BMJ Open
Progressive resistance training compared to neuromuscular exercise and the additive effect of exercise booster sessions: protocol for a multicentre cluster randomised controlled trial (The Hip Booster Trial)

Troels Kjeldsen 1,2, Ulrik Dalgas 3, Søren T Skou 2,4, Maurits van Tulder 5,6, Bo M Bibby 7, Inger Mechlenburg 1,8

ABSTRACT
Introduction The primary aim of this randomised controlled trial is to investigate the effectiveness of 3 months of progressive resistance training (PRT) compared to neuromuscular exercise (NEMEX) on functional performance in patients with hip osteoarthritis (OA). Secondary aims are to investigate the effectiveness of exercise booster sessions (EBS) in prolonging the effects of the initial exercise interventions as well as to investigate the cost-effectiveness of PRT, NEMEX and EBS at 12-month follow-up.

Methods and analysis This multicentre cluster randomised controlled trial will be conducted at hospitals and physiotherapy clinics across Denmark. A total of 160 participants with clinically diagnosed hip OA will be recruited. Participants will be cluster randomised to a 3-month intervention of either PRT or NEMEX and to receive EBS or not, resulting in four treatment arms.

The primary outcome is change in functional performance, measured by the 30-s chair stand test at 3 months for the primary comparison and at 12 months for the EBS comparisons. Secondary outcomes include changes in 40 m fast-paced walk test, 9-step timed stair climb test, leg extensor muscle power and maximal strength, Hip Disability and Osteoarthritis Outcome Score subscales, EuroQol Group 5-dimension, global perceived effect, physical activity and pain. Outcomes are measured at baseline, after the initial 3 months of intervention, and at 6-month, 9-month and 12-month follow-up. An intention-to-treat approach will be used for analysing changes in the primary and secondary outcome measures.

Ethics and dissemination The trial has been approved by the Central Denmark Region Committee on Biomedical Research Ethics (Journal No 1-10-72-267-20) and registered at the Danish Data Protection Agency (Journal No 1-16-02-11-21). Results will be published in international peer-reviewed scientific journals.

Trial registration number NCT04714047.

STRENGTHS AND LIMITATIONS OF THIS STUDY
⇒ This is a cluster randomised, controlled, assessor-blinded, parallel-group trial providing high-quality evidence on exercise therapy for hip osteoarthritis.
⇒ This pragmatic multicentre trial is carried out at several hospitals and physiotherapy clinics across Denmark increasing the feasibility of subsequent implementation of the interventions.
⇒ The primary outcome is an objective test of functional performance and is less susceptible to performance bias compared with subjective measures.
⇒ The physiotherapists and patients will not be blinded to treatment allocation, which may introduce bias.
⇒ We are not able to distinguish between the specific treatment effects and the contextual effects of the interventions.

INTRODUCTION
Osteoarthritis (OA) is the most common joint disease and a leading cause of disability worldwide. OA is becoming more prevalent in the ageing and increasingly obese population, placing a substantial burden on healthcare systems and causing large societal costs. Exercise is a safe, feasible and effective treatment for patients with hip OA, reducing pain and improving physical function, and is strongly recommended by professional societies. Content and dosage of exercise interventions seem to be important for the magnitude of the effects. However, different exercise modalities in hip OA have only sparsely been compared9 and exercise recommendations suggesting one type of exercise over another are not based on solid...
EBS is cost-effective compared with no EBS; (2) PRT is cost-effective compared with NEMEX regardless of allocation to EBS or no EBS; (3) PRT with EBS is cost-effective compared with PRT without EBS and NEMEX with EBS is cost-effective compared with NEMEX without EBS; (4) PRT followed by EBS is superior to PRT without EBS, and that NEMEX followed by EBS is superior to NEMEX without EBS in improving functional performance, measured by the 30 s CST and (5) PRT is superior to NEMEX in improving functional performance, measured by the 30 s CST regardless of allocation to EBS.

**METHODS AND ANALYSIS**

**Study design**

This multicentre, cluster randomised, controlled, parallel-group, assessor-blinded, superiority trial will be conducted at 5 hospitals and 10 physiotherapy clinics across Denmark. Following the baseline test, participants will be cluster randomised to a 3-month intervention of either PRT or NEMEX and additionally to receive EBS or not, resulting in four treatment arms (see figure 1). EBS will be provided at 1, 3, 5 and 7 months after conclusion of the initial intervention. After baseline testing, the intervention is started as soon as possible. The primary outcome is change in functional performance, measured by the 30 s CST, and the primary endpoint is at 3 months after starting the intervention for the comparison between PRT and NEMEX and at 12 months for the comparison of EBS and no EBS. Secondary outcomes will be measured at baseline, and at 3-month, 6-month, 9-month and 12-month follow-up.

