calculations will be performed using and STATA BE version 18.5 (Statacorp, College Station, Texas, USA).

FIGURES AND TABLES

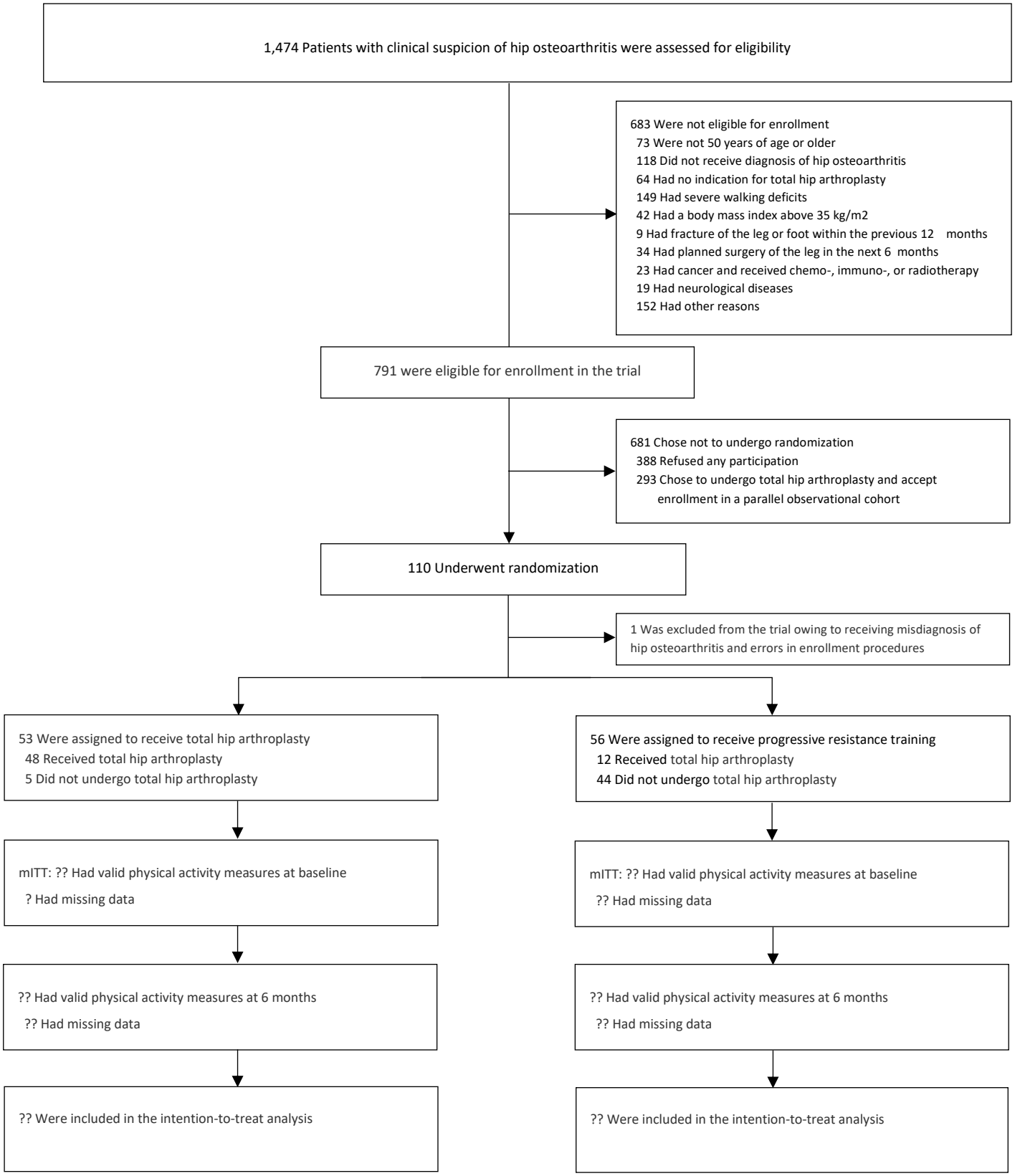


Figure 1 (Mockup). Participant Flow to 6 months Follow-up.

