Long-term effects of exercise therapy and patient education in patients with mild to moderate hip osteoarthritis

PhD Thesis

Ida Charlott Svege

The Norwegian Research Center for Active Rehabilitation (NAR)
Department of Orthopaedics, Oslo University Hospital

Faculty of Medicine, University of Oslo

February 2014

Series of dissertations submitted to the
Faculty of Medicine, University of Oslo
No. 1802

ISBN 978-82-8264-833-2

Cover: Inger Sandved Anfinsen.
Printed in Norway: AIT Oslo AS.

Produced in co-operation with Akademika Publishing.
The thesis is produced by Akademika Publishing merely in connection with the thesis defence. Kindly direct all inquiries regarding the thesis to the copyright holder or the unit which grants the doctorate.

From a certain point onward there is no longer any turning back.
That is the point that must be reached.

Franz Kafka

Content

ACKNOWLEDGEMENTS 7
ABBREVIATIONS 10
INCLUDED PAPERS 11
SUMMARY OF PAPERS I-III 12
PREFACE 14
INTRODUCTION 15
Osteoarthritis and its pathogenesis 15
Clinical symptoms, structural changes and the diagnosis of hip OA 15
The prevalence and burden of OA 16
Natural history and progression of hip OA 17
Treatment recommendations in hip OA 19
Outcome measures in OA 22
AIMS OF THE THESIS 25
MATERIAL AND METHODS 26
Ethical considerations 26
Study design 26
Subjects 26
Recruitment and screening for inclusion 28
Randomisation process 28
Interventions 30
Baseline assessments 31
Outcome measures 32
Sample size calculations 36
Statistical analyses 37
SUMMARY OF RESULTS 40
Study I 40
Study II 41
Study III 41
DISCUSSION 43
Main findings 43
Study design and methodological considerations 43
Subjects 45

Interventions 46
Outcome measures 47
Sample size and data analysis 49
Results 51
CONCLUSIONS 56
FUTURE PERSPECTIVES 57
PAPERS I-III 75

ACKNOWLEDGEMENTS

It is with mixed feelings I am writing the final words of this dissertation. I am grateful, proud and tired, but also sad to leave this period of time behind me. It has truly been five educational, challenging and eventful years, and I am so thankful for having had the opportunity of spending them as a part of the research group NAR – Norwegian Research Center for Active Rehabilitation. This work would never have been possible without the support of many people, and I would like to express my deepest gratitude and appreciation to all of you who have contributed to this thesis. In particular, I would like to thank:

May Arna Risberg, main supervisor and chair of NAR. I owe you my deepest gratitude for giving me this huge opportunity. Thank you for believing in me, for listening to my ideas, and for inspiring me to always set high standards for myself and be well prepared. I am also grateful for your patience during my two maternity leaves, and for your support and encouragement during more troublesome times. I truly admire your capacity, your initiative and enthusiasm, and your genuine belief in active rehabilitation. You have taught me so much about research!

Lars Nordsletten, co-supervisor. Thank you for your huge dedication to this project, for your contribution in screening patients and rating x-rays, and for your help during manuscript preparations. I sincerely appreciate your positivity and that you have encouraged me to aim high. Your genuine interest in research has been to my benefit, as you have generously shared your knowledge during these years.

Inger Holm, co-supervisor. I would in particular like to express my gratitude for your stand-in performance and for keeping me on track during May Arna's sabbatical Down Under. Thank you also for your contribution during the writing of the thesis and the manuscript of Paper I. Your large field of knowledge and ability to see the complete picture has been valuable to me.

Linda Fernandes, co-author, previous colleague and dearest friend. Thank you for letting me inherit your ~~baby~~ PhD-project. It has been a huge relief for me to have you to lean on in questions regarding the project, and I truly appreciate your patience and large recall capacity all the times I have bombarded you with questions on tiny details. Your manuscript feedback is always fast, to the point and redemptive. Your research skills impress me and your ability to see the whole picture has been extremely valuable. I love your sense of humour, and it has helped me to get out of the pit at several occasions during these years.

Elin Kolle, co-author. Thank you for your contribution on article III and for helping me to understand, at least a little bit of, the difficulties in physical activity research. I could not have done it without you.

I am very grateful to all who has been involved in administering the project, treating and assessing the patients, and contributing to the completion of this project in other ways. Kristin Bølstad, Emilie Jul-Larsen Aas, Karin Rydevik, Sigmund Ruud and Marte Lund, you are especially acknowledged for your significant contribution.

Leiv Sandvik and Ingar Holme. Your advices on statistical and methodological issues are highly appreciated.

Ingrid Eitzen. Thank you for your constructive feedback and for proofreading this dissertation.

The patients participating in the project. Thank you for your time and effort.

My friends and colleagues in NAR, Annika, Birgitta, Karin, Marte, Håvard, Silje, Hege, Britt-Elin, Linda, Ingrid, Kristin, Emilie, Kari Anne. Thank you for giving me the best social work environment possible! You have made me look forward to going to work – every day! In particular, I would like to express my deepest gratitude to Hege S, Britt Elin, Kari Anne, Linda, Ingrid and Kristin for showing interest in my work, for giving me feedback on ideas, methodological approaches and manuscripts, and for being supportive in times of frustration. During these years you have become some of my best friends! And Ingrid, after spending 2.5 year together with you in the 'hip girls office' I am so happy to pursue you to OUS.

Co-workers at NIMI. I truly admire your ability to incorporate new knowledge for the benefit of each patient. Because of you, I consider NIMI to be the #1 clinic in combining research and clinical practice.

Colleagues at Martina Hansens Hospital. Your encouragement and support during these years have been valuable to me.

My dear friends, Birgitte, Elisabeth, Fia, Ella, Inger Lene, Jannike. Thank you for showing interest in my work, but mostly for being fun to be around, and for being my social off-work alibi these past years.

Helene and Per. Your support during this hectic period has been crucial. Thank you for babysitting the kids whenever we needed time to work (or ski, run, kayak).

Mamma and pappa. You have always told me that that I could do what I set my mind to! Thank you for giving me what it takes, for believing in me, and for your pride in me. I also appreciate all your help on practical issues during these years.

Selma and Mats. Thank you for not understanding ANYTHING about what I have been doing at work during these years – that has helped me to take my mind off work and to put my work and research in a larger perspective. You did both arrive during my PhD-period, Selma after just three months and Mats after almost three years, and you have ensured that these five years have been the busiest, most tiring and most eventful period of my life. Also, the two of you have, together with this thesis, made sure that I have not wasted too many hours sleeping these past five years. Now, less work and more play (and sleep)!

Dag. Thank you for telling me that I am good enough – both in what I do and who I am! Thank you for advising me to be more cynical when I try to do 'everything' and please 'everyone', and for overlooking my unreasonableness during times of heavy workload and little sleep. Your support has been crucial in this and, not least, in previous challenges. With your background as an actuary, I would have expected you to have been more helpful in my struggles with statistical issues. You can talk for days about the Gaussian distribution, but we have to agree: You really don't know that much about biostatistics!

Ida C. Svege

Bærum, February 12th 2014

The work comprised in this dissertation was carried out at the Norwegian Research Center for Active Rehabilitation (NAR) in the period 2005-2011. NAR is a formal collaboration between The Department of Orthopaedics, Oslo University Hospital, NIMI and The Norwegian School of Sports Sciences, Department of Sports Medicine.

The main financial source for funding this dissertation was the former Research Council at Ullevål University Hospital (VIRUUS), later Oslo University Hospital. Furthermore, the project has been financially supported by the EXTRA funds from the Norwegian Foundation for Health and Rehabilitation, through the Norwegian Rheumatism Association, and the South-Eastern Norway Regional Health Authority Osteoarthritis research group. NIMI is acknowledged for providing NAR with rehabilitation facilities, research staff and office facilities.

ABBREVIATIONS

ACR – American College of Rheumatology
BMI – Body mass index
CI – Confidence interval
ESR – Erythrocyte sedimentation rate
EULAR – European League Against Rheumatism
HHS – Harris Hip Score
ICC – Intraclass correlation coefficient
IPAQ – International Physical Activity Questionnaire
ITT – Intention to treat
LoA – Limits of Agreement
LOCF – Last observation carried forward
MCID – Minimal clinically important difference
MD – Medical doctor
MDC – Minimal detectable change
MET – Metabolic Equivalent of Task
MJS – Minimum joint space
NAR – Norwegian research center for Active Rehabilitation
NIMI – Norwegian sports medicine clinic
NSAIDs – Non-steroidal anti-inflammatory drugs
OA – Osteoarthritis
OARSI – Osteoarthritis Research Society International
OMERACT - Outcome Measures in Rheumatology
PASE – Physical Activity Scale for the Elderly
PT – Physical therapist
RCT - Randomised, controlled trial
ROM – Range of motion
SD – Standard deviation
SEM – Standard error of the measurement
THR – Total hip replacement
VAS – Visual analogue scale
WHO – World Health Organisation
WOMAC – Western Ontario and McMaster Universities Osteoarthritis Index
6MWT – Six Minute Walk Test

INCLUDED PAPERS

I Svege I, Fernandes L, Nordsletten L, Holm I, Risberg MA. Effect of exercise therapy and patient education on impairments and activity limitations in patients with hip osteoarthritis. A randomized controlled trial with 29 months follow-up. Submitted to J Orthop Sports Phys Ther, Jan 14 2014

II Svege I, Nordsletten L, Fernandes L, Risberg MA. Exercise therapy may postpone total hip replacement surgery in patients with hip osteoarthritis: a long-term follow-up of a randomised trial. *Ann Rheum Dis* 2013 Nov 20 [Epub ahead of print]

III Svege I, Kolle E, Risberg MA. Reliability and validity of the Physical Activity Scale for the Elderly (PASE) in patients with hip osteoarthritis. *BMC Musculoskeletal Disorders 2012, 13:26*

SUMMARY OF PAPERS I-III

Paper I reports the results from a long-term (29 months) follow-up of a randomised, controlled trial (RCT) evaluating the effect of exercise therapy and patient education by outcome measures of specific impairments and activity limitations in patients with hip osteoarthritis (OA). One hundred and nine patients with radiographic and clinical hip OA, considered to have mild to moderate symptoms, were included. The mean age of the included patients was 57.8 years, and 54 % were women. Patients with knee pain/knee OA, back pain or specific comorbidities were excluded form participation. Patients were randomised to receive either 1) a patient education program followed by a 12 week exercise therapy program and (exercise therapy group) or 2) a patient education program only (control group). Assessments were conducted four, ten and 29 months after enrolment, and included hip range of motion (ROM), isokinetic concentric knee and hip flexion and extension muscle strength, the six minute walk test for distance (6MWT) and pain during walking assessed by a visual analogue scale (VAS) for pain. At the 29-month follow-up 10 % of the patients were lost to follow-up, 25 % had undergone total hip replacement (THR) surgery and did not participate at subsequent follow-ups, and 3 % did not complete the clinical/functional tests. Fifty three percent of the patients in the exercise therapy group were compliant to the program, defined as participating in 80 % of the exercise sessions. The results revealed that the exercise therapy group reported significantly less pain during walking over the 29 month follow-up ($p=0.018$), but no group differences were shown for ROM, muscle strength or walking capacity assessed by the 6MWT.

In Paper II we conducted a long-term follow-up of the 109 patients in the RCT to compare the 6-year survival of the native hip to THR between patients who were given exercise therapy and patient education and patients who were given patient education only. Additionally, we evaluated the treatment effect on self-reported pain, stiffness and physical function by using the Western Ontario and McMasters Osteoarthritis Index 3.1™ (WOMAC). Main outcome measure for this long-term follow-up was survival of the native hip to THR. Thus, 3.6-6.1 years after inclusion, in May 2011, we assessed whether and when patients had gone through THR. The WOMAC was used at the four, ten, 16 and 29 months follow-up, and was included as an outcome measure in order to compare levels of pain, stiffness and physical function between the exercise therapy group and the control group prior to THR or end of study. The response rate at the 3.6-6.1 years follow-up was 94 %. Patients who were lost to follow-up were treated as censored in the analyses. The 6-year survival of the native hip to THR was significantly higher in the exercise therapy group compared to the control group. The Hazard rate in the exercise therapy group compared to the control group was 0.56, indicating that the 6-year risk for THR was reduced by 44 % in the exercise therapy group. Furthermore, we found that the exercise group had significantly better self-reported physical function over the 29 months period compared to the control group ($p<0.01$). However, no group-differences were revealed for self-reported pain ($p=0.083$) or stiffness ($p=0.112$).

In Paper III we evaluated the reliability and validity of the Physical Activity Scale for the Elderly (PASE) for this hip OA patient population with mild to moderate symptoms. Forty patients with a mean age of 61.3 ±10.0 years (50 % women) were recruited from the RCT and included in the study. The test-retest reliability was examined by both relative and absolute measures of reliability. The construct validity was studied by comparing the PASE to both an accelerometer and the International Physical Activity Questionnaire short form (IPAQ). The test-retest reliability of the total PASE score was considered to be moderate, with an acceptable intraclass correlation coefficient (ICC) of 0.77, but with relatively large measurement error. The calculated minimal detectable change (MDC) was 87, implying a change in PASE score larger than 87 points to be considered to represent a real change. The correlation coefficient between the total PASE score and the categories representing total physical activity level of the IPAQ and the accelerometer recordings was 0.61 and 0.30, respectively. The correlation coefficients were in line with the a priory hypotheses. However, the correlation coefficient between the PASE and the accelerometer did not exceed 0.5, which has been defined as the cut-off for acceptable construct validity. In addition, we found that the specific items addressing light, moderate and vigorous physical activity to be of poor validity and test-retest reliability.
Hence, the PASE was found inadequate to provide valid data on physical activity level and intensity in patients with mild to moderate hip OA.

PREFACE

Exercise therapy is recommended to be offered to all individuals with OA of the hip or knee as a first line treatment. These recommendations are based on the evidential beneficial effect of exercise on self-reported pain and function in RCTs including patients with knee OA. However, guidelines based on the aggregated effect for hip and knee OA may be inappropriate, as the exercise effect and the underlying mechanisms may differ between the two joints. Previous studies suggest that exercise therapy may improve self-perceived pain and function in patients with hip OA, but evidence is limited due to few available RCTs including patients with hip OA only. Furthermore, knowledge is sparse regarding the long-term effect of exercise interventions. Thus, the need for high-quality research investigating the effect of exercise therapy in patients with isolated hip OA is highly emphasised.

To address this lack of evidence the Norwegian research center for Active Rehabilitation (NAR) launched an RCT in 2005 to evaluate the effect of exercise therapy and patient education in patients with hip OA. The study was initiated by Professor, Physical Therapist (PT), PhD May Arna Risberg, Professor, Medical Doctor (MD), PhD Lars Nordsletten, PT, PhD Kjersti Storheim and PT, PhD Linda Fernandes, as a collaboration between Ullevål University Hospital (later Oslo University Hospital), the Norwegian School of Sports Sciences and the Norwegian Sports Medicine Clinic (NIMI). Patient inclusion was conducted by Lars Nordsletten and Linda Fernandes at Ullevål University Hospital between April 2005 and October 2007. Linda Fernandes defended her PhD thesis, comprising four Papers which all emerged from this project, in January 2011.

This PhD thesis is a 2.4-6.1 year follow-up of the RCT and consists of three Papers. In Paper I and II the long-term results of exercise therapy and patients education in patients with hip OA is evaluated, whereas the validity and reliability of a questionnaire assessing physical activity is investigated in Paper III. A total of 109 patients were included in the RCT. Initially, all 109 patients participated in an education program, before the patients were randomised to either an exercise group, participating in a 12-week exercise therapy program especially developed for patients with hip OA, or to a control group, receiving no further treatment. Follow-ups were conducted at four, ten, 16 and 29 months, and at 3.6-6.1 years after inclusion. Assessments included outcome measures of impairments and activity limitations, patient-reported outcome measures, and time to THR surgery. Forty patients from the RCT were included in the reliability and validity study. Further long-term follow-up of the RCT with assessments at five and ten years after inclusion are on-going.

INTRODUCTION

Osteoarthritis and its pathogenesis

Osteoarthritis is a degenerative joint disease which affects the synovial joints resulting in pathological changes in the articular cartilage and its surrounding tissue[49, 62]. In healthy joints the primary function of the cartilage is to distribute joint loading and decrease friction between the joint surfaces. In early OA the biomechanical properties of the cartilage is comprised, and further disease development includes cartilage loss which eventually extends the cartilage and exposes the subchondral bone[62]. Remodelling and sclerosis of subchondral bone occurs, with formation of bone cysts and osteophytes at the joint margins[27, 62]. Thickening of the joint capsule, and hypertrophy and fibrosis of the synovium is common[62]. In early disease intermittent episodes of pain are common, as a dynamic process with alternating cartilage damage and attempts of repair take place. Later, in more severe disease irreversible and progressive pathologic changes may occur[111]. Traditionally OA have been regarded as a mechanical rather than inflammatory disease. Later, however, it has been found that inflammatory factors may contribute to disease development[20, 46], possibly influenced by a systemic metabolic component[198]. Hence, the development of OA seems to be complex and different phenotypes of the disease may exist.