This protocol is written in accordance with the SPIRIT guideline and the interventions are described according to the CERT guidelines and the mechanobiological determinants suggested by Toigo and Boutellier. The results of this trial will be reported following the CONSORT statement guidelines.

Participant enrolment started on January 2021 and is expected to be completed by October 2022. All participants are expected to have completed 12-month follow-up assessments by December 2023, marking the end of the study.

**Participants**

Participants will be recruited from 4 of the 5 hospitals and 10 physiotherapy clinics across three healthcare regions of Denmark. Participants are initially screened and given participant information verbally and in writing. After 24 hours, they are contacted by an assessor, with an invitation to baseline assessment. Before the baseline assessment participants give written informed consent.

**Inclusion criteria**

(1) Clinically diagnosed OA of the hip joint according to the National Institute for Health and Care Excellence criteria; (2) An event of pain during activity of at least 3 out of 10 on a Numerical Rating Scale (NRS) in the index...
hip within the last 2 weeks; (3) Age ≥ 45 years; (4) No hip joint morning stiffness or less than 30 min; (5) No surgery in the lower extremities 6 months prior to inclusion; (6) No comorbidity that markedly affects hip function; (7) Adequacy in written and spoken Danish and (8) Not being a candidate for total hip arthroplasty.

Exclusion criteria
(1) Body mass index score > 40; (2) Pregnancy; (3) PRT or NEMEX for the lower extremities exceeding 12 sessions over the last 6 months or 6 sessions over the last 3 months and (4) Planned vacation for more than 14 days within the initial 3-month intervention period without the possibility of prolonging the intervention accordingly.

Randomisation and treatment allocation
After recruitment and baseline assessment, participants will be randomised to either PRT, NEMEX, PRT+B or NEMEX+B by cluster randomisation stratified by recruitment site according to a randomly generated sequence of numbers. A member of the research team (IM) who is not involved in recruitment, assessment or treatment, generated the allocation sequence for each of the 14 sites by drawing tokens from a bag containing an even distribution of the four allocations. The sequence will be concealed to the physiotherapists who enrol participants. The cluster size is set at five participants and each cluster forms an exercise group that will attend group sessions together. However, to keep the waiting time at an acceptable level, groups of 1–4 participants are cluster randomised if they have waited > 14 days after inclusion. Cluster randomisation was chosen to allow timely formation of the exercise groups and to avoid treatment contamination between interventions.

Blinding
Outcome assessors will conduct baseline and follow-up assessment blinded to group allocation. Prior to the 3-month and 12-month assessment, participants will be instructed not to disclose the allocated treatment. The participants and the physiotherapists delivering the exercise interventions will not be blinded to group allocation.

Interventions
The exercise interventions will be performed at the collaborating hospitals and physiotherapy clinics. All sessions will be conducted as group sessions with one physiotherapist supervising the exercises. Each cluster...
will be supervised once by the principal investigator, to further ensure fidelity to the protocol. The duration and frequency of the interventions will be 3 months with two supervised sessions each week interspersed by 72 hours. This results in 24 sessions, each lasting 60 min. Each session consists of a 10 min submaximal warm-up on an exercise bike at an intensity of 13–14 on Borgs Rating of Perceived Exertion scale,34 followed by 50 min of PR T or NEMEX. If participants experience pain during exercise exceeding 5 out of 10 on an NRS, the physiotherapist will modify the exercise, decreasing the exercise intensity (load), modifying the range of motion or changing the tempo of the exercise. All unilateral exercises will be performed for both legs. During the exercise interventions, the physiotherapists informed the patients on pain flares, pain management and that exercising with OA is safe. There are no restrictions for concomitant care.

Neuromuscular exercise

The NEMEX intervention (described in table 1 and figure 2) will follow the programme as described by Ageberg et al42 with the exception that the 10 min cool-down is left out. Ten exercises are performed each session and progression is primarily provided by four levels of difficulty and second by varying the number of, direction, and velocity of the movements and/or changing the support surface.12 35 At the first session, the physiotherapist assesses each participant to determine at which level the participant will start. The exercises are tailored to the individual to improve movement quality, reduce pain or increase the difficulty of the movement.