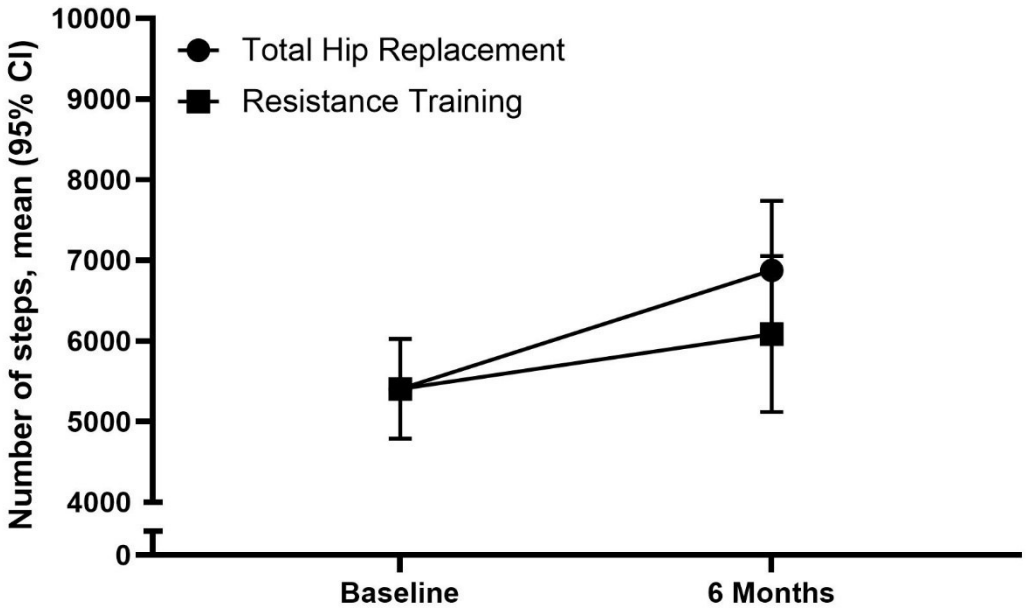


Figure 2 (Mockup). *Based on Simulated DATA*. Trajectories in of number of steps between the from baseline to 6 months in the two groups. The simulated data are based on values for the Danish background population.[35] Baseline values are based on the results for Danish persons 70 years of age or older, while the improvements are based on the results for Danish persons 60-69 years of age (Total Hip Arthroplasty) and 50-59 years of age (Progressive Resistance Training). The values plotted will be least-squares means, and I bars indicate 95% confidence intervals.

Figure 3 (Mockup). Scatter plot of individual changes in the number of steps per day from baseline to 6 months in the two groups. Horizontal bars indicate group medians.

Table 1 (Mockup). Baseline Characteristics of the Participants in the modified Intention-To-Treat Population*

Characteristic	Total Hip Arthroplasty (N=??)	Progressive Resistance Training (N=??)
Female sex, no. (%)		
Age, years		
Body mass index, kg/m2		
Duration of hip symptoms, years		
Previous total hip arthroplasty, no. (%)		
Previous total knee arthroplasty, no. (%)		
OHS, 0 to 48†		
UCLA activity score, 1 to 10‡		
40-m Fast Paced Walk Test, meters/sec		
30-sec Chair Stand Test, number of repetitions		
Physical activity variables		
Steps, number per day		
Sit-to-stand transfers, number per day		
Short walking bouts, number per day		
Average cadence, steps/min per day		
Very low-intensity activity, % per day		
Low-intensity activity, % per day		
Moderate-intensity activity, % per day		
High-intensity activity, % per day		
Walking, hours per day		
Standing, hours per day		
Sedentary, hours per day		
Cycling, min per day		
Cycle rotations, number per day		

* Values are presented as no. (%), mean (standard deviation) or median (interquartile range).
† The Oxford Hip Score (OHS) assess hip pain and function in one total score ranging from 0 (worst) to 48 (best).
‡ The University of California Los Angeles (UCLA) activity score measures physical activity level ranging from 1 (inactive) to 10 (regular physical activity with high intensity).