Clinical symptoms, structural changes and the diagnosis of hip OA

The main clinical symptoms in hip OA typically comprise pain, functional limitations, and joint stiffness[49, 62, 111]. Clinical examination reveals tenderness, restricted ROM, alterations in gait pattern, and muscular weakness and atrophy[111]. These features, accompanied by higher age, make it rather easy to suspect OA. The clinical criteria of the American College of Rheumatology (ACR)[5] include the presence of pain in combination with 1) hip internal rotation ≥15 degrees, pain during internal rotation, morning stiffness for ≤60 minutes, and age >50 years, or 2) hip internal rotation <15 degrees, and an erythrocyte sedimentation rate (ESR) ≤45 mm/hour or hip flexion ≤115 degrees if no ESR was obtained.

Radiographic examination is usually conducted to confirm the clinical diagnosis. Radiographic definition of hip OA is based on the presence of joint space narrowing, with or without accompanying osteophyte formation, subchondral bone cysts and sclerosis[114]. Several systems of assessing radiographs of the hip exist[8, 9, 36, 40, 128], with the minimum joint space (MJS) width method showing measurement properties that are both superior to other radiographic criteria for disease and also with a stronger association with self-reported pain[32, 36]. According to Chu Miow et al.[32] there is evidence for a weak association between the MJS and the patients' symptoms and for a predictive validity for subsequent THRs.

The severity of clinical symptoms in hip OA is poorly correlated with the structural changes[21, 32, 36, 142, 175]. Radiographic findings of structural joint damage consistent with OA are common in the elderly population with or without the presence of pain[38], and patients with advanced structural changes within the joint may have few or no clinical symptoms. On the contrary, patients presenting with pain and other symptoms of OA, may have little or no radiographic evidence of OA, as plain radiographs can only capture the structural changes in the joint which occurs relatively late in the disease process. This is clearly shown in the study by Birrell et al.[21] where 41 %, 56 %, and 3 % of patients presenting with *no* hip pain had none, mild or severe evidence of radiographic hip OA, respectively. Among patients presenting *with* hip pain, 29 %, 55 %, and 16 % had none, mild or severe radiographic hip OA, respectively[21]. As the articular cartilage is both aneural and avascular it is incapable of directly generating pain[111]. Loss of cartilage, resulting in the joint space narrowing visible on radiographs, is therefore not necessarily accompanied by pain. This can partly explain the poor correlation between clinical symptoms and structural changes. However, the subchondral bone, periosteum, periarticular ligaments and muscles, synoviom and joint capsule are richly innervated by nociceptive fibres[111]. Furthermore, joint pain in OA is typically exacerbated by activity and relieved by rest, while pain at rest and during the night may be present in more advanced disease[111, 112]. According to Dieppe and Lohmander[49], pain in OA should be assessed within a biopsychosocial framework including interaction between activity, incidents within the joint, pain sensitisation and the patient's perception of pain.

The prevalence and burden of OA

In a systematic review by Dagenais et al.[38] the estimated reported prevalence of radiographic hip OA varied from 0.9 % to 27.0 %, giving an overall mean prevalence of 8.0 % in the general adult population. Higher prevalence was associated with higher age (from 55 to 80 years), and studies using the Kellgren and Lawrence scoring system reported higher prevalence of hip OA[38]. Pereira et al.[177] conducted a systematic review to evaluate the differences in prevalence of OA according to case definition. Whereas the overall prevalence of hip OA was 10.9 %, radiographic OA presented a higher prevalence, and self-reported and symptomatic OA were less common. Twenty-seven studies on the prevalence of hip OA was included in the review, 19 of them reported prevalence of radiographic hip OA ranging from 1-45 %, four studies reported prevalence of self-reported hip OA from 5.5-9-7 %, and four studies reported prevalence of clinical and radiographic hip OA from 0.9-7.4 %[177]. In a population-based study conducted in North Carolina, US, 28 % of persons older than 45 years had radiographic findings consistent with hip OA, but only 10 % presented both clinical and radiographic evidence of the disease[121]. Other population-based studies have reported somewhat lower prevalence of clinical and radiographic hip OA, 2.8 % in Turkey[242] and 2.5 % in France[91]. In Norway, the prevalence of self-reported hip OA have been found to be 5.5 % in the general adult population[90]. These prevalence estimates confirm that only a proportion of individuals with pathological radiographic findings present clinical symptoms of OA. The prevalence of hip OA has remained relatively stable over many decades[39], but OA is considered to

represent a growing epidemic, due to both longer life expectancy, and the increasing incidence of obesity and inactivity[166, 238].

Musculoskeletal disorders are among the major contributors to the global burden of disease, particularly in high-income, western countries and in the elderly population[228]. Kingsbury et al.[130] studied the impact of OA in five large European countries and concluded that the impact of OA on health status and work productivity was substantial. Furthermore, expenses related to direct and indirect medical costs, and work disability in OA are extensive[81, 82, 179, 193]. In a Norwegian report, musculoskeletal disorders are referred to as the largest expenditure in the somatic specialist health services, with primary joint replacements being highlighted as especially costly[214]. To limit the impact of the disease on decreased health status and work ability and avoid a further increased THR rate, and thereby reduce healthcare costs, a shift towards more preventive interventions and early stage treatment is needed.

Natural history and progression of hip OA

Although OA is commonly regarded as a steadily progressive disease, its natural course is poorly studied. Disease progression may involve worsening of pain and symptoms as well as structural progression which may result in end-stage disease requiring surgical treatment. However, this seems to vary widely between persons and in some cases progression to severe disease never occurs[49, 111]. According to Dennison and Cooper in Brandt et al.[45], 19-83 % and 29-65 % of patients with hip OA experienced clinical worsening and radiographic progression, respectively, within 2.3-12 years. Van Dijk et al.[224] conducted a systematic review of the literature to evaluate the course of functioning in hip OA. They found that pain and functional status seemed to deteriorate slowly at the beginning, but after three years a worsening was evident[224]. Furthermore, the findings in a recent study by Eitzen et al.[57] (submitted manuscript) suggest that a subgroup of patients, initially presenting with mild to moderate hip OA, did not progress clinically or radiographically over 6.6 years. Based on the general consensus that total joint replacement surgery is appropriate only in advanced stages of the disease, joint replacement may be interpreted as a result of disease progression[1, 7, 52, 157]. While the lifetime risk of clinical and radiographic hip OA have been estimated to be 25.3 %[165], the lifetime risk of THR have been estimated to be 10.1-11.6 % in women and 7.1-9-9 % in men[24, 37].

A limited number of studies have investigated the long term course in subjects with radiographic evidence of hip OA, not necessarily accompanied by clinical symptoms. Franklin et al.[74] found that 17 % of patients presenting with only radiographic signs of hip OA had gone through THR surgery within 11 to 29 years. Additionally, Lane et al.[139] found that within 8.3 years, 13 % of women with radiographic hip OA had gone through THR, 23 % had increased lower limb disability, and 65 % had progressed radiographically or had gone through surgery. Disease progression was found to be greater in patients who presented clinical symptoms of hip OA in addition to radiographic

evidence, with 24 % going through THR within 8.3 years[139]. Ledingham et al.[143] and Dieppe et al.[47] studied disease progression in patients presenting with both clinical and radiographic evidence of OA. Ledingham et al[143] found that after 3-72 months, 66 % of the patients had increased pain, 10 % had experienced a functional decline, and 15 % had radiographic worsening, defined as a change in in Kellgren and Lawrence grade. However, 57 % were classified as having a Kellgren and Lawrence grade 4 at inclusion, and could thus not progress further. In total, 53 % of the hips had been replaced within 3-72 months[143]. Dieppe et al.[47] found that 39 % of the patients had improved and 43 % had worsened after three years, and 21 % had improved and 54 % had worsened after 8 years. Thirty-two percent and 48 % had undergone THR prior to the three and eight year follow-up, respectively[47]. In patients presenting with hip pain, but no radiographic evidence of hip OA, 15 % and 28 % had worsened after three and six years, respectively, and 12 % and 22 % had gone through THR within three and six years, respectively[149]. In the study by Fernandes et al.[65], which reported the short-term results of the RCT on which this thesis is based, 15 % of patients presenting with clinical and radiographic hip OA had gone through THR within 16 months, indicating that some patients experienced substantial progression in a relative short period of time.

Summed up, the findings in these studies suggest that the natural course of hip OA is multifaceted. Some patients experience diminutive clinical changes over extended periods, some remain unchanged or even temporarily improve, while others show a rapid decline with disabling impacts within few years[27]. Furthermore, rapidly developing radiographic changes have been found to have a negative effect on self-reported pain and function[235].

The reasons for this large variability in the natural course of the disease are not fully understood. However, higher age, female sex, narrower joint space, bony sclerosis, femoral osteophytes and increased hip pain have all been found to be associated with a more rapid disease progression[143, 240]. Additionally, older age[125, 240], higher body mass index (BMI)[12, 67, 68, 125, 188, 231], narrower joint space and femoral osteophytes[240], activity-demanding work[68], intensive exercise[161], higher education and willingness to consider surgery[101] have been identified as potential risk factors for THR, suggesting that some of these factors might accelerate disease progression. Furthermore, it has been suggested that idiopathic hip OA characterised by superior or medial migration of the femoral head follow different natural courses, with OA with medial migration having better long-term prognosis[98]. The increased risk of radiographic hip OA in siblings of patients with manifest severe hip OA indicate that hip OA is also under strong genetic influence[140]. Finally, the overexpression of inflammatory mediators in synovial tissue in early compared to late knee OA may be associated with the progressive cartilage degradation seen in some patients[19].

Treatment recommendations in hip OA

It is recommended that general information on the disease, advice on life-style alterations, exercise and physical activity should be provided to all patients with hip OA as a first-line treatment to empower self-management of the disease[34, 63, 106, 244]. Patient education and exercise therapy can, therefore, be considered as part of the core treatment in lower extremity OA. The use of pharmacological interventions should be considered in patients with consistent symptoms and clinical findings. Furthermore, invasive interventions are to be considered in cases with severe symptoms, where other treatment options have failed, with THR as the final treatment option for end-stage disease[34, 63, 106, 244]. According to the ACR 2012 Recommendations, patients with hip OA are also strongly recommended to participate in cardiovascular and/or resistance exercises, aquatic exercises, and to lose weight if overweight[106], but the evidence for the effect of these modalities in hip OA is sparse. Furthermore, all treatment, including information and education, lifestyle changes, and exercise therapy, should be individualised for each patient[63].

As there is consensus regarding physical activity and exercise recommendations across the many clinical guidelines, future effort should target implementation of these guidelines in primary and secondary health care practices[141, 169]. Only 42 % of patients with self-reported doctor-diagnosed OA reported that they had received advice on increasing their level of physical activity, even if it is considered as part of the core treatment in OA[69].

Patient education

For educational modalities, strong evidence has been established for a negligible to small, but statistically significant effect, with effect sizes of 0.06-0.12 for pain and 0.06-0.07 for physical function[232, 244]. According to the Osteoarthritis Research Society International (OARSI) recommendations for the management of hip and knee OA[244] and the European League Against Rheumatism (EULAR) 2013 Recommendations for the non-pharmacological core management[63], education and information should be included in the core treatment in patients with lower limb OA, as it is believed to be essential to achieve adequate self-management of the disease. However, in a recent review it was concluded that self-management programmes have minimal effect in patients with OA[135].

Exercise therapy

Exercise therapy has been found to be beneficial in reducing pain and improving physical function in manifest lower limb OA[75, 76, 103, 223, 244]. As few RCTs have included patients with hip OA only, evidence for the effect of exercise is mainly based on the many trials that have included patients with knee OA. Thus, the aggregated effect estimates of exercise in knee and hip OA reported in meta-analyses may be less valid for patients with hip OA only. Knowledge on the long-term effects of exercise therapy in OA is sparse, as

most studies only evaluate short-term effects[42]. Consequently, exercise has currently no evidential long-term effects on pain and/or physical function in patients with knee and/or hip OA[183]. However, it has been suggested that implementation of additional booster sessions may improve the long-term effect[183]. Furthermore, adherence to the recommended exercises has been found to be associated with better treatment outcomes in the long-term[182].

In hip OA, effect sizes of 0.33-0.38 for pain have been reported for exercise interventions[76, 103, 244]. These effect estimates are similar to or larger than the demonstrated effect of commonly used analgetisc, but without the potential side effects of pharmacological treatment. Fransen et al.[76] estimated an effect size of 0.10 for physical function after exercise therapy, but this estimate was based on only one RCT. Effect has been found to be larger for exercise interventions administered in person[103]. The limited number of RCTs conducted in the field of hip OA and their relatively small sample sizes diminishes the possibility to calculate effect estimates and restricts the confidence of the estimates.

The study by Tak et al.[209] was the only RCT to evaluate the effect of exercise therapy in patients with hip OA only published prior to the start of our RCT. They studied the effect of an exercise program designed for patients with hip OA, which comprised strengthening exercises, health education and ergonomic advice[44]. The exercise program, provided as group sessions over eight weeks, had a positive effect on both pain and self-reported physical function in patients with clinical diagnosis of hip OA (mean age of the included patients 68 years, 68 % women)[209]. In 2010 the results of the 16 month follow-up of our RCT was published (mean age of the included patients 58 years, 56 % women)[65]. No significant difference in self-reported pain was revealed between the group receiving both patient education *and* supervised exercises compared to the group receiving patient education *only*. However, a significant difference was found for self-reported physical function in favour of the exercise group[65]. The same year, Juhakoski et al.[124] showed that patients with clinical and radiographic hip OA (mean age of the included patients 67 years, 70 % women) who were given one weekly exercise session for 12 weeks and four additional booster sessions, had better effect in self-reported physical function after six and 18 months, compared to patients receiving usual care. In addition, French et al.[78] found that eight exercise sessions provided over an eight week period, with or without additional manual therapy, provided beneficial results in self-reported physical function nine months later, compared to waitlist controls. The patients included in the study had clinical and radiographic hip OA (mean age of the included patients 62 years, 64 % women)[79]. Summed up, these recent publications suggest that exercise therapy improve physical function in patients with hip OA over 6-18 months follow-up, but the results are less consistent for pain.

Pharmacological treatment

Optimal management of hip OA requires a combination of non-pharmacological and pharmacological treatment modalities[243]. The ACR 2012 Recommendations[106] and the EULAR evidence based recommendations[243] both recommend paracetamol to be the analgesic of first choice in patients with mild to moderate pain, while non-steroidal anti-inflammatory drugs (NSAIDs) should be added in patients who respond inadequately to paracetamol. The use of opioids is generally not encouraged, but Hochberg et al.[106] suggests that patients with inadequate response to other pharmacological treatment, who are unwilling to or cannot undergo THR, may use opioids. The effect sizes have been calculated to be 0.14 for paracetamol, 0.29 for NSAIDs, and 0.24-0.58 for other pharmacological treatments, including diacerin, cox-2 inhibitors, chondroitin sulfate and glucosamine sulphate[243]. Further investigation on whether long term use of symptomatic slow acting drugs in OA, including glucosamine sulphate, chondroitin sulphate and diacerin, can comprise a structure-modifying effect on joint cartilage, synovium and subchondral bone, and thus postpone joint replacement, is encouraged[106, 243].

Surgical treatment with THR

There is general consensus that THR is an appropriate treatment option in cases of advanced OA, where non-surgical treatment has failed. The effect of THR has not been evaluated by comparing it to a control group following a randomised study design. However, effect sizes of 1.7-2.6 for pain, 1.0-2.2 for stiffness, and 1.8-2.9 for physical function have been reported for post-operative outcome after THR in non-controlled follow-up studies[159]. According to Fitzpatrick et al.[66] 54-86 % of the patients were rated as having good to excellent postoperative results four to eleven years after surgery. Thus, THR has been considered a successful surgical treatment modality, of which the majority of patients can expect improved function and reduced pain for several years, and older patients may never need a revision of the prosthesis[66]. Cost benefit calculations have revealed a larger expected cost in younger patients, as they most likely will require future THR revisions[66] Nevertheless, the quality adjusted life-years has been shown to be higher in younger patients[66, 190]. Overall, THR thus seem to be effective in terms of improving health-related quality-of-life dimensions[60]. Men seem to benefit more than women, as do persons with poorer preoperative outcomes, while comorbidities have been found be associated with an inferior postoperative result[60, 100]. However, Räsänen et al.[190] found that only 58 % of the patients had experienced an improvement larger than the minimal clinically important difference (MCID) after one year. Furthermore, according to a review by Vissers et al.[227] full recovery of physical function after THR is not obtained. Perceived physical function improved from 50 % preoperatively to 80 % 6-8 months postoperatively of that of healthy controls. Furthermore, functional capacity improved from 70 % to 80 %, and actual daily activity increased from 80 % to 84 % of that of healthy controls[227]. Nilsdotter et al.[172] found that patients who underwent THR had poorer physical function 3.6 years after surgery compared to a reference group, but health related quality of life was similar in the two groups. Additionally, Harding et al.[96]

and De Groot et al.[43] found that despite improvement in pain, patient-reported function, physical performance and patient-reported physical activity after joint replacement surgery of the knee or hip, none or a very small increase was found for objectively measured physical activity level.