### Table 1 Descriptors of the progressive resistance training and neuromuscular exercise interventions

<table>
<thead>
<tr>
<th>Exercise type</th>
<th>Progressive resistance training</th>
<th>Neuromuscular exercise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Periodisation model</td>
<td>Block 1 (Weeks 1–4)</td>
<td>Block 2 (Weeks 5–8)</td>
</tr>
<tr>
<td>Volume (rep/set)</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>Exercise intensity</td>
<td>12 RM</td>
<td>10 RM</td>
</tr>
<tr>
<td>Sets/exercise (n)</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Muscle contraction types</td>
<td>Concentric</td>
<td>As fast as possible</td>
</tr>
<tr>
<td>Isometric</td>
<td>1 s</td>
<td>1 s</td>
</tr>
<tr>
<td>Eccentric</td>
<td>3 s</td>
<td>3 s</td>
</tr>
<tr>
<td>Time under tension (s/rep)</td>
<td>4–6</td>
<td>Not controlled for</td>
</tr>
<tr>
<td>Time between repetitions (s)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Time between sets (s)</td>
<td>60</td>
<td>Equivalent to completing one set</td>
</tr>
<tr>
<td>Session duration (m)</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>Sessions/week (n)</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Time between sessions (h)</td>
<td>&gt;72</td>
<td>&gt;72</td>
</tr>
<tr>
<td>Supervised and group based</td>
<td>Initial 12 weeks</td>
<td>Initial 12 weeks</td>
</tr>
<tr>
<td>Focus for exercises</td>
<td>Maximal intensity (exercise weight) and volume without compromising technique</td>
<td>Stability, postural function, postural orientation, lower extremity muscle strength, functional exercises</td>
</tr>
<tr>
<td>Range of motion</td>
<td>Full</td>
<td>Full</td>
</tr>
<tr>
<td>Volitional muscle failure</td>
<td>Third set only</td>
<td>Not controlled for</td>
</tr>
<tr>
<td>Order of adjusting exercises in case of pain exacerbation</td>
<td>Pace → range of motion → intensity</td>
<td>Pace → range of motion → no of repetitions → difficulty level</td>
</tr>
<tr>
<td>Progression</td>
<td>When able to perform all assigned repetitions in the third set, the weight is increased by 2%–10%</td>
<td>When able to perform 15 repetitions for three sets with good sensorimotor control, movement quality and acceptable exertion, progression is made to the next level of difficulty.</td>
</tr>
<tr>
<td>Equipment</td>
<td>Leg press machine, knee extension machine, hyperextension bench, dumbbells, cable pulley, ankle straps</td>
<td>Aerobic stepper, pilates exercise ball, elastic bands, sliding mat, chair with armrest, foam balance pad</td>
</tr>
</tbody>
</table>

h, hours; m, minutes; n, number; rep, repetitions; RM, repetition maximum; ROM, range of motion; s, seconds; w, weeks.
Progressive resistance training

Unilateral leg press

Hyperextensions

Unilateral hip extension (alternative to hyperextensions)

Unilateral hip abduction

Unilateral hip flexion

Bilateral knee extensions

Neuromuscular exercise (difficulty 4)

Pelvic lift

Sit ups

Slide-exercise forw-backw

Slide-exercise sideways

Hip abductors

Hip adductors

Knee extensors

Knee flexors

Chair stands

Stair climbing

Figure 2  Exercises at starting position (left) and at end-range of motion (right) for the two exercise interventions. Only difficulty level 4 is shown for the neuromuscular exercises while level 1–3 can be found as an additional file to the article by Ageberg et al.12

exercise. In every third set, participants are instructed to continue performing the exercise until volitional muscular failure. If a participant reaches the RM target repetitions or more for the third set, the exercise intensity (weight) will be increased by 2%-10% depending on the level of exertion.

