Statistical analysis plan (SAP)

Self-management program for greater trochanteric pain syndrome: a randomized controlled trial

Section 1: Administrative information

SAP version and date:	Version 2, date 20.06.24, updated secondary outcomes (PASS)
Protocol version:	This document has been written based on information contained in the study protocol (3)
REC no:	2023/590816
Roles and responsibility:	Marianne Bakke Johnsen, PhD, PT, project leader, Department of Physical Medicine and Rehabilitation, Oslo University Hospital, Institute of Rehabilitation Science and Health Technology, Oslo Metropolitan University
	Are Hugo Pripp, Professor, statistician, Oslo Centre of Biostatistics and Epidemiology, Oslo University Hospital and Faculty of Health Sciences, OsloMet – Oslo Metropolitan University
	Cecilie Røe, Professor, MD, Head of Department of Physical Medicine and Rehabilitation, Oslo University Hospital
	John Andreas Bjørneboe, MD, PhD, supervisor, Department of Physical Medicine and Rehabilitation, Oslo University Hospital
	Thea Morin Melås, PT, PhD student, Department of Physical Medicine and Rehabilitation, Oslo University Hospital, Institute of Rehabilitation Science and Health Technology, Oslo Metropolitan University
	Maren Lunder Wefring, MD, PhD student, Department of Physical Medicine and Rehabilitation, Oslo University Hospital

Section 2: Introduction

Background and rationale

This document describes the planned statistical analyses to be performed for the clinical trial: "*Self-management program for greater trochanteric pain syndrome: a randomized controlled trial"*.

The aim of the project is to to assess the effectiveness of a self-management program versus usual care for patients with greater trochanteric pain syndrome (GTPS).

Objectives

1. To investigate the clinical effectiveness of a self- management program versus usual care for patients with GTPS*.
2. To investigate if self-management is more cost-effective than usual care in the treatment of patients with GTPS.

*H0: There is no difference between a self- management program and usual care on pain and function in patients with GTPS.
*H1: There is a difference between a self-management program and usual care on pain and function in patients with GTPS.

Section 3: Study methods

Study design

This study is a single-centre, parallel-group randomized controlled trial (RCT) with 6 months primary outcome endpoint and 12 months follow-up.

Randomisation

The randomization sequence is computer-generated with blocks of variable size, which is unknown to any of the research team. Participants will be randomly allocated into one of two groups: 1) self-management program at OUH or 2) usual care in primary sector. The randomisation information will be stored in a separate folder on the server for sensitive data at OUH, only accessible for the research assistant that will not be involved in the treatment of the study participants.

Sample size

Based on validation of the primary outcome, the Norwegian version of the VISA-G (score 0-100), we calculated the sample size in the current RCT using a standard deviation of 17 and a mean difference of 10 between groups. This calculation resulted in a sample size of 94 participants (i.e. 47 in each group), assuming a two-sample means t-test, 80% power, and an alpha of 0.05. To account for potential drop out (15%) we plan to recruit 110 participants (55 per group).

Statistical framework

Superiority hypothesis testing will be performed to test the effectiveness of the self-management program compared to usual care, according to the null hypothesis as stated above. Superiority of the self-management over usual care will be claimed if the two-sided *P* value in the test comparing the mean difference between groups at 6 months in VISA-G-N

score is less than 5 %. This protocol is designed to address a single primary endpoint; measure of lateral hip pain and disability using VISA-G-N. A difference in the effect of interventions will be claimed if the null hypothesis is rejected, that is if the two-sided *P* value is less than 5 %.

Statistical interim analysis and stop guidance

There will be no interim analyses in this trial.

If severe medical events should occur, the manager of the department have access to unblind that particular patient.

Timing of final analysis

The analysis timing is stratified by planned length of follow-up. The primary endpoint is at 6 months follow-up, where analysis of the primary outcome will be performed. The secondary endpoint is at 12 months follow-up, where secondary analysis will be performed.