For the 5-year period 1996-2000, the average annual incidence rate for THR due to primary hip OA in Norway was 85 per 100.000, with an increasing incidence rate over the years[152]. A higher THR incidence was associated with older age and female sex[2]. In a large prospective population-based study in Norway, Flugsrud et al.[68] found that 1.3 % had gone through THR due to primary hip OA within nine years. This was in line with the findings by Frankel et al.[73] who estimated the prevalence of hip disease severe enough to require THR surgery to be 1.5 % in the general population. In 2012, 7786 THRs were conducted in Norwegian hospitals, representing a 17 % increase since 2007, and a 26 % increase since 2002. Over the same period the number of THR revisions have increased even more, with a 25 % increase since 2007, and a 36 % increase since 2002[167, 168].

According to Gossec et al.[87], clinical disease severity varies widely at the time of THR in different orthopaedic centres across Europe. Clearly defined indications for THR surgery would be helpful in selecting those patients who are likely to benefit most from surgical treatment. Several attempts have been made to identify criteria for surgery[54, 87, 158, 186] but still no general agreement regarding explicit criteria for THR exist. Based on the increasing number of THRs and the accompanying health care costs, and the increased attention towards the importance of early initiation of secondary preventive actions, a need for a shift towards non-surgical treatment approaches has been proposed[28, 84, 110]. According to Hunter[110], the focus in OA research has been on pain relieving end-stage treatment in persons with established OA, rather than treatments to prevent onset and progression of the disease.

Outcome measures in OA

Measuring treatment effect and Disease Progression

The efficacy of treatment interventions in hip OA can be evaluated by patient-reported assessments of pain, function, and stiffness as well as by patient-reported or clinician-rated evaluations of treatment response and disease status. Instruments used to measure outcome should be valid, reliable and responsive to change. Furthermore, the use of already published outcome measures has been encouraged, to allow comparison across trials and treatments[6]. To facilitate standardisation among the diverse measures that are available, the OARSI and the Outcome Measures in Rheumatology (OMERACT) have recommended a core set of outcome domains to be used in clinical trials[6, 18]. According to these recommendations, assessment of pain, physical function and patients' global rating should be included as a core set of outcome measures. In addition, radiographic assessment is suggested to be conducted in studies lasting longer than one year. Several other outcome measures, including measures of stiffness, functional performance and time to surgery, have been suggested to be included as optional, additional measures[6, 18].

Patient-reported outcome measures, including VAS, numeric rating scales, and disease-specific outcome measures, such as the Lequesne Index[146], the WOMAC[17], and the Hip disability and Osteoarthritis Outcome Score[171] have been widely used. Disease-specific questionnaires are easy to administer, their perspective is highly patient-focused, and many of them have been found to have acceptable validity and reliability in OA populations. Furthermore, performance based outcome measures can provide important supplementary information concerning physical function. The inclusion of both patient-reported outcome measures and outcome measures assessing activity limitations are therefore encouraged as they are likely to measure different constructs of function[241]. Recently, a consensus-derived set of performance based tests to assess physical function in patients with hip or knee OA have been proposed as suitable for use in clinical trials in lower limb OA[50]. Additionally, evaluation of treatment effect on specific impairments, including ROM and muscle strength, may be valuable, as they are suggested to be associated with functional performance and self-perceived physical function[13, 14, 156, 195, 206]. Additionally, decreased hip ROM and lower limb muscle strength has been demonstrated in patients with hip OA when compared to healthy controls[14, 195], with larger deficits in patients with more severe disease[122].

Based on the general consensus that total joint replacement surgery is appropriate in advanced OA, it may be interpreted as an end-point representing disease progression. Time to total joint replacement has been evaluated in trials evaluating the efficacy of medications with potential structure-modifying effect, including symptomatic slow acting drugs in osteoarthritis[26, 53, 162], but may also be applied in studies evaluating long-term results after non-surgical, non-pharmacological treatment modalities[181].

Measuring physical activity

Physical inactivity has been referred to as the greatest public health problem of our time [219]. Engagement in moderate to vigorous physical activity for at least 30 minutes per day is thus recommended[99, 174, 239]. Patients with knee and hip OA are found to be less physically active than the general population, with only 14-30 % fulfilling these recommendations[61, 105, 203]. However, the importance of exercise and physical activity in OA should be particularly emphasized because of its beneficial impact on symptoms and function[55]. Advice on physical activity and exercise are already incorporated in most clinical guidelines for management of lower limb OA[169]. However, the dose-response relationship of exercise therapy interventions is hard to establish, partly due to the difficulties in measuring the amount and intensity of exercise and physical activity. For the same reason it is challenging to evaluate if interventions aiming to increase physical activity actually induce changes. Thus, valid and reliable methods for measuring physical activity represent a key component in studying its health effects. Furthermore, dimensions of physical activity including frequency, duration and intensity are important when the effect of physical activity is evaluated[211].

Several methods for assessing physical activity are currently available. They can be categorised into three main groups; Self-reported methods (questionnaires, diaries); activity monitors (pedometers, accelerometers, heart rate monitors); and direct measures of energy expenditure (oxygen consumption, doubly labelled water method). Self-reported outcome measures are frequently used, as they are inexpensive and easy to administer in large clinical trials and epidemiological studies. They have the potential to capture intensity, duration, and frequency of activity, and also allow specification of type of exercise. However, their qualitative attributes and measurement properties are questioned[212], with the validity and reliability being potentially hampered by recall and reporting bias. In contrast, accelerometers represent a method for measuring body acceleration, which allows for quantification of both the amount, duration and intensity of movement[77]. Unfortunately, they are rather expensive, and can only capture data for the particular period they are worn. Thus, accelerometers are less frequently used in larger clinical trials with repeated follow-ups. Direct measure of energy expenditure by using the doubly labelled water method is often considered to be the gold-standard for measuring energy expenditure/physical activity[213]. However, it is a time-consuming, costly method which requires access to both technical expertise and equipment, and is therefore rarely used in clinical trials.

To summarise, hip OA is a common, potentially disabling disease which follow a multifaceted long-term course. A shift towards non-surgical treatment modalities has been advocated, and exercise is recommended as part of the core treatment in hip OA. Nevertheless, few RCTs evaluating the effect of exercise treatment in hip OA have been conducted, and knowledge on the long term effect of exercise is particularly sparse. Furthermore, it is unknown if exercise may serve as a secondary preventive factor, which can delay disease progression. The studies included in this thesis were conducted to enhance knowledge and explore possible causal relations in the field of effect of exercise in patients with hip OA.

AIMS OF THE THESIS

General aim

The overall aim of this thesis was to evaluate the long-term effects of exercise therapy in addition to patient education (exercise therapy group) compared to patient education only (control group) in patients with hip OA.

Specific aims

I To evaluate the long-term results (29 months) in the exercise therapy group compared to the control group by selected outcome measures of impairments and activity limitations, including hip ROM, muscle strength, walking capacity and pain during walking. (Paper I)

II To evaluate the cumulative 6-year survival of the native hip to THR in the exercise therapy group compared to the control group. (Paper II)

III To evaluate the long-term results (29 months) in the exercise therapy group compared to the control group in self-reported pain, stiffness and physical function. (Paper II)

IV To evaluate the test-retest reliability and the construct validity of the PASE, a questionnaire assessing level and intensity of physical activity, in patients with hip OA. (Paper III)

The following null hypotheses were generated:

I There is no difference between the exercise therapy group and the control group over the 29 month follow-up period in hip ROM, muscle strength, walking capacity and pain during walking. (Paper I)

II There is no difference between the exercise therapy group and the control group in the 6-year cumulative survival of the native hip to THR. (Paper II)

MATERIAL AND METHODS

Ethical considerations

All patients included in this thesis signed a written informed consent prior to inclusion. The informed consent included information about the background and the overall aim of the study interventions, in addition to the randomisation process, the follow-ups and outcome measures included. The patients were also informed that no health risks were associated with participation in the study, but that the exercise sessions might induce a temporary increase in pain. They were informed that they could withdraw their consent at any time without giving any reason. Based on the fact that all patients diagnosed with hip OA are recommended to be provided information about the disease, we considered the inclusion of a true control group receiving no treatment to be unethical. Thus, all patients included in the RCT participated in a structured patient education program. Therefore, instead of comparing exercise therapy to a control group receiving 'no treatment', we compared exercise therapy *and* patient education to patient education *only*.

Patients participating in the reliability and validity study (Paper III) signed an additional written informed consent including information about the purpose of the study, the study procedure and the assessments.

The study was approved by the Regional medical research Ethics Committee, and the patients' rights were protected as the study was carried out in compliance with the Helsinki Declaration.

Study design

The study was designed as a randomised, parallel group, single-blinded, controlled trial with repeated follow-up assessments (Figure 1). Paper I and Paper II were both based on the RCT design.

Additionally, a study was conducted to evaluate test-retest reliability and construct validity of the PASE (Paper III).

Subjects

The inclusion and exclusion criteria for the RCT was selected to ensure that the patients enrolled in the study had clinical and radiographic hip OA, without concomitant back or knee pain. Inclusion criteria comprised age between 40 and 80 years, hip pain for at least three months, radiographically verified hip OA, and an Harris Hip Score (HHS) between 60 and 95 points. Radiographic evidence of OA was based on assessment of the MJS according to Danielsson's criteria[40], i.e. MJS <3 mm in patients aged 70 years and older, MJS <4 mm in patients younger than 70 years, or >1 mm difference in MJS between the hip joints. Clinical symptoms of hip OA was defined as presence of hip pain for three

months or longer and an HHS ≤95 points. The HHS is a clinician administered, disease-specific tool to evaluate hip pain, hip function and hip ROM, which is rated worst to best on a 0-100 scale[97]. An HHS below 60 is considered to reflect severe hip OA, and has been used as a criterion for THR surgery at our institution. Therefore, an HHS between 60 and 95 points, considered to represent hip OA with mild to moderate symptoms, was used as an inclusion criterion in the trial. In patients with bilateral hip pain, the most painful joint was defined as the index joint. Patients were excluded from participation in the trial if they previously had undergone THR in the index joint, or if they had non-idiopathic hip OA, knee pain or knee OA, low back pain, or dysfunction in the lower extremities due to accident or disease. Furthermore, patients who had been diagnosed with rheumatoid arthritis, osteoporosis, cancer, or cardiovascular disease unable to tolerate exercise, were pregnant, or did not understand Norwegian were excluded from participation.

One hundred and nine patients were included in the RCT (Paper I and II). A sub group of 40 patients, recruited from the RCT, was included in the study evaluating the reliability and validity of the PASE (Paper III). In this study, patients who had undergone THR in the index joint in the period between inclusion in the RCT and September 2010 were excluded from participation.

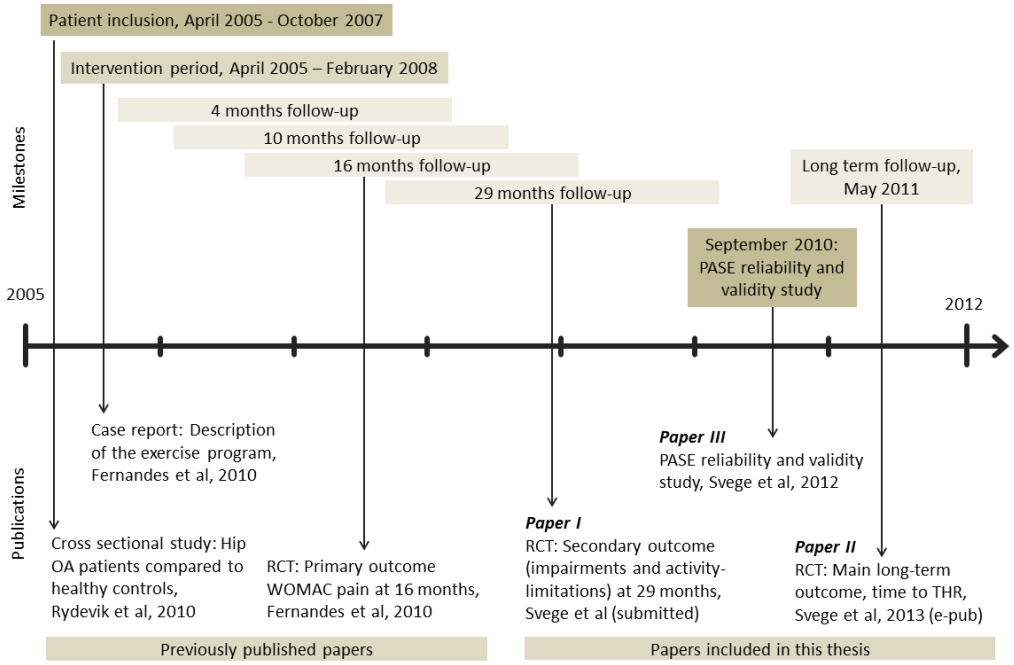

Figure 1. Time-line showing the milestones for the complete study from 2005 to 2012, and the publications originating from the study.

Recruitment and screening for inclusion

The majority of the patients were recruited from the outpatient clinic at the Department of Orthopaedics, Oslo University Hospital. To enhance enrolment Diakonhjemmet Hospital, the Norwegian Sports Medicine Clinic (NIMI), and general medical practitioners within the Oslo region assisted in patient recruitment. Furthermore, an advertisement was published in a local newspaper.

All screening for inclusion was conducted at the outpatient clinic at the Department of Orthopaedics, Oslo University Hospital from April 2005 to October 2007. Radiographic assessments were conducted at two radiographic clinics, or radiographs were obtained from the patient if he or she had gone through radiographic examination during the past year. All the standardised posteroanterior pelvic digitised radiographs were examined and the MJS was measured by one orthopaedic surgeon. Screening for eligibility in the study was performed by one trained physical therapist. Based on the inclusion and exclusion criteria the clinical assessment comprised of the HHS, questions on pain of the hip, knee, ankle and back, and questions on co-morbidities and additional lower limb dysfunction.

Screening and inclusion of patients for the reliability and validity study was conducted in September 2010. Sixty-one patients participating in the RCT were contacted by telephone, and patients who were able and willing to participate, and who had not undergone THR in the index joint, were included in the study.

Randomisation process

The patients were randomly allocated to 1) a group receiving exercise therapy in addition to patient education (exercise therapy group), or 2) a group receiving patient education only (control group) (Figure 2). A computer generated randomisation list (block length 10, allocation ratio 1:1) was drawn by a statistician prior to inclusion. Sequentially numbered, sealed envelopes were used to assign treatment for the included patients consecutively by a research coordinator not involved in the patient assessment or interventions. Allocation concealment was maintained until written informed consent was obtained. Assignment to treatment group was conducted after completion of the baseline assessments and patient education sessions. The randomisation sequence was also concealed from the study collaborators until treatment was assigned.

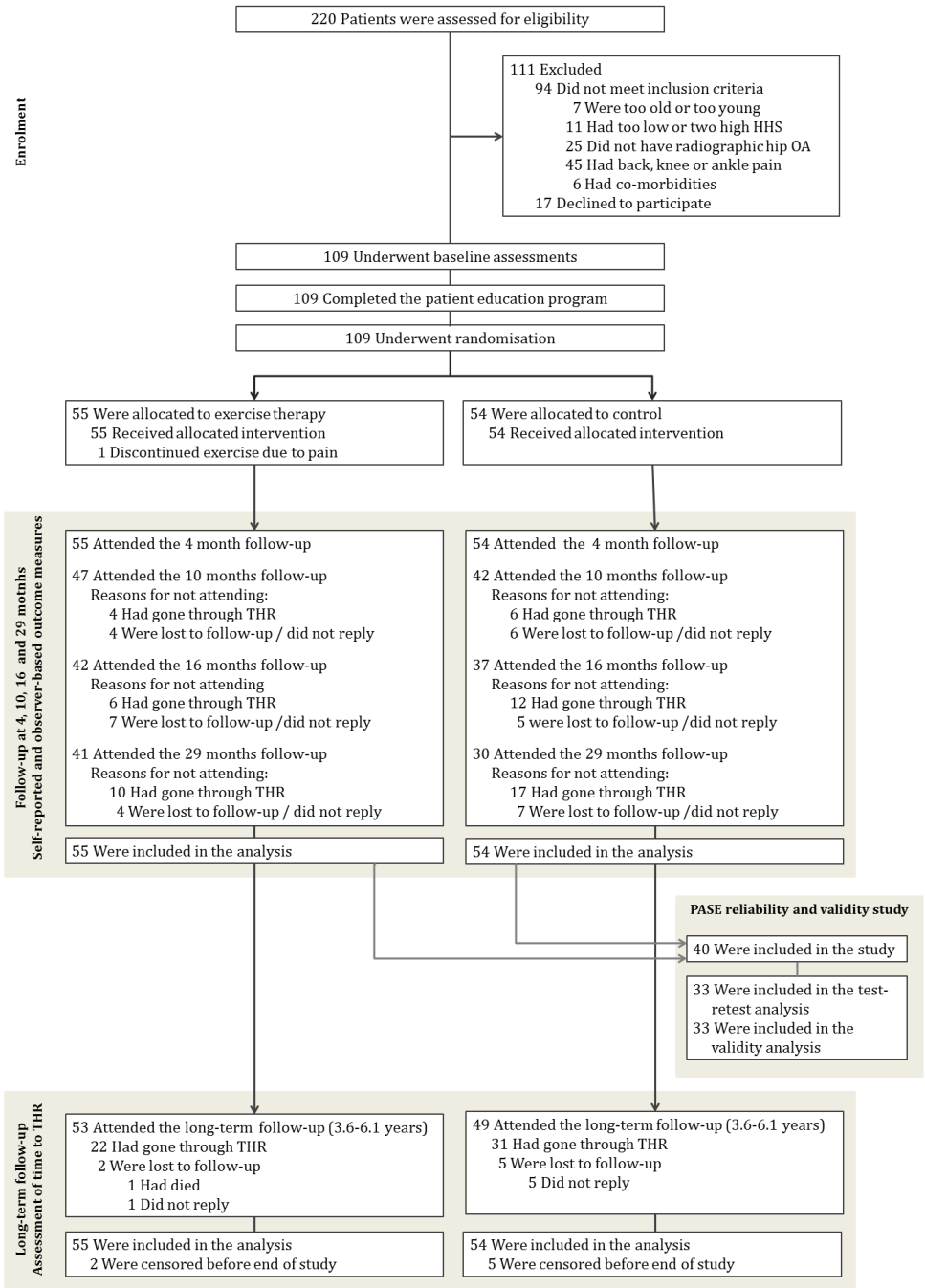

Figure 2. Flow-chart of enrolment, allocation, follow-up, and analysis of the patients included in the study.