Table 2 (Mockup). Outcomes at 6 months in the modified Intention-To-Treat Population.*

	6 Months follow-up		Between-Group Difference	
Outcomes	Total Hip Arthroplasty (N=??)	Progressive Resistance Training (N=??)	Adjusted Mean Difference (95% CI)	P value
Primary Outcome				
Steps, number per day				
Key Secondary Outcomes				
Sit-to-stand transfers, number per day				
Short walking bouts, number per day				
Other Secondary Outcomes				
Average cadence, steps/min				N/A
Very low-intensity activity, % per day				N/A
Low-intensity activity, % per day				N/A
Moderate-intensity activity, % per day				N/A
High-intensity activity, % per day				N/A
Walking, hours per day				N/A
Standing, hours per day				N/A
Sedentary, hours per day				N/A
Cycling, min per day				N/A
Cycle rotations, number per day				N/A

* All analyses will be based on the modified Intention-To-Treat population. For continuous outcomes, an analysis of covariance with multiple imputations for missing data will be used to estimate least squares means with standard errors and differences between groups, along with 95% confidence intervals. Between-group differences of continuous endpoints will be analyzed using an analysis of covariance model with randomized treatment (THA or PRT) and stratification groups (Vejle, Odense, Aarhus, or Næstved) as factors and the baseline endpoint value as a covariate. The 95% confidence intervals will not be adjusted for multiplicity and should not be used in place of hypothesis testing.

Appendix Table 1 (Mockup). Outcomes at 6 months in the hypothetical estimand population.*

	6 Months follow-up		Between-Group Difference
Outcomes	Total Hip Arthroplasty (N=??)	Progressive Resistance Training (N=??)	Adjusted Mean Difference (95% CI)
Primary Outcome			
Steps, number per day			
Key Secondary Outcomes			
Sit-to-stand transfers, number per day			
Short walking bouts, number per day			

* All analyses will be based on the hypothetical estimand population. Between-group differences of continuous endpoints will be analyzed using an analysis of covariance model with randomized treatment (THA or PRT) and stratification groups (Vejle, Odense, Aarhus, or Næstved) as factors and the baseline endpoint value as a covariate. The 95% confidence intervals will not be adjusted for multiplicity and should not be used in place of hypothesis testing.

Appendix Table 2 (Mockup). Outcomes at 6 months in the Complete Case population.*

	6 Months follow-up		Between-Group Difference
Outcomes	Total Hip Arthroplasty (N=??)	Progressive Resistance Training (N=??)	Adjusted Mean Difference (95% CI)
Primary Outcome			
Steps, number per day			
Key Secondary Outcomes			
Sit-to-stand transfers, number per day			
Short walking bouts, number per day			

* All analyses will be based on the complete case population. Between-group differences of continuous endpoints will be analyzed using an analysis of covariance model with randomized treatment (THA or PRT) and stratification groups (Vejle, Odense, Aarhus, or Næstved) as factors and the baseline endpoint value as a covariate. The 95% confidence intervals will not be adjusted for multiplicity and should not be used in place of hypothesis testing.

Appendix Table 3 (Mockup). Outcomes at 6 months in the modified Intention-To-Treat population, using Non-responder imputation.*

	6 Months follow-up		Between-Group Difference
Outcomes	Total Hip Arthroplasty (N=5?)	Progressive Resistance Training (N=5?)	Adjusted Mean Difference (95% CI)
Primary Outcome			
Steps, number per day			
Key Secondary Outcomes			
Sit-to-stand transfers, number per day			
Short walking bouts, number per day			

* All analyses will be based on the modified Intention-To-Treat population using non-responder imputation. Between-group differences of continuous endpoints will be analyzed using an analysis of covariance model with randomized treatment (THA or PRT) and stratification groups (Vejle, Odense, Aarhus, or Næstved) as factors and the baseline endpoint value as a covariate. The 95% confidence intervals will not be adjusted for multiplicity and should not be used in place of hypothesis testing.