Timing of outcome assessments

After the initial examination with a medical doctor at OUH, eligible patients who wish to participate in the study will be referred for screening with a physical examination by a doctor dedicated to the study. Participants who fulfil the inclusion and exclusion criteria, and consent to participate, will go on to a physiotherapist for general information and answer baseline questionnaires prior to randomisation. Outcomes are measured again 3, 6 and 12 months after randomisation (table 1).

Table 1 Timing of outcome measures in the study

Outcome measures	Baseline	3 months	6 months	12 months
Other assessments	X			
VISA-G-N (primary) (Victorian Institute of Sport Assessment Gluteal Questionnaire)	X	X	X	X
NRS (Numeric rating scale)	X	X	X	X
Painful sites	X	X	X	X
PSEQ (Pain self-efficacy questionnaire)	X	X	X	X
EQ-5D-5L (EuroQoL-5 dimensions-5 Level)	X	X	X	X
HSCL-10 (Hopkins Symptoms checklist-10)	X	X	X	X
Expectations	X	X	X	
iPCQ (iMTA Productivity Cost Questionnaire)	X	X	X	X
iMCQ (iMTA Medical Consumption Questionnaire)	X	X	X	X
Medicine consumption	X	X	X	X
GROC (Global rating of change)			X	
PASS (Patient Acceptable Symptom State)			X	

Section 4: Statistical principles

Confidence intervals and *P* values

All statistical tests will be two-sided and will be performed using a 5% significance level. Any confidence intervals presented will be 95% and two-sided.

Adherence and protocol deviations

Definition of adherence to the intervention (self-management program): For the **self-management program**, the patients must have attended at least 2 out of 3 sessions.

- % compliance of self-management sessions = (number of sessions/3 sessions) *100%

Adherence to the intervention will be presented as descriptive statistics indicating the percentage compliance (N, mean, median, minimum, maximum), and will be summarized by randomization group.

Definition of major protocol deviation in the intervention group is not meeting to any self-management sessions. Description of major protocol deviations will be presented as descriptive statistics indicating the percentage compliance (N, mean, median, minimum, maximum), and will be summarized by randomization group.

Analysis population

We will define the following patient population in this trial:

- Intention to treat population: will include all randomized participants, according to the treatment they were randomized to receive, regardless of what treatment (or not) they received.
- Per protocol population: will include all patients that were randomized and received treatment according to protocol, without major deviations (e.g. for the intervention group 0 sessions with the physiotherapist).

Section 5: Trial populations

Screening data

All patients at the outpatient clinic with lateral hip pain (diagnosis M76.0/M70.6) will be screened for eligibility.

Eligibility

The number of participants assessed for eligibility and those excluded due to violation of inclusion/exclusion criteria will be presented in a flow diagram[1].

Recruitment & Withdrawal/follow-up

The CONSORT diagram will comprise the number of people screened, eligible, consented, randomized, receiving their allocated treatment and the numbers withdrawing or lost to follow-up at each step.

Baseline patient characteristics

Baseline characteristics will be presented by trial arm. Continuous data will be summarized by mean, SD and range if data are normal distributed and median, IQR and range if data are skewed. Categorical data will be summarized by numbers and percentages. Tests of statistical significance will not be undertaken for baseline characteristics; rather the clinical importance of any imbalance will be noted. For each variable (continuous or categorical), the percentage of missing values will be reported.

Section 6: Analysis

Outcome definitions

Primary outcome

The primary outcome is a diagnose-specific questionnaire that consists of eight questions on pain and function (VISA-G-Norwegian)[2]. The responses are summarized, so that the total score ranges from 0-100, with higher scores representing less pain and disability[2].

Secondary outcomes

Pain intensity measured on **Numeric rating scale** (NRS). The patients are asked to rate the average severity of pain (at night, rest, in activity) over the last week on a scale ranging from 0 (no pain) to10 (worst possible pain)[3].

The patients are asked to depict the number of **painful sites** during the last 14 days, divided into18 anatomical regions. The number of painful sites is then added up to give a total score.

The Pain Self-Efficacy Questionnaire (PSEQ) assesses the confidence people with ongoing pain have in performing activities while in pain[4]. The questionnaire includes 10-items, where patients rate their confidence from 0 points (not at all confident) to 6 points (completely confident). Total scores are calculated by summing the individual items with a range from 0 points (less self-efficacy) to 60 points (more self-efficacy).