Interventions

Patient education

All patients included in the RCT participated in a structured patient education program. The patient education program was given after the baseline assessments had been conducted, but prior to randomisation. The patient education program applied was developed particularly for patients with hip OA and has previously been described by Klassbo et al.[132].

The patient education was led by a physical therapist and given as three group sessions of 1.5 hours over three weeks at NIMI. Oral information on OA was provided, accompanied by slides with text, photos and illustrations. The patients were encouraged to ask questions during the lessons, and topics of interest were discussed. The aim of the program was to reduce fear of pain, and to educate the patients about pain management, how to prevent or reduce impairments, and how to improve or maintain physical ability. Key elements at the first group meeting were information about hip anatomy, pathogenesis and diagnosis of hip OA, and clinical symptoms and natural course of the disease. At the second meeting the patients were given information about possible reasons for functional impairments, hereunder impaired muscle strength and restricted ROM, they were encouraged to be physically active, to preserve or enhance ROM, and a sheet describing therapeutic exercises were provided. The third meeting focused on pain and techniques for self-management of pain, and information on additional treatment options, including physical therapy, pharmacological treatment and surgical options, were given.

Patients who were randomised to the control group attended a visit to a physical therapist two months after finishing the patient education program at NIMI. A two month visit at a physical therapist was originally described as part of the patient education program[132], but was not provided to the patients randomised to the exercise therapy group because they already attended weekly sessions with a physical therapist through the exercise therapy program.

Exercise therapy

The applied exercise therapy program was developed within our research group and has previously been published by Fernandes et al.[64]. It was especially designed to target impairments and functional disability in patients with hip OA, aiming to reduce pain, and improve muscle strength, flexibility and overall physical function. Thus, an exercise pool consisting of 26 exercises on warm-up, resistance exercises, functional exercises and stretching/flexibility exercises was composed.

Warm-up was conducted by 5-10 minutes of cycling on a stationary bike or walking on a treadmill at an intensity of 12-13 on the Borg Rating of Perceived Exertion Scale[25]. Ten strength training exercises for hip and core muscles, including leg extension, leg curl, hip

extension, hip abduction, heel-raise, crunches, and bridging were included in the program. Dosage, progression and frequency of the strengthening exercises were based on the recommendations by the American College of Sports Medicine for exercise in healthy adults[133]. Thus, three sets of eight repetitions, aiming at 70-80% of one repetition maximum (1RM), were prescribed in the program, and the resistance was to be increased when the patient could perform more than eight repetitions. This was also in accordance with later described and published recommendations based on a meta-analysis evaluating the dose-response relationship of resistance training in older adults[205]. Seven functional exercises, including squats, single leg stance/squats, lunges, sideway lunges and step-up/step-down, were further included in the exercise program. The exercises were chosen to imitate movement during typical activities of daily living. Movement quality was emphasised during execution of the functional exercises. Three sets of ten repetitions were used, and progression was assured by increasing the level of difficulty of each exercise when the patients were able to perform ten repetitions or more with acceptable movement quality. Finally, two flexibility exercises and four stretching exercises were included in the exercise program. The flexibility exercises were performed with the patient lying on a mat with one leg suspended in a sling. Relaxed, repetitive movements of the leg in hip flexion and extension and in hip abduction and adduction were conducted for two minutes in each direction. The four static stretching exercises comprised extension, abduction, internal rotation and external rotation, and the patients were instructed to hold the static stretch for 30 seconds at the end position in each direction.

Each exercise session consisted of a minimum of eight resistance and functional exercises, and lasted approximately one hour and 15 minutes. The patients in the exercise therapy group were to perform two to three weekly exercise sessions for 12 weeks. At least once weekly they were supervised by a physical therapist. The physical therapist assisted the patients in choosing exercises from the pool of 26 exercises, setting the initial resistance and regulating the progression. Furthermore, a pain scale was introduced by the physical therapist to assist the patient in deciding when to modify the exercise or intensity if unacceptable pain occurred during an exercise. The patients were informed that exercising could be somewhat pain provoking, but that pain should not exceed the limit for acceptable pain, defined as ≤3 on a VAS for pain. Compliance was based on training diaries filled in weekly by the patients during the 12 weeks intervention period. Attending at least 20 of a total of 24 sessions (≥80 %) was defined as satisfactory adherence. Patients in the control group did not have access to the exercise therapy program during the intervention period.

Baseline assessments

All baseline assessments were carried out prior to the patient education program and randomisation process. Patient characteristics included age, gender, height, weight, work status, education level, uni- or bilateral hip pain, pain duration, MJS, and HHS.

Outcome measures

The outcome measures used in this thesis included measures of impairments and activity limitations (Paper I), time to THR (Paper II), a disease specific questionnaire (Paper II), physical activity questionnaires (Paper I, II, III), and accelerometer registrations of physical activity (Paper III) (Table 1).

The primary end point in the RCT was the previously published 16-months outcome of the WOMAC pain subscale[65]. Thus, the outcome measures included in this thesis, used to evaluate long-term outcome of the RCT, are to be regarded as secondary outcome measures in the RCT.

Table 1. Outcome measures in Paper I, II and III

Paper I-III	**Included outcome measures**
Paper I Effect of exercise therapy and patient education on impairments and activity limitations in patients with hip osteoarthritis. A randomized controlled trial with 29 months follow-up	Impairments and activity-limitations • Hip range of motion • Isokinetic muscle strength tests of knee and hip flexion and extension • Six minute walk test Activity-related pain • Visual analogue scale for pain during walking Physical activity level • Physical Activity Scale for the Elderly
Paper II Exercise therapy may postpone total hip replacement surgery in patients with hip osteoarthritis: a long-term follow-up of a randomised trial	Main long-term outcome • Survival of the native hip to total hip replacement Self-reported pain, stiffness and physical function • Western Ontario and McMasters Universities Osteoarthritis Index 3.1™ Physical activity level • Physical Activity Scale for the Elderly
Paper III Reliability and validity of the Physical Activity Scale for the Elderly (PASE) in patients with hip osteoarthritis	Physical activity level and intensity • Physical Activity Scale for the Elderly • International Physical Activity Questionnaire, short form • Actigraph GT1M accelerometer

Impairments and Activity limitations (Paper I)

The outcome measures of impairments and activity limitations included hip ROM, isokinetic concentric muscle strength testing of knee and hip flexion and extension, the 6MWT, and pain during walking assessed on a VAS for pain. Assessments were conducted

at the follow-ups at four, ten and 29 months after inclusion. Assessment at baseline, four months and ten months were conducted at the Norwegian School of Sports Sciences. The 29-month follow-up was conducted at NIMI. The outcome assessors remained blinded to group allocation throughout the trial.

Passive hip joint ROM was measured by using a half-circle 1°-increment plastic goniometer with a moveable arm. Flexion, abduction and adduction were measured in supine position, with the opposite thigh fixated[173]. Internal and external rotation was measured in prone position, with the hip extended and the knee 90° flexed[173]. Hip extension was measured by using the modified Thomas test[58, 173]. Fair to good intra-rater reliability have been demonstrated for goniometric measurements of hip ROM in patients with hip OA, with ICCs of 0.50-0.94[109, 131] for the different hip movements. Measurement error (Altman's repeatability) has been calculated to be 21.5° for the total hip ROM, and between 3.9° and 10.7° for the different hip movements, smallest for extension and largest for internal rotation[109]. The standard error of measurement (SEM) have been calculated to be between 2.3° and 5.3° for the different hip movements, giving MDC ranging from 6.4° to 14.7°, smallest for extension and largest for external rotation[131]. Inter-rater reliability have been found to be poorer with ICCs below 0.73[184, 215].

Isokinetic muscle strength of hip and knee flexion and extension was tested by using an isokinetic dynamometer at baseline, four months follow-up, ten months follow-up (REV 9000; Technogym SpA, Gambettola, Italy), and 29 months follow-up (Biodex 6000, Biodex Medical Systems Inc., Shirley, New York, US). Tests of knee muscle strength were conducted with the patient in a sitting position, with straps fixating the trunk and thigh. Knee joint testing range was set to 5-85° of knee flexion. Hip muscle strength was tested with the patient in a supine position, with fixation over the pelvis, and the opposite leg extended at the knee and hip. Hip joint testing range was set to 35-75° of hip flexion. The angular velocity was set to 60° per second (60°/s) for both knee and hip muscle strength testing. The test protocol included a warm-up of four repetitions, followed by a 20 second rest prior to the five test repetitions. For the analysis the highest peak torque value of the five test repetitions, measured in Newton-meter (N-m), was used. For isokinetic knee muscle strength testing ICCs between 0.92 and 0.94 have been reported for both flexion and extension[30, 127], with a MDC of 22 N-m for knee extension[148]. In patients with hip OA ICCs of 0.89 and 0.75 have been reported for isokinetic hip flexion and hip extension muscle strength, respectively[13].

The 6MWT[59] was conducted with the patients walking back and forth in a 20 m long corridor. Patients were instructed to walk as far as possible, without running, within six minutes. The walking distance was registered in meters and time was monitored by using a stopwatch. The 6MWT has been found to have acceptable reliability with an ICC of 0.94 and a MDC of 34-61 m in patients with OA of the hip or knee[117, 129], and a responsiveness superior to other functional tests and the WOMAC physical function subscale[80, 129]. The

MCID for the 6MWT has not been determined in patients with OA, but Pereira et al.[178] have estimated a substantial meaningful change to be 50 m in older adults.

Immediately following the 6MWT the patients were asked to score their hip pain during walking on a VAS ranging from 0 to 100 mm, with 0 representing no pain and 100 representing extreme pain. The MCID for improvement has been estimated to be -15.3 mm in patients with hip OA[221]. However, as it was found to be influenced by symptom severity, the threshold for minimal clinical improvement was estimated to be -7.2 mm in patients with mild to moderate symptoms[221].

Total hip replacement surgery (Paper II)

Based on the consensus that THR is appropriate only in cases of advanced disease, it may be interpreted as an expression of disease progression. Thus, time to THR was used as the main outcome measure at 3.6-6.1 year follow up, to compare the survival of the native hip to THR between the two groups (Paper II). All patients were instructed to consecutively report if and when they went through THR surgery during follow-up. Additionally, data on THR was recorded at the follow-ups conducted four, ten, 16 and 29 months after inclusion, and by contacting all patients by telephone in April and May, 2011. The mean time from inclusion to May 15th 2011 (end of study) was 4.8 years, ranging from 3.6 to 6.1 years. The outcome assessor was blinded to group allocation.

According to Abadie et al.[1], THR is a highly relevant clinical outcome, but as no clearly defined criteria for when THR is to be conducted exists, it is potentially biased by non-disease related factors. Total hip replacement is suggested to be a sensitive outcome measure, but its specificity and responsiveness has been questioned[7].

Self-reported pain, stiffness and physical function (Paper II)

Self-reported pain, stiffness and physical function was assessed by the WOMAC. It is a self-administered, disease-specific questionnaire comprising three domains, pain, stiffness and physical function. A score for each domain is calculated separately, expressed best to worst on a 0-100 scale[17]. Calculation of the WOMAC subscale for pain is based on five items, with each item scored from 'no pain' to 'extreme pain' on a VAS. The score of each item was measured in mm by a ruler, with the average of the five items representing the total score for the WOMAC pain subscale. The different items addressed pain while 1) walking on a flat surface, 2) going up or down stairs, 3) in bed at night, 4) sitting or lying down, and 5) standing upright. Calculation of the WOMAC subscale for stiffness is based on two items addressing sensation of joint stiffness while 1) getting up in the morning and 2) after sitting for 30 minutes or more. Each item is scored from 'no stiffness' to 'extreme stiffness' on a VAS, and the average of the two items represents the total score for the WOMAC stiffness subscale. Calculation of the WOMAC subscale for physical function is based on 17 items addressing difficulties experienced while performing different activities, with each item scored from 'not difficult' to 'extremely

difficult' on a VAS. The average of the 17 items represents the total score for the WOMAC physical function subscale.

The WOMAC has been proven to be a reliable and valid outcome measure in patients with lower limb OA, with adequate internal consistency demonstrated by Chronbach's alpha coefficients of 0.81-0.96, acceptable test-retest reliability with ICCs between 0.43-0.93, Spearman's ρ of 0.87, construct validity estimates of 0.43-0.59, and adequate responsiveness with effect sizes of 1.0-4.8 for THR[17, 159, 201, 208]. According to Angst et al.[10, 11] effects larger than 12-22 % of the baseline score can be interpreted as a clinically meaningful change in patients with lower limb OA. The MCID has been estimated to be 7.5 for the pain subscale and 6.7-7.9 for the physical function subscale[11, 221]. The MCID was found to be affected by the initial degree of severity of the symptoms. Thus, the MCID for the WOMAC physical function subscale in patients with mild to moderate symptoms was estimated to be 5.3[221].

Physical activity assessments (Paper I, II and III)

The PASE was used to assess overall activity level at baseline, and at the four, ten, 16 and 29 months follow-up of the RCT (Paper I and II). Furthermore, the number of weekly exercise sessions was assessed at baseline and at the four months follow up. Additionally, the test-retest reliability and the construct validity of the PASE were investigated in patients with hip OA, since that had not previously been reported (Paper III). In the reliability and validity study we also included other outcome measures intended to assess physical activity level and intensity. The IPAQ was included as well as an Actigraph GT1M accelerometer (Acti-Graph, LLC, Pensacola, FL, USA) to examine if the PASE could validly assess physical activity level and intensity.

The PASE is a self-administered, 7-day recall questionnaire designed to assess physical activity in older adults[234]. The Norwegian version of the PASE, which was used in this trial, has been slightly adapted during the translation process due to cultural differences[153]. Thus, the Norwegian version consists of 24 questions giving an overall PASE score ranging from 0-315, with 0 representing no physical activity and 315 representing an extremely high activity level. The questions included in PASE address leisure-time, household and work-related physical activity. Participation in leisure-time physical activity, including light, moderate and vigorous intensity, and strengthening activities, is recorded as never, seldom (1-2 days per week), sometimes (3-4 days per week), and often (5-7 days per week). Duration is categorised as less than 1 hour, 1-2 hours, 2-4 hours and more than 4 hours. Housework activities are recorded as yes or no, and paid or unpaid work, requiring some physical activity, is recorded in hours per week. The total PASE score was computed by multiplying time spent (hours per day) in each leisure and work-related activities or participation (yes/no) in household related activities, by empirically derived weighting, and then summarising all items. Additionally, the score for household-/work-related activities, the score for leisure-time physical activity, and the score of the items addressing light, moderate and vigorous intensity

were calculated separately. To evaluate its test-retest reliability, the PASE was filled in by the patients included in the reliability and validity study at two occasions with seven days in between.

The IPAQ is a self-administered, 7-day recall questionnaire designed for assessing physical activity in adults[35]. The Norwegian version of the IPAQ short form was used in this study (http://www.ipaq.ki.se/). It consists of seven questions which include physical activity in all contexts of everyday life, addressing time spent on vigorous physical activity, moderate physical activity and walking. The IPAQ is scored by using the Metabolic Equivalent of Task (MET) method, where different activities and levels of intensity are given different MET estimates. Total MET-minutes per week, representing the overall activity level, was calculated. In addition, MET-minutes per week for walking activities, moderate activities and vigorous activities were calculated. The IPAQ was included in the reliability and validity study to evaluate construct validity of the PASE, and was thus only assessed at the first test occasion for the patients participating in this study.