Self-administered exercise and EBS

After the initial 3-month interventions, participants in all four groups are given a training programme and encouraged to continue performing the same exercises independently twice weekly through the 9-month follow-up period, starting at the level they reached during the initial
intervention. The NEMEX and NEMEX+B groups will be provided with the equipment for NEMEX in their home (one aerobic stepper, one pilates exercise ball and two elastic bands). The PRT and PRT+B groups will be offered a membership to a training facility. The facility staff will be told not to give any supervision to the study participants. The EBS groups (PRT+B and NEMEX+B) will receive four EBS provided at 1, 3, 5 and 7 months after the initial interventions. At each EBS, the participants are performing an as-usual group-based training session while the physiotherapist reviews the self-administered programmes following a standardised screening of potential complications and provides recommendations for modification of the programmes.

Outcomes measures
All outcome measures will be assessed at the participating hospital departments at baseline, within 1 week after the end of the 3-month group-based intervention, and at 12-month follow-up (see table 2). The EuroQol Group 5-dimensional (EQ-5D-5L), iMTA Productivity Costs Questionnaire (IPCQ) and Health Utilisation Questionnaire (HUQ) will additionally be assessed at 6-month and 9-month follow-up. The outcome assessors are physiotherapists, who are trained in performing the assessments according to a standardised protocol.

Participant characteristics
The following information will be obtained from the participants at baseline: gender, age, height, weight (also at three and 12 months), civil status, educational level, employment status, substance use (alcohol and smoking), duration of symptoms, index hip, previous treatment, pain medication, joint replacements (also at 12 months), and other diseases.

Primary outcome
Thirty-second chair stand test
The 30 s CST is a valid and responsive measure with excellent reliability evaluating sit-to-stand function (number of repetitions).38–40 The 30 s CST was chosen as the primary outcome because it is an objective measure of functional performance (low risk of performance bias) that is easily standardised between test locations while sit-to-stand function is affected in people with hip OA negatively impacting activities of daily living.41 Identical chairs (44 cm seat height, no armrest) are provided for each test location.

Key secondary outcomes
Hip disability and Osteoarthritis Outcomes Score
The Hip disability and Osteoarthritis Outcomes Score (HOOS) is a 40-item patient-reported questionnaire consisting of five subscales. Each subscale gives a score ranging from 0 (worst) to 100 (best).44 HOOS is a valid, reliable and responsive measure in patients with hip OA.45 The pain and hip-related quality of life subscales are chosen as key secondary outcomes and will help guide the conclusion.

Table 2  Outcome measures and corresponding instruments and time points for data collection

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Instrument</th>
<th>Time points</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Functional tests</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sit-to-stand function (PO)</td>
<td>30 s chair stand test</td>
<td>B, 3 and 12 months</td>
</tr>
<tr>
<td>Maximum walking speed</td>
<td>40 m fast-paced walk test</td>
<td>B, 3 and 12 months</td>
</tr>
<tr>
<td>Stair negotiation</td>
<td>9-step timed stair climb test</td>
<td>B, 3 and 12 months</td>
</tr>
<tr>
<td>Leg extensor muscle power</td>
<td>Nottingham Power Rig</td>
<td>B, 3 and 12 months</td>
</tr>
<tr>
<td>Leg extensor muscle strength</td>
<td>Unilateral 1RM leg press test</td>
<td>B, 3 and 12 months</td>
</tr>
<tr>
<td><strong>Patient-reported outcomes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health-related QoL</td>
<td>Questionnaire (EQ-5D-5L)</td>
<td>B, 3, 6, 9 and 12 months</td>
</tr>
<tr>
<td>Health utilisation</td>
<td>Questionnaire (HUQ)</td>
<td>B, 3, 6, 9 and 12 months</td>
</tr>
<tr>
<td>Productivity losses</td>
<td>Questionnaire (IPCQ)</td>
<td>3, 6, 9 and 12 months</td>
</tr>
<tr>
<td>Symptoms</td>
<td>Questionnaire (HOOS)</td>
<td>B, 3 and 12 months</td>
</tr>
<tr>
<td>Pain</td>
<td>Questionnaire (HOOS, NRS)</td>
<td>B, 3 and 12 months</td>
</tr>
<tr>
<td>ADL function</td>
<td>Questionnaire (HOOS)</td>
<td>B, 3 and 12 months</td>
</tr>
<tr>
<td>Sport/recreation</td>
<td>Questionnaire (HOOS)</td>
<td>B, 3 and 12 months</td>
</tr>
<tr>
<td>Hip-related QoL</td>
<td>Questionnaire (HOOS)</td>
<td>B, 3 and 12 months</td>
</tr>
<tr>
<td>Global perceived effect</td>
<td>Questionnaire (GPE)</td>
<td>3 and 12 months</td>
</tr>
<tr>
<td>Physical activity</td>
<td>Questionnaire (SNBHw)</td>
<td>B, 3 and 12 months</td>
</tr>
<tr>
<td>THA (yes/no)</td>
<td>Questionnaire</td>
<td>B and 12 months</td>
</tr>
<tr>
<td><strong>Patient characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>Questionnaire</td>
<td>B</td>
</tr>
<tr>
<td>Age</td>
<td>Questionnaire</td>
<td>B</td>
</tr>
<tr>
<td>Height</td>
<td>Questionnaire</td>
<td>B</td>
</tr>
<tr>
<td>Weight</td>
<td>Questionnaire</td>
<td>B, 3 and 12 months</td>
</tr>
<tr>
<td>Civil status</td>
<td>Questionnaire</td>
<td>B</td>
</tr>
<tr>
<td>Educational level</td>
<td>Questionnaire</td>
<td>B</td>
</tr>
<tr>
<td>Employment status</td>
<td>Questionnaire</td>
<td>B</td>
</tr>
<tr>
<td>Substance use</td>
<td>Questionnaire</td>
<td>B</td>
</tr>
<tr>
<td>Duration of symptoms</td>
<td>Questionnaire</td>
<td>B</td>
</tr>
<tr>
<td>Index hip</td>
<td>Questionnaire</td>
<td>B</td>
</tr>
<tr>
<td>Previous treatment</td>
<td>Questionnaire</td>
<td>B</td>
</tr>
<tr>
<td>Pain medication</td>
<td>Questionnaire</td>
<td>B</td>
</tr>
<tr>
<td>Other diseases</td>
<td>Questionnaire</td>
<td>B</td>
</tr>
<tr>
<td>Treatment-related variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adverse events</td>
<td>Supervision and questionnaire</td>
<td>Throughout study period</td>
</tr>
</tbody>
</table>