REFERENCES

1. GBD 2021 Osteoarthritis Collaborators. Global, regional, and national burden of osteoarthritis, 1990-2020 and projections to 2050: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet Rheumatol*. Sep 2023;5(9):e508-e522. doi:10.1016/s2665-9913(23)00163-7

2. Fu M, Zhou H, Li Y, Jin H, Liu X. Global, regional, and national burdens of hip osteoarthritis from 1990 to 2019: estimates from the 2019 Global Burden of Disease Study. *Arthritis Res Ther*. Jan 3 2022;24(1):8. doi:10.1186/s13075-021-02705-6

3. Long H, Liu Q, Yin H, et al. Prevalence Trends of Site-Specific Osteoarthritis From 1990 to 2019: Findings From the Global Burden of Disease Study 2019. *Arthritis Rheumatol*. Jul 2022;74(7):1172-1183. doi:10.1002/art.42089

4. Kloppenburg M, Namane M, Cicuttini F. Osteoarthritis. *Lancet*. Jan 4 2025;405(10472):71-85. doi:10.1016/S0140-6736(24)02322-5

5. Tang S, Zhang C, Oo WM, et al. Osteoarthritis. *Nat Rev Dis Primers*. Feb 13 2025;11(1):10. doi:10.1038/s41572-025-00594-6

6. Achakri H, Ben-Shlomo Y, Blom A, et al. National Joint Registry Annual Reports. *The National Joint Registry 20th Annual Report 2023*. National Joint Registry © National Joint Registry 2022.; 2022.

7. The Danish Hip Arthroplasty Register (DHR). National Annual Report 2023. 2024.

8. Ferguson RJ, Palmer AJ, Taylor A, Porter ML, Malchau H, Glyn-Jones S. Hip replacement. *Lancet*. Nov 3 2018;392(10158):1662-1671. doi:10.1016/s0140-6736(18)31777-x

9. Blom AW, Donovan RL, Beswick AD, Whitehouse MR, Kunutsor SK. Common elective orthopaedic procedures and their clinical effectiveness: umbrella review of level 1 evidence. *Bmj*. Jul 7 2021;374:n1511. doi:10.1136/bmj.n1511

10. Skou ST, Poulsen E, Bricca A, et al. Benefits and Harms of Interventions With Surgery Compared to Interventions Without Surgery for Musculoskeletal Conditions: A Systematic Review With Meta-analysis. *J Orthop Sports Phys Ther*. Jun 2022;52(6):312-344. doi:10.2519/jospt.2022.11075

11. Gibbs AJ, Gray B, Wallis JA, et al. Recommendations for the management of hip and knee osteoarthritis: A systematic review of clinical practice guidelines. *Osteoarthritis Cartilage*. Oct 2023;31(10):1280-1292. doi:10.1016/j.joca.2023.05.015

12. Frydendal T, Christensen R, Mechlenburg I, et al. Total Hip Replacement or Resistance Training for Severe Hip Osteoarthritis. *N Engl J Med*. Oct 31 2024;391(17):1610-1620. doi:10.1056/NEJMoa2400141

13. Young JJ, Zywiel MG, Skou ST, et al. Total Hip Arthroplasty Versus Education and Exercise: A Propensity-Matched Analysis of 266 Patients Who Have Hip Osteoarthritis. *J Arthroplasty*. Sep 2024;39(9S2):S261-S269. doi:10.1016/j.arth.2024.04.072

14. Arnold JB, Walters JL, Ferrar KE. Does Physical Activity Increase After Total Hip or Knee Arthroplasty for Osteoarthritis? A Systematic Review. *J Orthop Sports Phys Ther*. Jun 2016;46(6):431-42. doi:10.2519/jospt.2016.6449

15. Hammett T, Simonian A, Austin M, et al. Changes in Physical Activity After Total Hip or Knee Arthroplasty: A Systematic Review and Meta-Analysis of Six- and Twelve-Month Outcomes. *Arthritis Care Res (Hoboken)*. Jun 2018;70(6):892-901. doi:10.1002/acr.23415

16. Smith TO, Latham S, Maskrey V, Blyth A. Patients' perceptions of physical activity before and after joint replacement: a systematic review with meta-ethnographic analysis. *Postgrad Med J*. Sep 2015;91(1079):483-91. doi:10.1136/postgradmedj-2015-133507

17. Webster F, Perruccio AV, Jenkinson R, et al. Understanding why people do or do not engage in activities following total joint replacement: a longitudinal qualitative study. *Osteoarthritis Cartilage*. Jun 2015;23(6):860-7. doi:10.1016/j.joca.2015.02.013