The Actigraph GT1M (ActiGraph, LLC, Pensacola, FL, USA) is an electronic motion sensor comprising a single plane accelerometer. Movement in the vertical plane is detected as a combined function of the frequency and intensity of the movement. For the analyses, a valid day was defined as having ten or more hours of monitor wear. Six or more valid days of registration were considered sufficient. The software program ActiLife (ActiGraph, LLC, Pensacola, FL, US) was used to initialise the Actigraphp GT1M and to download the recordings, and the software program SCA Analyzer (SAS Institute Inc., Cary, North Carolina, USA) were used to reduce data. Average counts per minute representing the overall activity level were calculated. In addition, the minutes spent in light, moderate and vigorous physical activity intensity was calculated, based on the cut-offs described by Troiano[217] and Hansen et al.[95]. Furthermore, the average minutes spent in moderate to vigorous physical activity per day were calculated, to evaluate the proportion of patients who fulfilled the recommendations of at least 30 minutes of daily moderate to vigorous physical activity. The accelerometer recordings was used to evaluate the construct validity of the PASE. Therefore, the patients participating in the reliability and validity study to evaluate the measurement properties of the PASE, completed a seven-day Actigraph GT1M measurement period prior to filling in the PASE.

Sample size calculations

To evaluate sample size, a clinically relevant outcome measure must be determined prior to study start, together with the expected variance and the minimum difference considered to be clinically important for the specific outcome measure, as well as the risk for type 1 and type II errors (risk for a false positive or a false negative conclusion). The WOMAC subscale for pain was defined as the primary outcome measure for the short term outcome[65], and risk levels for type I and type II errors were set at 0.05 and 0.10, respectively. Additionally, the standard deviation (SD) for the change of WOMAC pain

score was set to 23 mm and the mean group difference was set to 15 mm. The sample size calculations revealed that 49 patients were needed in each group in the 16 months follow-up of the RCT, published by Fernandes et al.[65] (ClinicalTrials.gov number NCT00319423). To compensate for a possible drop-out rate of 10 %, inclusion of at least 54 patients in each group was considered sufficient[65]. For the long-term outcome of the RCT, no new a priori statistical power analysis was performed. However, post power analyses were conducted for the outcome measures included in Paper I, and included in the Discussion section of this thesis.

Developing new knowledge based on innovative ideas is considered a primary task for research. Hence, we decided to conduct a long-term follow-up study of the RCT based on the surprisingly high number of patients who had gone through THR after just 16 months[65]. An additional protocol was prepared in which we specified that the patients were to be followed for minimum 3.5 years to assess time to THR (Paper II) (ClinicalTrials.gov number NCT01338532). Since these patients were included over a 2.5 year period, the longest follow-up would exceed six years.

Sample size for the reliability and validity study (Paper I) was based on previous studies [70, 233] and on calculations suggesting that inclusion of 35 patients would provide correlation estimates between 0.64 and 0.89. Also, according to the model for sample size requirements by Donner and Eliasziw[51] inclusion of 25-30 subjects seemed to be sufficient, based on an estimated correlation >0.85 between two measurements. Hence, we decided to include 40 patients in the study.

Statistical analyses

The statistical analyses included descriptive analyses (Paper I-III), mean comparison within and/or between groups (Paper I-III), group comparisons by mixed model analysis (Paper I and II), Kaplan-Meier analysis (Paper II), group comparisons by Log Rank test and Cox-proportional hazard model (Paper II), two-ways analysis of variance (Paper III) and bivariate correlation analysis (Spearman's ρ) (Paper III).

All analyses were performed by IBM SPSS Statistics, version 19.0 (IBM Corp., Somers, New York, USA), except for Paper III where the PASW Statistics, version 18 (IBM Corp., Somers, New York, USA) was used. Descriptive statistics were presented as mean and SD or frequency and percentage. P-values below 0.05 were considered statistically significant for all analyses.

Paper I

To evaluate between-group differences in ROM, muscle strength, walking capacity and pain during walking, a linear mixed model (variance component structure with time and the interaction between time and group as fixed effects and time as random effect intercept and slope) was used. The analyses were based on the intention to treat (ITT) principle. Because many patients had undergone THR during follow-up, and thus were

not assessed at the subsequent follow-ups, a missing value analysis was conducted to evaluate the pattern of missing values. The analysis revealed that missing was not completely at random. Thus, as a sensitivity analysis, a linear mixed model using the last observation carried forward technique (LOCF) was conducted to account for the poorer pre-operative results of those patients who had gone through THR during follow-up. To evaluate within-group differences from baseline to four month follow-up paired t-tests were applied. A linear mixed model was used to compare total PASE scores between the exercise therapy group and the control group. Weekly number of exercise sessions was reported as mean and SD.

Paper II

A Kaplan-Meier survival analysis was constructed to evaluate cumulative six year survival, and group difference was assessed with the Log Rank test. The analyses were based on the ITT principle. Patients were followed until time of THR in the index joint, or until death, drop-out or end of study. THR in the index joint was defined as event, while patients who were lost to follow-up, were deceased or were followed until end of study were treated as censored in the analysis. Time to THR was reported as median and 95 % confidence interval (CI). A Cox-proportional hazard model was used to calculate hazard ratio and 95 % CI between groups. No adjusted analysis was conducted, due to equality of groups at baseline.

A linear mixed model (variance component structure with time and the interaction of time and group as fixed effects and time as random effect intercept and slope) was used to compare WOMAC scores between the exercise therapy group and the control group over the 29-month follow-up period. A linear mixed model was also applied to compare WOMAC scores prior to THR surgery or end of study between patients who went through THR and patients who did not. A linear mixed model was also used to compare total PASE scores between the exercise therapy group and the control group. Weekly number of exercise sessions was reported as mean and SD.

Paper III

Test-retest reliability of the PASE score was evaluated by calculating the ICC (two-way random effect model, absolute agreement). The SEM, the MDC and the limits of agreement (LoA) were calculated to evaluate measurement error. Additionally, a Bland Altman plot was conducted for visual judgment of the relationship between the individual mean total PASE score of the test and retest, and the difference in total PASE score between test and retest[23].

The construct validity of the PASE was evaluated by calculating the Spearman's ρ between the total PASE score and the Actigraph GT1M total counts per minute, and between the total PASE score and the IPAQ total MET-minutes per week. A priori, we hypothesised a low to moderate positive correlation (ρ between 0.15 and 0.5) between the total PASE score and the Actigraph GT1M total counts per minute, and a moderate to strong positive correlation (ρ between 0.6 and 0.9) between the total PASE score and the

IPAQ total MET-minutes per week. A correlation coefficient above 0.50 between the total PASE score and the Actigraph GT1M counts per minute was defined as a cut-off for acceptable validity, as suggested by Terwee et al.[213]. Additionally, Spearman's ρ were calculated for the PASE items for light, moderate and vigorous PA intensity and the respective intensity categories of the Actigraph GT1M and the IPAQ.

SUMMARY OF RESULTS

Study I

Two hundred and twenty patients were screened for eligibility between April 2005 and October 2007. One hundred and nine patients (54 % women), with a mean age of 57.9 ± 9.9 years were included in the trial, and randomised to the exercise therapy group (n=55) or the control group (n=54)(Figure 2). The patients had experienced pain for 48.4 ± 51.9 months (median 30 months), and 70 % presented bilateral hip pain. The mean MJS at baseline was 2.0 ± 1.1 mm, and the mean HHS was 78.4 ± 8.1. The patients randomised to the exercise therapy group completed a median of 20 (interquartile range 16-24) exercise sessions over the 12-week period, with 53 % completing ≥20 exercise sessions. One patient discontinued exercise after three sessions due to increasing hip pain, but no other adverse events were registered.

At the 29 months follow-up eleven patients were lost to follow-up, three patients did not complete the clinical/functional tests, and 27 patients had gone through THR prior to the follow-up. Hence, 38 % of the patients presented missing data for the outcome measures of impairments and activity limitation at the 29 month follow-up. At the four month follow-up, immediately after completion of the exercise program, the exercise group had 5.4 (95 % CI from 0.5 to 10.3) degrees better external rotation ROM compared to the control group. No other significant group differences were found at the four month follow-up. Within the exercise therapy group, significant improvement from baseline to four months follow up was found for external rotation ROM, 6MWT, and pain on VAS during walking, and significant decline was found for hip flexion. Within the control group significant improvement was found for external rotation ROM, and significant decline was found for hip adduction ROM and for hip flexion muscle strength. No other within group differences was found between baseline results and four months results. Over the total follow-up period of 29 months, patients in the exercise group had significantly less pain on VAS during walking compared to the control group (p=0.018). No other significant between group differences were found. The results of the sensitivity analysis corresponded with the findings of the main analysis. Overall, the estimated means for each group improved or remained relatively stable over the follow-up period. No group differences were found for the total PASE score over the 29 months follow-up period (p=0.397). The mean number of self-reported exercise sessions per week was 3.2 and 3.7 at baseline and four months, respectively, in both groups.

This study shows that exercise therapy in addition to patient education provided no benefit over patient education only in hip ROM, knee and hip muscle strength and walking capacity, but the exercise therapy group reported less pain during walking in the long term.

Study II

Inclusion, baseline characteristics and compliance in Study II were identical with Study I. Data on whether THR had been performed or not was obtained from 102 patients, giving a response rate of 94 % at the 3.6-6.1 year follow-up. The remaining seven patients were treated as censored at the time of death, last follow-up, or contact during the follow-up period. Twenty-two patients in the exercise therapy group and 31 patients in the control group underwent THR within the 3.6 to 6.1 year follow-up period, giving a six year cumulative survival of the native hip of 41 % and 25 %, respectively (p=0.034). Cox proportional hazard analysis showed that participating in both exercise therapy and patient education had a protective effect against THR compared to patient education only (Hazard ratio=0.56, 95 % CI from 0.32 to 0.96, p=0.036). Median time to THR was 5.4 years in the exercise group and 3.5 years in the control group.

The exercise therapy group had better self-reported hip function prior to THR or end of study evaluated by the WOMAC physical function (p<0.01), but no significant difference were found for WOMAC pain (p=0.083) or WOMAC stiffness (p=0.112). There was no significant difference in PASE scores between the groups (p=0.397). The number of self-reported exercise sessions per week was equal across the groups, 3.2 at baseline and 3.7 at the four months follow-up. Mean MJS at baseline was 1.5 ±0.9 mm in patients who went through THR, compared to 2.5 ±1.0 mm in the patients who did not (p<0.01). At baseline, there was no significant difference between patients who went through THR and patients who did not, neither for WOMAC pain, stiffness or physical function. The 53 patients who underwent THR before end of study had worse preoperative score in all WOMAC subscales over the 29 months follow-up period compared to the patients who did not go through THR or were censored at end of study (p<0.01).

This study showed that participating in exercise therapy in addition to patient education may reduce or postpone the need for THR in patients with hip OA.

Study III

In 2010, 20 men and 20 women were recruited from the RCT for participation in a study evaluating the reliability and validity of the PASE. Hence, they presented patient characteristics similar to the RCT but were slightly older with a mean age of 61.3 ± 10.0 years because the study was conducted three years after the inclusion in the RCT was completed.

The test-retest reliability was calculated based on the 33 patients who had completed the PASE at both test and retest. Mean days between test and retest was nine days. Mean total PASE score was 143 ± 71 and 125 ± 56 at test and retest, respectively. The decline in the total PASE score from test to retest was significant (p = 0.02). The ICC for the total PASE score was 0.77 (95 % CI from 0.56 to 0.88). The estimated SEM and MDC were 31 and 87 points, respectively, and the calculated LoA ranged from -65 to 100. Furthermore, the ICC for household- and work activities was 0.69 (95 % CI from 0.46 to 0.84), and the ICC for

leisure-time activity was 0.53 (95 % CI from 0.24 to 0.74). The ICC for the different intensity items of the PASE was 0.46 (95 % CI from 0.15 to 0.69) for light intensity, 0.20 (95 % CI from -0.16 to 0.51) for moderate intensity, and 0.68 (95 % CI from 0.44 to 0.83) for vigorous intensity.

Construct validity was evaluated by calculating the Spearman's ρ between the PASE and the Actigraph GT1M registrations, which was based on the 33 patients who had readable Actigraph GT1M files with six days or more of registration. The comparisons between the PASE and the IPAQ, was based on the 25 patients who had completed the IPAQ without missing values. The Spearman's ρ between the total PASE score and the Actigraph GT1M total counts per minute and between the total PASE score and the IPAQ MET-minutes per week was 0.30 ($p=0.089$) and 0.61 ($p<0.01$), respectively. The Spearman's ρ between the PASE items assessing different levels of physical activity intensity and the respective categories of the Actigraph GT1M registrations ranged from 0.20 to 0.38, being highest for moderate intensity. The Spearman's ρ between the PASE items assessing different levels of physical activity intensity and the respective items of the IPAQ ranged from 0.29 to 0.75, being highest for vigorous intensity.

Based on the Actigraph GT1M measurements 67 % of the patients fulfilled the recommendations of at least 30 minutes of accumulated moderate to vigorous physical activity per day, and the mean time spent on moderate to vigorous physical activity was 45 ± 32 minutes per day.

These findings suggest that the PASE has moderate test-retest reliability, with acceptable ICC, but with large measurement error. The construct validity of the PASE was considered to be poor, based on a Spearman's ρ of 0.30. Hence, the PASE was not considered to be able to provide valid estimates of physical activity levels and intensity.

DISCUSSION

Main findings

Patients with mild to moderate hip OA who participated in exercise therapy and patient education revealed significantly less pain during walking over the 29-months follow-up period compared to patients participating in patient education only. However, no additional effect of exercise therapy was found for hip ROM, knee or hip muscle strength or walking capacity. Furthermore, we demonstrated a higher six year cumulative survival of the native hip to THR in the exercise therapy group compared to the control group, suggesting that exercise therapy can postpone the need for THR in patients with hip OA. Exercise therapy in addition to patient education significantly improved self-reported physical function, but no additional effect of exercise therapy was found for self-reported pain or stiffness. The reliability and validity study revealed that the PASE had poor construct validity, with a correlation coefficient of 0.30 when compared to an accelerometer. Thus, the use of PASE to assess physical activity level in patients with hip OA cannot be recommended, despite acceptable test-retest reliability.

Study design and methodological considerations

This thesis is based on three studies originating from a RCT evaluating the effect of exercise therapy and patient education in patients with hip OA. In an experimental design, the aim is to establish a cause-and-effect relationship by manipulating an independent variable to evaluate its effect on a dependent variable[194, 216]. Randomised controlled trials are required to study causality inferences, i.e. the link between the cause and its effect, as random assignment most likely ensure equality across groups[89]. As the aims of Paper I and Paper II were to evaluate the efficacy of two different treatment modalities, the RCT design was a natural choice.

All research designs have strengths and weaknesses influencing the validity of the research and the generalisability of the results. Internal validity concerns the degree to which all variables in a study has been successfully controlled for, to enable elimination of explanations other than the independent variable for the observed outcome[216]. The randomisation procedure aims at securing equal distribution of potentially confounding factors across the treatment groups, and can thus control for pre-study events, statistical regression, selection bias (for selecting treatment) and learning effect[216]. Furthermore, blinding of the outcome assessor can minimise observer bias. Inclusion of a group receiving 'no treatment' would allow comparison with the time-dependent course of the disease. As all patients were given the patient education program due to ethical reasons, the control group in our RCT, who received patient education only, cannot be considered to represent a 'true' control group, even though the expected benefits of the patient education are negligible[135]. Additionally, as no group receiving sham exercise was included in our RCT, we could not rule out a possible placebo effect.

Minimising the drop-out rate is considered to be crucial to ensure high internal validity. It has been suggested that a drop-out rate greater than 20 % may represent an important threat to the validity of the result[197]. At the 3.6-6.1 year follow-up (Paper II) the follow-up rate was excellent, as only six percent were lost to follow-up. Hence, the results of this analysis are less likely to be biased by drop-out. Outcome measures of impairments and activity limitations (Paper I) and patient-reported outcome measures (Paper II) was assessed at the four, ten, 16 (only patient-reported) and 29 month follow-up. Unfortunately, a relatively large percentage of the patients had missing observations at the latest follow-ups due to drop-out or THR. Consequently, overall power was reduced, which causes an increased risk for Type II error, i.e. to incorrectly retain the null hypothesis. The relatively large percentage of patients presenting with missing observations due to THR may have biased the results, as their presumably poorer pre-operative results were not included in the analysis. Furthermore, the uneven distribution of THRs across the groups, with more patients in the control group having gone through THR, may have resulted in an underestimation of the effect of exercise therapy.