Continued
Table 2  Continued

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Instrument</th>
<th>Time points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adherence to interventions</td>
<td>Supervision</td>
<td>Training sessions</td>
</tr>
<tr>
<td>Pain during exercise</td>
<td>NRS</td>
<td>Before and after each training session</td>
</tr>
<tr>
<td>Drop-outs</td>
<td>Supervision</td>
<td>Throughout study period</td>
</tr>
<tr>
<td>Lost to follow-up</td>
<td>Supervision</td>
<td>Throughout study period</td>
</tr>
</tbody>
</table>

3 and 12 months, follow-up test sessions 3 and 12 months after starting the intervention; B, baseline test; EQ-5D-5L, EuroQol Group 5-dimension; GPE, Global perceived effect; HOOS, Hip disability and Osteoarthritis Outcomes Score; HUQ, Health Utilisation Questionnaire; IPCQ, Productivity Costs Questionnaire; NRS, Numeric Rating Scale; PO, primary outcome; QoL, quality of life; 1RM, one-repetition-maximum; SNBHW, The Swedish National Board of Health and Welfare questionnaire; THA, total hip arthroplasty.

Secondary outcomes

Hip disability and Osteoarthritis Outcomes Score

The HOOS subscales for symptoms, activities of daily life function and sport/recreation are secondary outcome measures.

Forty metre fast-paced walk test

The 40 m fast-paced walk test measures the total time (in seconds) it takes to walk 4×10 m (s) excluding turns. It is a valid and responsive measure of short distance maximum walking speed with excellent reliability.

Nine-step timed stair climb test

The 9-step timed stair climb test measures the time (in seconds) spent to ascend and descend nine steps and has excellent reliability in patients with symptomatic hip OA.

Nottingham leg extensor power rig

Leg extensor muscle power (watt/kg) will be measured using the Nottingham Power Rig, which has excellent reliability in patients with symptomatic hip OA. Leg extensor muscle power is a clinically important measure strongly correlated to physical function. For each leg, participants perform two warm-up trials followed by at least five trials with 30s rest between trials, until the participant does not improve in two successive trials, or at a maximum of 10 trials.

Unilateral one-repetition-maximum leg press

Maximal leg extensor strength of the index hip is measured by a 1RM test in a leg press resistance training machine, which is a highly reliable test in elderly populations. Warm-up consists of a set of 10 repetitions against a weight equivalent to 50% bodyweight. The weight for the first 1RM attempt is set corresponding to the rate of exertion during the warm-up set and the best result is found within five 1RM attempts.