18. Ahmadi MN, Rezende LFM, Ferrari G, Del Pozo Cruz B, Lee IM, Stamatakis E. Do the associations of daily steps with mortality and incident cardiovascular disease differ by sedentary time levels? A device-based cohort study. *Br J Sports Med*. Mar 8 2024;58(5):261-268. doi:10.1136/bjsports-2023-107221

19. Frydendal T, Christensen R, Mechlenburg I, Mikkelsen LR, Overgaard S, Ingwersen KG. Total hip arthroplasty versus progressive resistance training in patients with severe hip osteoarthritis: protocol for a multicentre, parallel-group, randomised controlled superiority trial. *BMJ Open*. Oct 22 2021;11(10):e051392. doi:10.1136/bmjopen-2021-051392

20. Gamble C, Krishan A, Stocken D, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. *Jama*. Dec 19 2017;318(23):2337-2343. doi:10.1001/jama.2017.18556

21. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *Bmj*. Mar 23 2010;340:c869. doi:10.1136/bmj.c869

22. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. Apr 2009;42(2):377-81. doi:10.1016/j.jbi.2008.08.010

23. Sliepen M, Lipperts M, Tjur M, Mechlenburg I. Use of accelerometer-based activity monitoring in orthopaedics: benefits, impact and practical considerations. *EFORT Open Rev*. 2020;4(12):678-685. doi:10.1302/2058-5241.4.180041

24. Detry MA, Lewis RJ. The intention-to-treat principle: how to assess the true effect of choosing a medical treatment. *Jama*. Jul 2 2014;312(1):85-6. doi:10.1001/jama.2014.7523

25. Little RJ, D'Agostino R, Cohen ML, et al. The prevention and treatment of missing data in clinical trials. *N Engl J Med*. Oct 4 2012;367(14):1355-60. doi:10.1056/NEJMsr1203730

26. Frydendal T, Thomsen KS, Mechlenburg I, et al. Patient and public involvement to inform the protocol of a clinical trial comparing total hip arthroplasty with exercise: an exploratory qualitative case study. *BMJ Open*. Apr 24 2023;13(4):e070866. doi:10.1136/bmjopen-2022-070866

27. Lipperts M, van Laarhoven S, Senden R, Heyligers I, Grimm B. Clinical validation of a body-fixed 3D accelerometer and algorithm for activity monitoring in orthopaedic patients. *J Orthop Translat*. Oct 2017;11:19-29. doi:10.1016/j.jot.2017.02.003

28. Sliepen M, Lipperts M, Tjur M, Mechlenburg I. Use of accelerometer-based activity monitoring in orthopaedics: benefits, impact and practical considerations. *EFORT Open Rev*. Dec 2019;4(12):678-685. doi:10.1302/2058-5241.4.180041

29. Christensen R, Bliddal H, Henriksen M. Enhancing the reporting and transparency of rheumatology research: a guide to reporting guidelines. *Arthritis Res Ther*. Feb 28 2013;15(1):109. doi:10.1186/ar4145

30. Christensen R, Ranstam J, Overgaard S, Wagner P. Guidelines for a structured manuscript: Statistical methods and reporting in biomedical research journals. *Acta Orthop*. May 10 2023;94:243-249. doi:10.2340/17453674.2023.11656

31. Vickers AJ, Altman DG. Statistics notes: Analysing controlled trials with baseline and follow up measurements. *Bmj*. Nov 10 2001;323(7321):1123-4. doi:10.1136/bmj.323.7321.1123

32. Little RJA, Rubin DB. *Statistical analysis with missing data*. 1st ed. John Wiley and Sons; 1987.

33. Jarvinen TL, Sihvonen R, Bhandari M, et al. Blinded interpretation of study results can feasibly and effectively diminish interpretation bias. *Journal of clinical epidemiology*. Jul 2014;67(7):769-72. doi:10.1016/j.jclinepi.2013.11.011

34. White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. *Bmj*. Feb 7 2011;342:d40. doi:10.1136/bmj.d40

35. Schrøder CK, Tønning LU, Tjur M, Kristensen PK, Mechlenburg I. Reference Values for Daily Physical Activity Measured with Accelerometers in a Danish Background Population between 18 and 80 Years of Age. *Applied Sciences*. 2023;13(3):1443. doi:10.3390/app13031443