The external validity addresses the question on generalisability, and is dependent on the process of patient recruitment and the eligibility criteria. The generalisability can be maximised by selecting participants, treatments, experimental situation, and tests that represent a larger population[216]. Selection bias may pose a threat to the external validity, and Altman et al[7] have suggested that patients recruited to non-surgical treatment trials may have a stronger desire to go through with non-surgical treatment strategies. To evaluate the external validity potential discrepancies between the study patients and the population from which it is drawn must be considered[216]. In RCTs, inclusion and exclusion criteria are typically strictly formed. The inclusion rate is often relatively low, due to both strict exclusion criteria and unwillingness to undergo randomisation. Low inclusion rate can be a threat to the external validity, especially if the excluded patients differ from the included. In this trial the inclusion rate was 50 %. Most patients were excluded because they lacked radiographic evidence of hip OA, had too high or too low HHS, or because they had knee pain or knee OA. Thus, our findings are applicable for patients with clinical and radiographic evidence of hip OA, presenting with mild to moderate symptoms, without concomitant knee pain, knee OA or low back pain. In a Norwegian population-based study 36 % of patients who reported to have been diagnosed with hip OA, also reported to have knee OA. Thus, by excluding patients with knee pain/knee OA we cannot generalise our findings to the approximately 1/3 of patients with hip OA who have concomitant knee pain/knee OA.

The interventions provided in a RCT may not have the same effect when given in another setting and by other providers. This also poses a threat to the external validity, which can be minimised by avoiding complex interventions and offer proper descriptions of the applied interventions. The content of both the patient education program and the exercise therapy program utilised in our study is easy to adopt in clinical settings, and detailed descriptions are readily available[64, 132]. According to Bagshaw and Bellomo [15],

actions aiming at optimising internal validity may compromise the overall generalisability, and advantages and disadvantages of the different actions should be carefully balanced.

Paper III was conducted to evaluate the test-retest reliability and the construct validity of the PASE in this particular subpopulation of hip OA patients, since the measurement properties of an outcome measure is considered to be related to the population and context in which it is being used[164]. The importance of similar conditions and appropriate time intervals between test and retest has been emphasized[213]. Thus, we administered the PASE by mail with seven days in between, and the patients completed the questionnaire at home. However, the week prior to the first PASE test occasion the patients had worn the Actigraph GT1M accelerometer. It is suggested that accelerometry may modify activity patterns[191], and the conditions of the seven day recall period for the first and second measurement occasion was therefore not fully identical. This was reflected by a significant decrease in PASE score from test to retest. We included both relative and absolute measures of reliability, as recommended by Terwee et al.[213]. To evaluate if the PASE truly measure the construct of physical activity level, we evaluated the construct validity by comparing it to outcome measures expected to measure the same construct. A correlation coefficient between the PASE and the Actigraph GT1M larger than 0.5 has been defined as acceptable validity[213]. Construct validity is often used to evaluate the validity of an instrument when no gold-standard is available to evaluate the criterion validity. However, the doubly labelled water method is considered to represent a gold standard for assessing energy expenditure/physical activity, and it could have been used as a comparison to establish the criterion validity of the PASE. However, we did not consider to include it in this study because it is too expensive, time-consuming and expertise demanding, and because it measures energy expenditure rather than level and intensity of physical activity.

Subjects

An age between 40 to 80 years was chosen as an inclusion criterion to reflect the age group with the highest incidence/prevalence of hip OA[39, 90, 115, 116, 187, 245]. The criteria by Danielsson[40] was used to assess radiographic evidence of hip OA. By using these criteria, which are less stringent compared to other suggested cut-offs[36], we aimed to include patients at an early stage of the disease, for which the exercise and educational interventions were considered to be particularly suitable. Clinical evidence of hip OA was based on the presence of hip pain for three months or longer and a HHS between 60 and 95 points. Patients with a HHS above 95 and below 60 were excluded from participation because their hip symptoms were considered to be minimal (>95) or severe enough for the patients to be considered as candidates for THR surgery (≤60).
Thus, the 109 included patients had radiographic OA, with mild to moderate clinical symptoms, and none of them were candidates for THR surgery at time of inclusion. They had no concomitant knee OA/knee pain or low-back pain, and no co-morbidities potentially hampering compliance to the interventions. Only three previous studies have

evaluated the effect of exercise interventions in patients with hip OA only[79, 124, 209]. Of these, only the study by Tak et al.[209] had been published prior to the start of our RCT. Hence, the inclusion and exclusion criteria were carefully chosen to enable us to contribute to increase the level of evidence for the effect of exercise therapy in this specific patient group.

The baseline characteristics of the patients included in this trial revealed that they were younger, had lower BMI, and a larger proportion of them were men, still employed and higher educated, compared to the patients included in other studies evaluating the effect of exercise treatment in hip OA[78, 124, 209]. Furthermore, they presented WOMAC pain, stiffness and physical function scores of 26.6, 34.6, and 22.3, respectively, indicating that they had mild to moderate pain and functional disability at the time of inclusion. Additionally, at baseline the patients reported a mean of as many as 3.2 exercise sessions per week, and their baseline activity level evaluated by the PASE was similar to healthy elderly[153]. Furthermore, the results of the Actigraph GT1M accelerometer recordings in the reliability and validity study of the PASE further emphasised that the patients included in our study were more physically active compared to what has been reported in other studies evaluating physical activity level in patients with OA[61, 105, 203]. Sixty-seven percent of the patients fulfilled the recommendations of 30 minutes or more of aggregated moderate to vigorous activity per day. Thus, a larger percentage of the patients in our study fulfilled the physical activity recommendations compared to the general adult and elderly population of Norway[95].

Interventions

Both interventions offered in this trial (a patient education program[132] and a supervised exercise therapy program[64]) were especially developed for patients with hip OA. As information is recommended to all patients with hip OA because it is assumed to be crucial to enhance self-management of the disease[34, 63, 106, 244], all study patients participated in the patient education program The patient education program focused also on the importance of physical activity, and provided suggestions on appropriate activity forms and exercises. Consequently, patients might have been encouraged to implement lifestyle changes, increase their physical activity level, or to adapt activity forms or exercises considered to be particularly valuable in hip OA. In a non-randomised trial, Klassbo et al.[132] demonstrated that patient who were given the same education program which we applied in our study, remained stable over a 12-month period, while the control group experienced increased pain and functional limitations. However, a recently published Cochrane-review concluded that self-management education programmes provide none or small benefits in patients with OA, but that they are unlikely to cause any harm[135].

The supervised exercise therapy was provided in an individualised, semi-standardised manner[64] to the 55 patients in the exercise therapy group as outpatient treatment by one of two physical therapists at NIMI. In knee OA, an approach combining exercises to

increase strength, flexibility, and aerobic capacity has been recommended, but no conclusion regarding the most efficient type of exercise have been established for hip OA[223]. However, strengthening exercises of the lower limb were particularly emphasised in our program, and nine resistance exercises and eight functional exercises containing a strengthening component were included. The patients were to perform three sets of eight repetitions of the resistance exercises, aiming at a resistance of 70-80 % of 1RM, with 2-3 weekly exercise sessions were used. This was in line with the recommendations by Kraemer et al.[134] and Steib et al.[205]. Training diary data revealed that only 53 % of the patients had completed 20 or more exercise sessions during the 12 week period. Hence, many patients did not exercise frequently enough to increase muscle strength. Furthermore, pain during training may have resulted in inadequate training intensity. The two flexibility exercises and four stretching exercises was included in the exercise therapy program aiming to improve hip ROM. Difficulties in maintaining an adequate stretching position and too short holding time may have resulted in inadequate stimulus to induce an increase in ROM. Adding a manual component to the stretching exercises might have increased the effect[107].

Despite that patient education has been considered as a part of the core treatment in OA, providing information on the benefits of exercise is probably insufficient to achieve long lasting behavioural changes[135]. Hurley et al.[113] suggested that for patients with chronic, painful diseases, positive experiences with performing an exercise program is necessary to implement it and to believe in its effectiveness. In patients with knee pain, programs combining education and exercise improved the patients exercise beliefs and self-efficacy[113, 118], and patient interviews have suggested that these improvements reflect the patients understanding and belief in exercise as an effective self-management of the disease[113]. Thus, we cannot rule out a possible interaction effect of the two interventions for the patients in the exercise therapy group.

Outcome measures

The outcome measures included in this thesis comprised assessments of impairments and activity limitations, disease specific questionnaires, and survival of the native hip to THR. As the results regarding the primary outcome measure of the RCT, WOMAC pain at 16 months follow-up, have been previously published[65], the outcome measures presented in this thesis are to be considered as secondary, exploratory, pre-defined outcome measures.

In Paper I we reported the treatment effect on specific impairments and activity limitations, including hip ROM, isokinetic knee and hip muscle strength, 6MWT and pain on VAS during walking. These measures were chosen because they address body impairments and activity limitations found in patients with hip OA[14, 195], and had been targeted in the exercise therapy program[64]. After our study was initiated, the OARSI recently recommended a set of five performance-based tests to be used as a complement to patient-reported measures in patients with hip OA[50]. This set of tests comprised the 30

seconds chair-stand test, 40 meter fast-paced walk test, a stair-climb test, timed up-and-go test and the 6MWT. Of the performance-based tests recommended by OARSI[50], only the 6MWT was included in our RCT. Furthermore, we assessed pain during walking on a VAS immediately after the 6MWT because it was considered to represent activity-related pain. Pain in OA, particularly at an early stage of the disease, is typically described as being aggravated by activity and relieved by rest. Thus, we considered it to be of particular interest to assess activity-related pain in this patient group, presenting with mild to moderate pain and symptoms at inclusion. In contrast, in the WOMAC pain subscale three of five items is regarding pain at rest[17], and it might thus be less responsive for improvement in this particular patient group.

Time to THR was defined as the main outcome measure for the 3.6-6.1 year follow-up (Paper II). Previous studies have also assessed time to THR to evaluate the effect of exercise and potentially disease-modifiable pharmacological interventions in lower limb OA[26, 53, 162, 180]. However, the use of THR as an outcome measure is hampered by the lack of generally accepted criteria for THR. Higher levels of pain and functional disability have been found to be associated with a higher THR rate[151], but as clinical severity varies widely at the time of THR self-reported pain and function cannot identify patients who are in need of surgery[48, 88]. According to Abadie et al.[1] and Altman et al.[7], THR is a highly relevant clinical outcome, but it is potentially biased by comorbidities and contraindications for surgery, as well as the patient's preferred choice of treatment. Still, we consider the use of THR to evaluate long-term treatment effect in this trial to be appropriate, due to the randomisation procedure which presumably secure an even distribution of potentially confounding factors across the groups. Unfortunately, we did not specify criteria for when THR was indicated prior to the start of the study. Also, we did not assess MJS, HHS, pain level and indication for THR immediately prior to surgery. However, we assume that the majority of the patients who underwent THR during follow-up had an HHS below 60 points at the time of surgery. This assumption was based on 65 % of the patients being operated at Oslo University Hospital, where an HHS below 60 points has been interpreted as an indication for THR. This HHS cut-off has also been used at other hospitals. Furthermore, pain experience, including pain intensity, impact on quality of life and ability to cope with pain, has been identified as the main reason for patients considering themselves as candidates for joint replacement surgery[72].

The disease-specific WOMAC questionnaire was included in Paper II with the purpose of comparing self-reported levels of pain, stiffness and physical function prior to THR or end of study between the exercise therapy group and the control group. The WOMAC has been extensively used and it has been suggested to be a reliable and valid instrument for measuring self-perceived pain and function in patients with hip OA[17, 208]. McConnell et al.[159] stated that the validity, reliability and responsiveness of the pain and physical function subscales are satisfactory, but the measurement properties of the stiffness subscale have not been well demonstrated. Additionally, the WOMAC may lack discriminant validity[185], and its responsiveness have been questioned in patients with

less severe symptoms due to floor effects[171]. Pain and physical function are related concepts, and since the WOMAC pain is filled in immediately prior to the WOMAC physical function and the two subscales have a parallel item content, it is not unlikely that reflections of pain can influence the WOMAC physical function score[185].

In 2005, when our RCT started, few studies had evaluated the measurement properties of the PASE. The findings in the studies by Washburn[233, 234], Schuit et al.[196] and Loland et al.[153] suggested the PASE to have acceptable measurement properties for assessing physical activity level in healthy elderly. Nevertheless, the use of self-reported tools to assess physical activity levels have been questioned, and the properties of the PASE had not been evaluated specifically for patients with hip OA. Hence, we decided to conduct a methodological study to evaluate both the test-retest reliability and the construct validity of the PASE in patients with clinical and radiographic hip OA. To evaluate construct validity we compared the PASE to the Actigraph GT1M accelerometer and the IPAQ, based on that they were considered to measure the same construct as the PASE. Accelerometers have been found to have acceptable criterion validity when compared to measures of free-living total energy expenditure measured by the doubly labelled water method, with correlation coefficient ranging from 0.56-0.83[33, 83, 160, 189]. Previous studies have reported correlation coefficients ranging from 0.29-0.55 between the IPAQ total MET-minutes per week and accelerometer measured counts per minute[56, 94, 138, 155].

The Norwegian version of the PASE[153], which was used in this study, was a slightly adapted version. Due to cultural differences, i.e. the question on walking activities in the original version was removed, and walking activities were instead incorporated in the three questions which address light, moderate and vigorous activity. The rationale for this was, according to Loland et al.[153], that walking activities is the most common form for physical activity in Norway. By removing one question and slightly transforming three others, the possible total achievable score was also influenced. Hence, the results of the Norwegian version of the PASE cannot easily be compared PASE results reported in foreign studies.

Sample size and data analysis

Post power calculations (90 % power and significance level of 5 %) was conducted for the outcome measures included in Paper I. Based on estimated MDCs for aggregated ROM[109], isokinetic knee extension muscle strength[148] and 6MWT[117, 129], the number of patients needed in each group was calculated to be 66, 60, and 42-119, respectively. Based on the MCID for pain on VAS in patients with mild to moderate hip OA[221], the estimated number needed in each group was estimated to be 80. Hence, we seem to lack power to detect possible significant group differences for these specific outcome measures. The certainty of these post power calculations can be questioned, due to lack of studies evaluating the MCID for hip ROM, isokinetic muscle strength and 6MWT in this specific patient group.

In the a priori sample size calculation based on the WOMAC pain subscale[65] it was accounted for a 10 % drop-out rate. This estimate was close to the actual rate of patients who were lost to follow-up or declined participation at subsequent follow-ups. However, a larger proportion than we anticipated underwent THR and did not participate at subsequent follow-ups. Hence, the overall statistical power was reduced for the analyses of the self-reported outcome measures and the outcome measures of impairments and activity limitations.

All analyses of efficacy in Paper I and Paper II were conducted according the ITT principle[108], i.e. all patients were included in the analyses. The linear mixed model was used to compare treatment effect across the two groups over a 29 months follow-up period. It has been suggested to be less vulnerable to missing values than other models[136]. However, due to the relatively large amount of missing data caused by THRs, we suspected that missing was not at random. This was based on the assumption that THR is to be performed in cases of advanced disease; thus we expected the patient in our study who decided to undergo THR to have experienced a negative disease progression. The results of the missing value analysis (Paper I) confirmed that missing was not at random. This may have introduced bias to the between-group analyses of ROM, muscle strength, 6MWT, pain during walking, WOMAC and PASE. Consequently, the long term mean estimates may have been slightly overestimated in both the exercise therapy group and the control group, as the supposedly poorer preoperative results of the patients who went through THR are missing. Furthermore, the between group difference in favour of the exercise group may have been underestimated as more patients in the control group had gone through THR. Because missing was considered to be not at random, we conducted a sensitivity analysis in Paper I by using the LOCF technique to evaluate the robustness of the results and to a greater extent account for the poorer preoperative performance of the patients who eventually underwent THR. The use of LOCF has been criticised, because it has been suggested to introduce bias[176, 237]. Other methods for imputation of missing values have been suggested[237], but no general accepted imputation method exist when 'missing not at random' is the case. Wood et al[236] stated that the choice of model is dependent on an understanding of the context. Thus, in this specific setting, where many patients had missing data due to THR, we considered the LOCF to be an alternative for evaluating preoperative group differences. Another possibility would have been to perform a strict ITT analysis, with all patients attending all follow-ups regardless of their choice of having a THR or not during the follow-up period. However, patients undergoing THR might still have biased the evaluation of the treatment effect, because more patients in the control group underwent THR.

Based on the sample size in previous studies evaluating measurement properties of physical activity questionnaires, and after consulting a statistician, we decided to include 40 patients in the reliability and validity study of the PASE. Sample size calculations in reliability and validity studies are rarely performed, since the objective of these studies is to calculate an estimate rather than to test a specific hypothesis. Terwee et al.[213]

suggested a sample size of at least 50 subjects to be adequate in studies evaluating measurement properties of physical activity questionnaires. Furthermore, inclusion of 30-50 subjects has been considered adequate to enable data for calculating MDC[147]. Several models for calculation of the ICC exists[200]. The two-way model was chosen because it enabled us to include two or more covariates (patients and test occasions). Furthermore, the 'absolute agreement' was used to include the variance of each measure when evaluating the degree of correlation, instead of excluding systematic measurement errors. The SEM, MDC and LoA were calculated to provide information about the precision of the measure expressed in its specific unit of measurement. To evaluate the construct validity of the PASE, we compiled a priory hypotheses as suggested by Mokkink et al.[164], defining the expected direction and magnitude of the correlations.