Global perceived effect

The global perceived effect (GPE) will be assessed for three domains; pain, activities of daily living and quality of life, on a 7-point Likert scale.

Adverse events and serious adverse events

Throughout the trial, there will be continuous registration of adverse events (AE) and serious AEs (SAEs) as defined by The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. Physiotherapists supervising the exercise sessions will monitor events. During the 3-month and 12-month follow-up, the participants will be asked about potential AE and SAE using open-probe questions to be reported according to recommendations given by the CONSORT Group. SAEs will be reported to the Central Denmark Region Committee on Biomedical Research Ethics. In case of an unexpected high rate of AE and SAE in either group, the project group will consider study termination.

Adherence and drop-outs

Adherence to training will be registered by the physiotherapists supervising the exercise sessions. High adherence for the initial 3 months is defined as ≥80% attendance to the supervised exercise sessions. High adherence for the 9-months self-administered exercise is defined as ≥80% completion of the self-managed exercise sessions. Number of drop-outs and lost to follow-up will be registered throughout the study.

Other outcomes

Physical activity

Physical activity and sedentary behaviour are measured by The Swedish National Board of Health and Welfare questionnaire, which is a three-domain categorical answer patient-reported questionnaire.

Numerical rating scale for pain

Pain intensity is measured on a 11-item NRS where 0 represents no pain and 10 represent the worst pain imaginable.

Societal costs

Health utilisation is measured by a nine-item patient-reported cost questionnaire (HUQ) assessing healthcare and medicine usage. Productivity losses of paid work due to absenteeism, presenteeism and losses related to unpaid work is measured by the IPCQ.

EuroQol Group 5-dimension

The EQ-5D-5L is a patient-reported, valid and reliable questionnaire in patients with hip OA assessing health-related quality of life. The EQ-5D-5L measures the five dimensions: mobility, self-care, daily activities, pain/discomfort and anxiety/depression. Each dimension consists of one item, distinguished in five levels (no, slight, moderate, severe problems, unable to do). The EQ-5D also includes a visual analogue scale on which the
patients rate their health on a scale from 0 (worst imaginable health) to 100 (best imaginable health). 57,58

**Data management**

The project is registered at the Danish Data Protection Agency (Journal No 1-16-02-11-21). Before inclusion, all participants will have to give their written, informed consent in accordance with the Declaration of Helsinki II. All data collected in this trial is directly entered into the Research Electronic Data Capture (REDCap) for safe storage and will be treated confidentially by the research staff. The patient-reported data is entered into the REDCap database by the patient. Questionnaires are sent by email to participants at 6 months and 9 months follow-up. No data monitoring committee was established.

**Sample size**

The sample size calculation is based on the expected between-group difference in the 30 s CST from baseline to 3-month follow-up. Due to lack of hip OA specific data, the sample size calculation relies on data from knee OA. A mean change of 2.5 chair stands was found by Skoffer et al.59 in knee OA patients after 4 weeks of PRT and a mean change of 1.0 chair stands was found by Bennell et al.60 after 12 weeks of NEMEX also in knee OA patients, resulting in a difference between treatments of 1.5 chair stands. An SD of 2.52 for the 30 s CST is calculated from the 95% CI of the change in the intervention group of the study by Skoffer et al.59 Given a power of 0.90 and two-sided significance level α=0.05, the estimated sample size for a two-sample means test comparing PRT to NEMEX yields 122 participants. With an anticipated dropout rate of 30%, a total of 160 participants is the estimated sample size. For the primary 12-month comparison, the difference between groups receiving booster sessions and groups not receiving booster sessions is expected to be larger than for the comparison of PRT and NEMEX. Hence, we expect this study to be adequately powered for both comparisons.