EQ-5D-5L (EuroQoL-5 dimensions-5 Level) is a validated generic health related quality of life measure[5]. The first part is based on 5 questions which are answered on a 5-step scale. It gives an index between 0-1, with a higher index representing better quality of life. The second part, EQ-VAS is a vertical line in which the participants mark the point best describing their quality of life. The score ranges from 0-100, where 100 represents best and 0 worst imaginable health states.

The **Hopkins Symptoms checklist-10** (HSCL-10) is a short version of the Hopkins Checklist-25 that has been developed to evaluate psychological distress[6]. The HSCL-10 asks the respondent about symptoms related to anxiety and depression on a scale of 1 (not at all) to 4 (extremely). The mean score is calculated producing a range of scores from 1-4 where higher score corresponds to more psychological distress. An average score $\geq$1.85 has commonly been considered a cut-off to identify cases.

Expectations will be measured at baseline through expected improvement in VISA-G score at 6 months. In addition, they will be asked about expected improvement in function and pain on NRS.

The **iMTA Medical Consumption Questionnaire** (iMCQ), which is an instrument for measuring medical consumption [7], will be used to measure healthcare utilization.

In addition, **medicine consumption** the last month before each follow-up will be registered and analysed according to defined daily doses using Anatomical Therapeutic Classification (ATC) codes[8].

For the measurement of productivity loss (absenteeism) we will use the iMTA **Productivity Cost Questionnaire** (iPCQ)[9].**Global Rating of Change** (GROC) will be used to measure the patients perceived change in lateral hip pain from baseline to follow-up at 6 months on an 11-point likert scale ranging from -5 (much worse) to 5 (completely recovered)[10].

Patient Acceptable Symptom State (PASS) will be used to measure patient satisfaction with current symptom state, by asking the question: "Taking into account all the activities you have during your daily life, your level of hip pain, and also your functional impairment, do you consider that your current state is satisfactory?" The response options are "yes" or "no" [11].

For timing of outcome measurements see table 1.

Analysis methods

The primary analysis will compare the two intervention groups (self-management vs usual care) on their mean difference in pain and disability (VISA-G-N score) at baseline, 3 and 6 months. The estimated mean difference between groups at 6 months (main endpoint) will be analyzed using a longitudinal mixed effects model analysis of covariance[12]. The model will include fixed effects of time, intervention, the interaction between time and intervention and the outcome variable at baseline as a covariate with a subject-specific random intercept. The primary effect analysis will use the intention-to-treat population.

Secondary outcomes assessed at multiple time points (baseline, 3, 6 and 12 months) will be analyzed by the same approach as described for the primary outcome, on both the intention-to-treat population and the per protocol population.

Baseline VISA-G-N score (primary outcome) will be included as a covariate in all the analyses.

Missing data

The degree of missingness will be explored. We will perform sensitivity analysis with multiple imputation and/or the use of an alternative repeated measures mixed model analysis accordingly, to assess the robustness of data[12].

Additional analyses

Multivariable logistic and linear regression analysis will be used to explore predictive factors, such as demographics, expectations, clinical findings, and psychosocial factors for primary and secondary outcomes. Model building will be done in a way that is appropriate for the given sample sizes, by restricting the number of potential predictive factors and considering shrinkage methods to stabilize predictions [13].

Cost-effectiveness analyses will be conducted as differences between groups in Quality Adjusted Life-Year (QALY) and costs. Differences between groups in QALYs will be describes with means (95% CI) and evaluated with ANCOVA to adjust for baseline scores unequally distributed across the groups at baseline. Costs of healthcare utilization per patient will be estimated by multiplying frequency of use by unit costs collected from national pricelists. Costs of productivity loss per patient will be estimated by multiplying number of workdays with complete productivity loss by an estimated average wage rate including taxes and social costs. Difference in costs will be evaluated with Student's t-test. Cost-effectiveness will be estimated with mean incremental cost-effectiveness ratios (ICERs). ICERs are calculated by dividing mean differences in costs by mean differences in effects. To adjust for uncertainty in ICER calculation, we will conduct bootstrapping (10 000 replicated datasets) of cost, and effect pairs will be plotted on cost-effectiveness planes (CE-planes) and cost-effectiveness acceptability curves (CEACs).