Results

Paper I and II

In line with our findings, Juhakoski et al.[124] found that exercise therapy had no significant short-term or long-term effect on hip ROM, lower limb muscle strength, or distance covered during the 6MWT[124]. Furthermore, no effect of exercise has been found in other walk tests or stair walk tests[79, 124, 209] In contrast, French et al.[79] demonstrated that patients who were given exercise therapy achieved better aggregated ROM compared to a control group. The post power calculations combined with the large amount of missing observations suggest that the study was underpowered to detect long-term group differences. However, the estimated between group differences in our study were small at all follow-ups. Relatively small deficits were present at baseline for both ROM, muscle strength, and walking capacity when compared to healthy controls and normative data[13, 14, 29, 31, 41, 85, 122, 195, 204, 218]. Thus, the potential for improvement in these outcome measures may have been somewhat limited. Also, the patients in this trial had activity levels superior to what have been reported in other studies in patients with OA, more comparable to healthy individuals. This may indicate that relatively large doses of exercise would be necessary to alter their clinical impairments and functional performance. Furthermore, the lack of effect in isokinetic muscle strength may have been due to inadequate frequency and intensity of the strengthening exercises, below the suggested requirements for muscle strength increase[133, 205] .

Additionally, the relatively small impairments of the patients included in this trial may play a negligible role in the patients' self-perceived functional level. Johnson et al.[119] found that significant improvement in WOMAC pain and physical function was not accompanied by similar improvements in ROM, disease activity and radiographic findings. The discrepancy between self-reported and observer-based measures was suggested to be related to the patients' adaptations to living with a chronic disease[120], but may also suggest that self-reported outcome measures are more responsive to changes over time. It is important to emphasise that neither of the groups in our study demonstrated a decline in mean results over the 29 month follow-up period. In patients with manifest

symptomatic hip OA a mean decline over time would have been expected. Consequently, a comparison with a true control group which received no treatment might have revealed a treatment effect over time for both our exercise therapy group and our control group.

The beneficial pain relief during walking found in the exercise therapy group over the 29 month follow-up period in our study was not accompanied by a significant better WOMAC pain score, in which three of five questions address resting pain. These results suggest that exercise therapy may influence activity related pain to a larger extent than resting pain in patients with hip OA with mild to moderate pain and symptoms. Furthermore, this can be related to the demonstrated effect in WOMAC physical function in the exercise therapy group, as activity related pain is suggested to influence self-reported function[207]. The functional exercises included in the exercise program was based on previously reported ADL difficulties in patients with OA[64]. This task-specificity may have been decisive for the effect in self-reported physical function which was evident in the exercise group after 16 months[65] and 29 months (Paper II). When we consider the beneficial effects for WOMAC physical function and pain during walking in relation to the demonstrated reduced risk for THR, it may suggests that activity related pain and the perceived ability of the patients to perform ADL without difficulty is important when deciding on whether and when THR is to be performed. The previously suggested estimated MCIDs for WOMAC physical function and VAS for pain in patients with hip OA with less severe hip OA, suggest that the demonstrated effect in these outcome measures at the ten, 16 and 29 months follow-up in our study may be clinically meaningful. Juhakoski et al.[124] and French et al.[79] found, in line with our results, a beneficial effect of exercise therapy on WOMAC physical function, but no significant effect on WOMAC pain, while Tak et al.[209] demonstrated less pain and better score in the clinician-rated HHS.

It has been suggested that that the largest treatment effect is present immediately after completion of an exercise intervention, with a decline in treatment effect over time[42, 183]. However, we demonstrated a beneficial effect of exercise therapy on WOMAC physical function and pain on VAS over a 29 months period, and Juhakoski et al.[124] found a significant effect on WOMAC physical function after 18 months in favour of the exercise group. Furthermore, we found that exercise therapy reduced the 6-year risk for THR by 44 %. This was supported by the findings in the study by Pisters et al.[180], who demonstrated a reduced 5-year risk for surgery after individualised exercise treatment in patients presenting with clinical diagnose of hip OA. The beneficial long-term effects in the exercise therapy group may be a result of continuation of the exercise therapy program beyond the 12 weeks exercise period. Long-term adherence to exercise seems to be required to sustain its beneficial effect[182], and higher leisure-time physical activity may also have a protective effect against THR[2]. Unfortunately, we did not obtain data on continuation of exercising in accordance with the exercise therapy program after the end of the 12-week exercise intervention period, but the data on exercise frequency and physical activity level did not differ between the two groups over the 29 month follow-up

period. However, these data should be interpreted with caution, as the main finding in Paper III was that the PASE lack validity to measure level and intensity of physical activity.

The mechanisms for the beneficial effect of exercise in OA are poorly understood. Most research within this field is restricted to patients with knee OA, or a combination of knee and hip OA. Beckwee et al.[16] have identified factors related to body impairments, activity limitations and general fitness, factors within the joint and in the surrounding tissue, and psychosocial factors as potential mechanisms which may explain the exercise induced improvements in pain and function in patients with knee OA. In patients with impaired ROM and muscle strength, improvements in these underlying factors may be beneficial, since ROM and muscle strength have been found to be associated with self-perceived function in hip OA[13, 14, 195]. The studies by French et al,[78] and Hoeksma et al.[107] both found that improvements in hip ROM were accompanied by improvements in self-reported outcome measures. Furthermore, it is suggested that exercise treatment may serve as a disease-modifier by inducing beneficial structural, biochemical and anti-inflammatory effects. These potentially positive effects of exercise on joint health might be relevant for disease deceleration and associated with improvements in pain and function. Still, only a limited number of studies have explored its effect on cartilage health and other intra-articular factors related to OA progression, all concerning the knee joint[102, 163, 188, 192, 222]. Psychosocial factors, including pain coping and self-efficacy, are considered as important for the ability of each individual to adjust to and manage a painful, chronic disease. Hence, positive changes in these factors may contribute to the beneficial effect of exercise therapy in OA. Patients with OA who report high levels of self-efficacy are more likely to have less pain and functional impairments[154]. Furthermore, self-efficacy has been identified as a mediator for the relationship between pain catastrophising and physical disability[199, 202]. Exercise may also influent OA symptoms by enhancing general well-being and reducing depression, and the placebo effect may be considerable in interventions including exercise treatment[16]. Furthermore, positive experiences with patient information and exercise may affect the patients' preferred choice of treatment.

Total hip replacement surgery is undoubtedly a good treatment option in cases of advanced hip OA. However, as the number of primary THRs have increased substantially the last decades[137], concern has been raised because they will result in a greater number of patients experiencing surgical complications and, subsequently, an increase in future revisions[84]. According to the Norwegian Arthroplasty Register the relative number of revisions increase even more than the number of primary THRs[167, 168]. Of a random selection of 700 orthopaedic complaints reported to the Norwegian System of Compensation to Patients between 1993 and 1999, most complaints were referring to primary THRs due to hip OA[22]. Mechanisms delimiting the future THR and revision burden are highly requested[84, 137], and include both consensus regarding criteria for THR and interventions potentially slowing down disease progression or improving the ability of the patients to cope with the disease. According to our findings, exercise therapy may represent a safe and easily administered treatment modality, with few complications,

which potentially can reduce the burden of THR. However, it has been debated whether postponing surgery is beneficial for the patients in the long-term[71, 226], as preoperative status has been associated with postoperative outcome. Interestingly, some studies have found better postoperative results in patients with less pain and functional limitations prior to surgery[172, 230], whereas other find patients with poorer preoperative outcome to experience larger post-operative improvements[100, 123, 230]. Additionally, Wang et al.[229] and Gilbey et al.[86] found that preoperative exercise improved post-operative physical function. These findings suggest that postponing surgery is safe if the preoperative condition remains stable or improves. It seems as if a treatment window exists in less severe disease where exercise therapy can be beneficial. Thus, in patients who are presenting acceptable pain and functional limitations, and who are able to participate in desired activities and social events, we consider postponing surgery to be appropriate, as it may reduce the future need for THR or additional THR revision surgery, and its accompanying potential for complications.

Paper III

In paper III we found the PASE to have acceptable test-retest reliability with an ICC of 0.77 for the total score, based on that an ICC above 0.70 has been considered to represent acceptable reliability[70, 213]. However, the estimated measurement error was relatively large, with a MDC of 87, meaning that a change in total PASE score would have to exceed 87 to represent a true change larger than the measurement error. Studies evaluating the test-retest reliability of the PASE in the general elderly population have reported ICC values ranging from 0.65 to 0.81[92, 170, 225, 234].

The construct validity of the PASE was considered to poor, as the correlation coefficient between the PASE and the Actigraph GT1M did not exceed the cut-off for acceptable validity. Perfect correlation between self-reported physical activity level and objectively measures of activity is unlikely[213], and Lee et al.[144] concluded that a correlation of 0.3–0.4 is perhaps as close as can be expected when comparing these instruments. Washburn[233] found a relatively high correlation of 0.49 between the PASE score and an accelerometer. However, other studies have reported a correlation between the PASE and accelerometry of 0.16[93, 150], more in line with our findings. The validity of the PASE has also been evaluated by comparing it to total energy expenditure calculated by the doubly labelled water method, giving correlation coefficients of 0.20 and 0.68[33, 196]. However, none of these previous studies have evaluated the validity of the PASE in patients with OA. Furthermore, it has not been evaluated if the PASE can provide reliable and valid data for physical activity intensity. However, our results indicated that the different intensity items of the PASE had poor construct validity and test-retest reliability.

A possible reason for the discrepancy we and others have found between the PASE and the Actigraph GT1M may be that questionnaires seem to over-report physical activity[56, 94], while accelerometers on the other hand have been found to under-report activity compared to the doubly labelled water method[145]. Problems with over-reporting when

using self-administered questionnaires can be related to recall bias and difficulties in evaluating true intensity level. Dyrstad et al.[56] found that over-reporting physical activity was more common in men and in less educated subjects. Problems with under-reporting physical activity by accelerometry may be related to its limitations. Accelerometers are unable to capture activities like cycling and upper limb exercise, cannot be used to measure swimming and other water activities, and can only provide measurements for the particular time it is worn. Four days of registration[210] with a minimum wear time of 13 hours per day[104] seems to be required to provide valid measures of physical activity. Hence, we consider our results to be valid, with a mean of 15 hours of registration per day over a mean of seven days.

In the RCT the PASE was included to assess overall activity level at repeated follow-ups. Unfortunately, despite having acceptable reliability, the PASE fails to validly measure amount and intensity of physical activity in patients with hip OA with mild to moderate symptoms. Hence, the estimated PASE results and between group differences reported in Paper I and II must be interpreted with caution. Kayes and McPherson[126] points out that both self-reported and objective measures of physical activity have limitations, but because they are likely to measure slightly different aspects of physical activity combined use of both methods can be useful. According to Ainsworth et al.[3] questionnaires are suitable to measure physical activity in most settings, and have further suggested the use of a conceptual framework to improve accuracy of estimates derived from physical activity questionnaires[4]. In contrast, Trost et al.[220] emphasises the advantages of the objective measures of physical activities and its appropriateness for use in clinical settings.

CONCLUSIONS

In this thesis the long-term effect of exercise therapy and patient education in patients with hip OA presenting with mild to moderate symptoms has been investigated. Additionally, the measurement properties of a questionnaire, PASE, assessing physical activity level have been evaluated. Since few RCTs has been conducted in this field, the study is considered to add new and important knowledge with respect to long-term results after non-surgical, active treatment modalities evaluated by different outcome measures in hip OA. The following conclusions can be drawn:

I. Patients presenting with mild to moderate hip OA who were given a 12 week exercise program and a patient education program reported significantly less pain during walking over the 29 months follow-up period compared to patients who were given patient education only. However, exercise therapy in addition to patient education provided no benefit over patient education only evaluated by hip ROM, muscle strength or walking capacity.

II. The 6-year survival of the native hip to THR was significantly higher in patients with mild to moderate hip OA who participated in a 12 week exercise therapy program and patient educations compared to patients who participated in patient education only, suggesting that exercise may reduce or postpone the need for THR.

III. Patients with mild to moderate hip OA who were given both exercise therapy and patient education had significantly better self-reported physical function, evaluated by WOMAC physical function, compared to patients who were given patient education only. No group differences were found for WOMAC pain or WOMAC stiffness.

IV. The PASE was found to have moderate test-retest reliability with an acceptable ICC-value, but with large measurement error. The PASE was not considered to provide valid measures for overall physical activity level and specific physical activity intensities in patients with mild to moderate hip OA, based on its poor correlation with an accelerometer.

FUTURE PERSPECTIVES

- More high quality RCTs evaluating the effect of exercise treatment are needed to increase the level of evidence for these modalities in patients with hip OA

- Future studies should be aiming to identify specific exercise components which are particularly effective in hip OA

- Studies evaluating the effect of exercise treatment should also
    - explore potential mechanisms for its beneficial effect, including mechanical factors, systemic factors, factors related to body function and impairments, and psychological factors
    - study treatment effect in specific sub-groups of patients with hip OA
    - identify patients who are most likely to benefit from exercise interventions

- There is need for future studies to confirm that exercise therapy may postpone the need for THR

- Effort should be made to reach general consensus on criteria for THR

- The development of easy, valid and reliable methods for assessing physical activity and adherence to exercise interventions is highly encouraged. The use of accelerometry should be considered in clinical trials where valid data on physical activity is required.

REFERENCES

1. Abadie E, Ethgen D, Avouac B et al. Recommendations for the use of new methods to assess the efficacy of disease-modifying drugs in the treatment of osteoarthritis. Osteoarthritis Cartilage 2004;12(4):263-268.

2. Ageberg E, Engstrom G, Gerhardsson d, V, Rollof J, Roos EM, Lohmander LS. Effect of leisure time physical activity on severe knee or hip osteoarthritis leading to total joint replacement: a population-based prospective cohort study. BMC Musculoskelet Disord 2012;13:73.

3. Ainsworth BE. How do I measure physical activity in my patients? Questionnaires and objective methods. Br J Sports Med 2009;43(1):6-9.

4. Ainsworth BE, Caspersen CJ, Matthews CE, Masse LC, Baranowski T, Zhu W. Recommendations to improve the accuracy of estimates of physical activity derived from self report. J Phys Act Health 2012;9 Suppl 1:S76-S84.

5. Altman R, Alarcon G, Appelrouth D et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. Arthritis Rheum 1991;34(5):505-514.

6. Altman R, Brandt K, Hochberg M et al. Design and conduct of clinical trials in patients with osteoarthritis: recommendations from a task force of the Osteoarthritis Research Society. Results from a workshop. Osteoarthritis Cartilage 1996;4(4):217-243.

7. Altman RD, Abadie E, Avouac B et al. Total joint replacement of hip or knee as an outcome measure for structure modifying trials in osteoarthritis. Osteoarthritis Cartilage 2005;13(1):13-19.

8. Altman RD, Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. Osteoarthritis Cartilage 2007;15 Suppl A:A1-56.

9. Altman RD, Hochberg M, Murphy WA, Jr., Wolfe F, Lequesne M. Atlas of individual radiographic features in osteoarthritis. Osteoarthritis Cartilage 1995;3 Suppl A:3-70.

10. Angst F, Aeschlimann A, Michel BA, Stucki G. Minimal clinically important rehabilitation effects in patients with osteoarthritis of the lower extremities. J Rheumatol 2002;29(1):131-138.

11. Angst F, Aeschlimann A, Stucki G. Smallest detectable and minimal clinically important differences of rehabilitation intervention with their implications for required sample sizes using WOMAC and SF-36 quality of life measurement instruments in patients with osteoarthritis of the lower extremities. Arthritis Rheum 2001;45(4):384-391.

12. Apold H, Meyer HE, Espehaug B, Nordsletten L, Havelin LI, Flugsrud GB. Weight gain and the risk of total hip replacement a population-based prospective cohort study of 265,725 individuals. Osteoarthritis Cartilage 2011;19(7):809-815.

13. Arokoski MH, Arokoski JP, Haara M et al. Hip muscle strength and muscle cross sectional area in men with and without hip osteoarthritis. J Rheumatol 2002;29(10):2185-2195.

14. Arokoski MH, Haara M, Helminen HJ, Arokoski JP. Physical function in men with and without hip osteoarthritis. Arch Phys Med Rehabil 2004;85(4):574-581.

15. Bagshaw SM, Bellomo R. The need to reform our assessment of evidence from clinical trials: a commentary. Philos Ethics Humanit Med 2008;3:23.

16. Beckwee D, Vaes P, Cnudde M, Swinnen E, Bautmans I. Osteoarthritis of the knee: why does exercise work? A qualitative study of the literature. Ageing Res Rev 2013;12(1):226-236.

17. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15(12):1833-1840.