**Statistical analysis**

An intention-to-treat approach will be used for analysing all changes in primary and secondary outcome measures including all enrolled participants according to randomisation group. Between-group comparisons of change from baseline to follow-up in the primary and secondary continuous outcomes will be analysed using a repeated measures mixed model with participants, clusters and sites as random effects, visits and treatment arms (that is treatment in case of the 3-month follow-up and the interaction between treatment and EBS for the 12-month follow-up) as fixed effects. Descriptive statistics for baseline demographics and clinical characteristics will be used to assess comparability of the groups and to adjust for potential confounders in per-protocol analyses. A per-protocol analysis will be conducted on participants who have a high adherence to interventions (≥80%) and have not undergone hip surgery. Data will be tested for normal distribution. The statistical level of significance will be set to p<0.05 and outcomes will be presented as means with 95% CIs. A difference in mean change between groups in the 30 s CST of 2.1 chair stands is considered a major clinically important improvement (MCII), as defined by Wright et al.46 To further guide the clinical interpretation, difference between groups in proportions of patients achieving the MCII for within-patients score change, as defined by Wright et al of 2.6 chair stands,46 will be analysed using a threshold of 20% between-group difference.61 Less than 20% is regarded as no meaningful difference between treatments and ≥20% as a meaningful difference between treatments. Furthermore, we will calculate the trial-specific minimal important difference by subtracting the mean 30 s CST score for participants reporting to have experienced a ‘small but not important change’ in GPE from those reporting ‘important change’ in GPE at 3 months. The statistical analyses and interpretation of data will be blinded to group allocation.62

**Economic evaluation of interventions**

The economic evaluation will be performed from both a societal perspective and a healthcare perspective and will use the intention-to-treat principle. An incremental cost–utility ratio will be calculated by dividing the difference in costs by the difference in effects. Cost of healthcare utilisation, productivity loss and total costs will be estimated. The EQ-5D-5L will be used to calculate quality-adjusted life-years (QALYs). Danish population norms will be used to calculate QALYs using linear interpolation between measurement points. Missing data will be imputed using multiple imputations by chained equations using Rubin’s rules.63 In order to account for the possible clustering of data, analyses will be performed using linear mixed models.64 Accounting for the possible clustering of data (eg, at the hospital level) is very important, as most economic evaluations fail to do so, whereas ignoring the possible clustering of data might lead to inaccurate levels of uncertainty and inaccurate point estimates.64 Bootstrap techniques will be used to estimate the uncertainty surrounding the cost-effectiveness estimates; cost-effectiveness planes and cost-effectiveness acceptability curves will be presented. Various one-way sensitivity analyses will be performed to test the robustness of the study results (eg, complete–case analysis and per-protocol analysis). The 12-month intervention period is the time horizon of the analysis. The economic evaluation will be performed in STATA and R.

**Patient involvement**

Before developing the design for this trial, two focus group interviews were conducted with patients with hip OA exercising at physiotherapy clinics. Open-ended questions were asked regarding optimal delivery of the exercise interventions, meaningful outcome measures and practical considerations, that is, how far they would commute for exercise and test sessions. These considerations guided the design of this trial.
DISCUSSION
This is the first trial to investigate the effectiveness of PRT compared with NEMEX and the effectiveness of EBS in patients with hip OA compared with a group with a matching initial intervention but without EBS. This study will contribute with novel and clinically important evidence enabling evidence-based recommendations and implementation of specific exercise modalities.

One of the major strengths of this trial is the multicentre, randomised, controlled, assessor blinded, design with participants recruited from four hospitals and 10 clinics making the findings highly generalisable to patients treated in the healthcare system. Moreover, the long-term follow-up will provide insights into the sustainability of the effects of exercise as treatment in hip OA and whether EBS can address the common problem of exercise effects diminishing over time. Last, the cost-effectiveness assessment will provide important information on societal costs associated with implementation of the interventions in healthcare systems and whether the effects are worth the costs.

The limitations of this trial are that the physiotherapists delivering the interventions and the participants receiving the interventions cannot be blinded to treatment allocation due to the nature of exercise interventions. This may introduce performance bias, especially in the patient-reported outcomes. However, the primary outcome and other functional performance tests are objective outcome measurements and less affected by performance bias. Second, there is no passive control group in this trial and as such, we can only compare the effects of the exercise modalities and not distinguish between the specific treatment effects and the contextual effects. Third, the sample size only provides power to detect a difference between PRT and NEMEX of 1.5 chair stands while the MCII has been found to be 2.1 chair stands. Hence, the clinical relevance of a difference in mean change of 1.5 chair stands is uncertain. Consequently, the clinical interpretation will also include a comparison of the proportion of patients in each group reaching the threshold for MCII and a trial specific minimal important difference is calculated. Fourth, the comparison of PRT and NEMEX as EBS is not only a comparison of exercise type but also different modes of delivery since participants are either given a fitness membership (PRT) or equipment to exercise at home (NEMEX). Hence, the setup and surroundings are different and will affect the comparison.