Mediation analysis will be performed to explore the causal pathway between treatment allocation and the primary outcome of pain and disability by considering, amongst others, pain self-efficacy and emotional distress as potential mediators that may be part of the causal pathway between intervention and outcome[14].

Harms

Any complications or adverse events will be continuously registered.

Statistical software

All statistical analyses will be performed using the IBM SPSS, STATA or other appropriate software.

References

1. Gamble C, Krishan A, Stocken D, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. Jama 2017;318(23):2337-2343. (In eng). DOI: 10.1001/jama.2017.18556.
2. Sveinall H, Wenstad PK, Fearon AM, et al. The reliability and validity of the Norwegian version of the Victorian Institute of Sports Assessment for gluteal tendinopathy questionnaire (VISA-G-Norwegian) for patients with greater trochanteric pain syndrome. BMC Musculoskelet Disord 2023;24(1):769. (In eng). DOI: 10.1186/s12891-023-06901-0.
3. Ferreira-Valente MA, Pais-Ribeiro JL, Jensen MP. Validity of four pain intensity rating scales. Pain 2011;152(10):2399-2404. (In eng). DOI: 10.1016/j.pain.2011.07.005.
4. Nicholas MK. The pain self-efficacy questionnaire: Taking pain into account. Eur J Pain 2007;11(2):153-63. (In eng). DOI: 10.1016/j.ejpain.2005.12.008.
5. Devlin NJ, Shah KK, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: An EQ-5D-5L value set for England. Health Econ 2018;27(1):7-22. (In eng). DOI: 10.1002/hec.3564.
6. Strand BH, Dalgard OS, Tambs K, Rognerud M. Measuring the mental health status of the Norwegian population: a comparison of the instruments SCL-25, SCL-10, SCL-5 and MHI-5 (SF-36). Nord J Psychiatry 2003;57(2):113-8. (In eng). DOI: 10.1080/08039480310000932.
7. iMTA EUR. iMTA Productivity and Health Research Group. Manual iMTA Medical Cost Questionnaire (iMCQ). 2018.
8. Brox JI, Staff PH, Ljunggren AE, Brevik JI. Arthroscopic surgery compared with supervised exercises in patients with rotator cuff disease (stage II impingement syndrome). BMJ 1993;307(6909):899-903. (http://www.ncbi.nlm.nih.gov/pubmed/8241852).
9. Bouwmans C, Krol M, Severens H, Koopmanschap M, Brouwer W, Hakkaart-van Roijen L. The iMTA Productivity Cost Questionnaire: A Standardized Instrument for Measuring and Valuing Health-Related Productivity Losses. Value Health 2015;18(6):753-8. DOI: 10.1016/j.jval.2015.05.009.
10. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. J Man Manip Ther 2009;17(3):163-70. (In eng). DOI: 10.1179/jmt.2009.17.3.163.
11. Kvien TK, Heiberg T, Hagen KB. Minimal clinically important improvement/difference (MCII/MCID) and patient acceptable symptom state (PASS): what do these concepts mean? Ann Rheum Dis 2007;66 Suppl 3(Suppl 3):iii40-1. (In eng). DOI: 10.1136/ard.2007.079798.
12. Twisk JWR. Analysis of Data from Randomized Controlled Trials. A Practical Guide: Springer; 2021:1-167.
13. Hingorani AD, Windt DA, Riley RD, et al. Prognosis research strategy (PROGRESS) 4: stratified medicine research. Bmj 2013;346:e5793. (In eng). DOI: 10.1136/bmj.e5793.

14. Lee H, Cashin AG, Lamb SE, et al. A Guideline for Reporting Mediation Analyses of Randomized Trials and Observational Studies: The AGReMA Statement. Jama 2021;326(11):1045-1056. (In eng). DOI: 10.1001/jama.2021.14075.