18. Bellamy N, Kirwan J, Boers M et al. Recommendations for a core set of outcome measures for future phase III clinical trials in knee, hip, and hand osteoarthritis. Consensus development at OMERACT III. J Rheumatol 1997;24(4):799-802.

19. Benito MJ, Veale DJ, FitzGerald O, van den Berg WB, Bresnihan B. Synovial tissue inflammation in early and late osteoarthritis. Ann Rheum Dis 2005;64(9):1263-1267.

20. Berenbaum F. Osteoarthritis as an inflammatory disease (osteoarthritis is not osteoarthrosis!). Osteoarthritis Cartilage 2013;21(1):16-21.

21. Birrell F, Lunt M, Macfarlane G, Silman A. Association between pain in the hip region and radiographic changes of osteoarthritis: results from a population-based study. Rheumatology (Oxford) 2005;44(3):337-341.

22. Bjerkreim I, Steen H. [Complaints in connection with total hip replacement in hip osteoarthritis]. Tidsskr Nor Laegeforen 2003;123(4):468-469.

23. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res 1999;8(2):135-160.

24. Bohensky M, Ackerman I, de SR, Gorelik A, Brand C, Adjunct. Lifetime risk of total hip replacement surgery and temporal trends in utilisation: A population-based analysis. Arthritis Care Res (Hoboken) 2014. [Epub ahead of print]

25. Borg GA. Psychophysical bases of perceived exertion. Med Sci Sports Exerc 1982;14(5):377-381.

26. Bruyere O, Pavelka K, Rovati LC et al. Total joint replacement after glucosamine sulphate treatment in knee osteoarthritis: results of a mean 8-year observation of patients from two previous 3-year, randomised, placebo-controlled trials. Osteoarthritis Cartilage 2008;16(2):254-260.

27. Buckwalter JA, Martin JA. Osteoarthritis. Adv Drug Deliv Rev 2006;58(2):150-167.

28. Buckwalter JA, Stanish WD, Rosier RN, Schenck RC, Jr., Dennis DA, Coutts RD. The increasing need for nonoperative treatment of patients with osteoarthritis. Clin Orthop Relat Res 2001;(385):36-45.

29. Camarri B, Eastwood PR, Cecins NM, Thompson PJ, Jenkins S. Six minute walk distance in healthy subjects aged 55-75 years. Respir Med 2006;100(4):658-665.

30. Carpenter MR, Carpenter RL, Peel J et al. The reliability of isokinetic and isometric leg strength measures among individuals with symptoms of mild osteoarthritis. J Sports Med Phys Fitness 2006;46(4):585-589.

31. Casanova C, Celli BR, Barria P et al. The 6-min walk distance in healthy subjects: reference standards from seven countries. Eur Respir J 2011;37(1):150-156.

32. Chu Miow LD, Reichmann WM, Gossec L, Losina E, Conaghan PG, Maillefert JF. Validity and responsiveness of radiographic joint space width metric measurement in hip osteoarthritis: a systematic review. Osteoarthritis Cartilage 2011;19(5):543-549.

33. Colbert LH, Matthews CE, Havighurst TC, Kim K, Schoeller DA. Comparative validity of physical activity measures in older adults. Med Sci Sports Exerc 2011;43(5):867-876.

34. Conaghan PG, Dickson J, Grant RL. Care and management of osteoarthritis in adults: summary of NICE guidance. BMJ 2008;336(7642):502-503.

35. Craig CL, Marshall AL, Sjostrom M et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003;35(8):1381-1395.

36. Croft P, Cooper C, Wickham C, Coggon D. Defining osteoarthritis of the hip for epidemiologic studies. Am J Epidemiol 1990;132(3):514-522.

37. Culliford DJ, Maskell J, Kiran A et al. The lifetime risk of total hip and knee arthroplasty: results from the UK general practice research database. Osteoarthritis Cartilage 2012;20(6):519-524.

38. Dagenais S, Garbedian S, Wai EK. Systematic review of the prevalence of radiographic primary hip osteoarthritis. Clin Orthop Relat Res 2009;467(3):623-637.

39. Danielsson L, Lindberg H. Prevalence of coxarthrosis in an urban population during four decades. Clin Orthop Relat Res 1997;(342):106-110.

40. Danielsson LG. INCIDENCE AND PROGNOSIS OF COXARTHROSIS. Acta Orthop Scand Suppl 1964;66:SUPPL-114.

41. Danneskiold-Samsoe B, Bartels EM, Bulow PM et al. Isokinetic and isometric muscle strength in a healthy population with special reference to age and gender. Acta Physiol (Oxf) 2009;197 Suppl 673:1-68.

42. Davis AM, MacKay C. Osteoarthritis year in review: outcome of rehabilitation. Osteoarthritis Cartilage 2013;21(10):1414-1424.

43. de Groot IB, Bussmann HJ, Stam HJ, Verhaar JA. Small increase of actual physical activity 6 months after total hip or knee arthroplasty. Clin Orthop Relat Res 2008;466(9):2201-2208.

44. de Jong OR, Hopman-Rock M, Tak EC, Klazinga NS. An implementation study of two evidence-based exercise and health education programmes for older adults with osteoarthritis of the knee and hip. Health Educ Res 2004;19(3):316-325.

45. Dennison E, Cooper C. The natural history and prognosis of osteoarhtritis. In: Brandt K, Doherty M, Lohmander LS, editors. Osteoarthritis. 2 ed. Oxford University Press; 2014.

46. Dieppe P. Developments in osteoarthritis. Rheumatology (Oxford) 2011;50(2):245-247.

47. Dieppe P, Cushnaghan J, Tucker M, Browning S, Shepstone L. The Bristol 'OA500 study': progression and impact of the disease after 8 years. Osteoarthritis Cartilage 2000;8(2):63-68.

48. Dieppe P, Judge A, Williams S et al. Variations in the pre-operative status of patients coming to primary hip replacement for osteoarthritis in European orthopaedic centres. BMC Musculoskelet Disord 2009;10:19.

49. Dieppe PA, Lohmander LS. Pathogenesis and management of pain in osteoarthritis. Lancet 2005;365(9463):965-973.

50. Dobson F, Hinman RS, Roos EM et al. OARSI recommended performance-based tests to assess physical function in people diagnosed with hip or knee osteoarthritis. Osteoarthritis Cartilage 2013;21(8):1042-1052.

51. Donner A, Eliasziw M. Sample size requirements for reliability studies. Stat Med 1987;6(4):441-448.

52. Dougados M, Gueguen A, Nguyen M et al. Requirement for total hip arthroplasty: an outcome measure of hip osteoarthritis? J Rheumatol 1999;26(4):855-861.

53. Dougados M, Nguyen M, Berdah L, Mazieres B, Vignon E, Lequesne M. Evaluation of the structure-modifying effects of diacerein in hip osteoarthritis: ECHODIAH, a three-year, placebo-controlled trial. Evaluation of the Chondromodulating Effect of Diacerein in OA of the Hip. Arthritis Rheum 2001;44(11):2539-2547.

54. Dreinhofer KE, Dieppe P, Sturmer T et al. Indications for total hip replacement: comparison of assessments of orthopaedic surgeons and referring physicians. Ann Rheum Dis 2006;65(10):1346-1350.

55. Dunlop DD, Song J, Semanik PA, Sharma L, Chang RW. Physical activity levels and functional performance in the osteoarthritis initiative: a graded relationship. Arthritis Rheum 2011;63(1):127-136.

56. Dyrstad SM, Hansen BH, Holme IM, Anderssen SA. Comparison of self-reported versus accelerometer-measured physical activity. Med Sci Sports Exerc 2014;46(1):99-106.

57. Eitzen I, Kallerud H, Fernandes L, Nordsletten L, Knarr B, Risberg MA. Gait characteristics, symptoms and function in subjects with hip osteoarhtirits: A prospective study with 6-7 years follow-up. 2014. [Submitted work]

58. Eland DC, Singleton TN, Conaster RR et al. The "iliacus test": new information for the evaluation of hip extension dysfunction. J Am Osteopath Assoc 2002;102(3):130-142.

59. Enright PL. The six-minute walk test. Respir Care 2003;48(8):783-785.

60. Ethgen O, Bruyere O, Richy F, Dardennes C, Reginster JY. Health-related quality of life in total hip and total knee arthroplasty. A qualitative and systematic review of the literature. J Bone Joint Surg Am 2004;86-A(5):963-974.

61. Farr JN, Going SB, Lohman TG et al. Physical activity levels in patients with early knee osteoarthritis measured by accelerometry. Arthritis Rheum 2008;59(9):1229-1236.

62. Felson DT. An update on the pathogenesis and epidemiology of osteoarthritis. Radiol Clin North Am 2004;42(1):1-9.

63. Fernandes L, Hagen KB, Bijlsma JW et al. EULAR recommendations for the non-pharmacological core management of hip and knee osteoarthritis. Ann Rheum Dis 2013;72(7):1125-1135.

64. Fernandes L, Storheim K, Nordsletten L, Risberg MA. Development of a therapeutic exercise program for patients with osteoarthritis of the hip. Phys Ther 2010;90(4):592-601.

65. Fernandes L, Storheim K, Sandvik L, Nordsletten L, Risberg MA. Efficacy of patient education and supervised exercise vs patient education alone in patients with hip osteoarthritis: a single blind randomized clinical trial. Osteoarthritis Cartilage 2010;18(10):1237-1243.

66. Fitzpatrick R, Shortall E, Sculpher M et al. Primary total hip replacement surgery: a systematic review of outcomes and modelling of cost-effectiveness associated with different prostheses. Health Technol Assess 1998;2(20):1-64.

67. Flugsrud GB, Nordsletten L, Espehaug B, Havelin LI, Engeland A, Meyer HE. The impact of body mass index on later total hip arthroplasty for primary osteoarthritis: a cohort study in 1.2 million persons. Arthritis Rheum 2006;54(3):802-807.

68. Flugsrud GB, Nordsletten L, Espehaug B, Havelin LI, Meyer HE. Risk factors for total hip replacement due to primary osteoarthritis: a cohort study in 50,034 persons. Arthritis Rheum 2002;46(3):675-682.

69. Fontaine KR, Bartlett SJ, Heo M. Are health care professionals advising adults with arthritis to become more physically active? Arthritis Rheum 2005;53(2):279-283.

70. Forsen L, Loland NW, Vuillemin A et al. Self-administered physical activity questionnaires for the elderly: a systematic review of measurement properties. Sports Med 2010;40(7):601-623.

71. Fortin PR, Penrod JR, Clarke AE et al. Timing of total joint replacement affects clinical outcomes among patients with osteoarthritis of the hip or knee. Arthritis Rheum 2002;46(12):3327-3330.

72. Frankel L, Sanmartin C, Conner-Spady B et al. Osteoarthritis patients' perceptions of "appropriateness" for total joint replacement surgery. Osteoarthritis Cartilage 2012;20(9):967-973.

73. Frankel S, Eachus J, Pearson N et al. Population requirement for primary hip-replacement surgery: a cross-sectional study. Lancet 1999;353(9161):1304-1309.

74. Franklin J, Ingvarsson T, Englund M, Ingimarsson O, Robertsson O, Lohmander LS. Natural history of radiographic hip osteoarthritis: A retrospective cohort study with 11-28 years of followup. Arthritis Care Res (Hoboken) 2011;63(5):689-695.

75. Fransen M, McConnell S. Exercise for osteoarthritis of the knee. Cochrane Database Syst Rev 2008;(4):CD004376.

76. Fransen M, McConnell S, Hernandez-Molina G, Reichenbach S. Exercise for osteoarthritis of the hip. Cochrane Database Syst Rev 2009;(3):CD007912.

77. Freedson P, Pober D, Janz KF. Calibration of accelerometer output for children. Med Sci Sports Exerc 2005;37(11 Suppl):S523-S530.

78. French HP, Cusack T, Brennan A et al. Exercise and manual physiotherapy arthritis research trial (EMPART) for osteoarthritis of the hip: a multicenter randomized controlled trial. Arch Phys Med Rehabil 2013;94(2):302-314.

79. French HP, Cusack T, Brennan A et al. Exercise and manual physiotherapy arthritis research trial (EMPART) for osteoarthritis of the hip: a multicenter randomized controlled trial. Arch Phys Med Rehabil 2013;94(2):302-314.

80. French HP, Fitzpatrick M, FitzGerald O. Responsiveness of physical function outcomes following physiotherapy intervention for osteoarthritis of the knee: an outcome comparison study. Physiotherapy 2011;97(4):302-308.

81. Gabriel SE, Crowson CS, Campion ME, O'Fallon WM. Direct medical costs unique to people with arthritis. J Rheumatol 1997;24(4):719-725.

82. Gabriel SE, Crowson CS, Campion ME, O'Fallon WM. Indirect and nonmedical costs among people with rheumatoid arthritis and osteoarthritis compared with nonarthritic controls. J Rheumatol 1997;24(1):43-48.

83. Gardner AW, Poehlman ET. Assessment of free-living daily physical activity in older claudicants: validation against the doubly labeled water technique. J Gerontol A Biol Sci Med Sci 1998;53(4):M275-M280.

84. Ghomrawi HM, Schackman BR, Mushlin AI. Appropriateness criteria and elective procedures--total joint arthroplasty. N Engl J Med 2012;367(26):2467-2469.

85. Gibbons WJ, Fruchter N, Sloan S, Levy RD. Reference values for a multiple repetition 6-minute walk test in healthy adults older than 20 years. J Cardiopulm Rehabil 2001;21(2):87-93.

86. Gilbey HJ, Ackland TR, Wang AW, Morton AR, Trouchet T, Tapper J. Exercise improves early functional recovery after total hip arthroplasty. Clin Orthop Relat Res 2003;(408):193-200.

87. Gossec L, Paternotte S, Bingham CO, III et al. OARSI/OMERACT initiative to define states of severity and indication for joint replacement in hip and knee osteoarthritis. An OMERACT 10 Special Interest Group. J Rheumatol 2011;38(8):1765-1769.

88. Gossec L, Paternotte S, Maillefert JF et al. The role of pain and functional impairment in the decision to recommend total joint replacement in hip and knee osteoarthritis: an international cross-sectional study of 1909 patients. Report of the OARSI-OMERACT Task Force on total joint replacement. Osteoarthritis Cartilage 2011;19(2):147-154.

89. Gray-Donald K, Kramer MS. Causality inference in observational vs. experimental studies. An empirical comparison. Am J Epidemiol 1988;127(5):885-892.

90. Grotle M, Hagen KB, Natvig B, Dahl FA, Kvien TK. Prevalence and burden of osteoarthritis: results from a population survey in Norway. J Rheumatol 2008;35(4):677-684.

91. Guillemin F, Rat AC, Mazieres B et al. Prevalence of symptomatic hip and knee osteoarthritis: a two-phase population-based survey. Osteoarthritis Cartilage 2011;19(11):1314-1322.

92. Hagiwara A, Ito N, Sawai K, Kazuma K. Validity and reliability of the Physical Activity Scale for the Elderly (PASE) in Japanese elderly people. Geriatr Gerontol Int 2008;8(3):143-151.

93. Hagiwara A, Ito N, Sawai K, Kazuma K. Validity and reliability of the Physical Activity Scale for the Elderly (PASE) in Japanese elderly people. Geriatr Gerontol Int 2008;8(3):143-151.

94. Hagstromer M, Oja P, Sjostrom M. The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity. Public Health Nutr 2006;9(6):755-762.

95. Hansen BH, Kolle E, Dyrstad SM, Holme I, Anderssen SA. Accelerometer-determined physical activity in adults and older people. Med Sci Sports Exerc 2012;44(2):266-272.

96. Harding P, Holland AE, Delany C, Hinman RS. Do Activity Levels Increase After Total Hip and Knee Arthroplasty? Clin Orthop Relat Res 2013. [Epub ahead of print]

97. Harris WH. Traumatic arthritis of the hip after dislocation and acetabular fractures: treatment by mold arthroplasty. An end-result study using a new method of result evaluation. J Bone Joint Surg Am 1969;51(4):737-755.

98. Hartofilakidis G, Karachalios T. Idiopathic osteoarthritis of the hip: incidence, classification, and natural history of 272 cases. Orthopedics 2003;26(2):161-166.

99. Haskell WL, Lee IM, Pate RR et al. Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. Med Sci Sports Exerc 2007;39(8):1423-1434.

100. Hawker GA, Badley EM, Borkhoff CM et al. Which patients are most likely to benefit from total joint arthroplasty? Arthritis Rheum 2013;65(5):1243-1252.

101. Hawker GA, Guan J, Croxford R et al. A prospective population-based study of the predictors of undergoing total joint arthroplasty. Arthritis Rheum 2006;54(10):3212-3220.

102. Helmark IC, Petersen MC, Christensen HE, Kjaer M, Langberg H. Moderate loading of the human osteoarthritic knee joint leads to lowering of intraarticular cartilage oligomeric matrix protein. Rheumatol Int 2012;32(4):1009-1014.

103. Hernandez-Molina G, Reichenbach S, Zhang B, Lavalley M, Felson DT. Effect of therapeutic exercise for hip osteoarthritis pain: results of a meta-analysis. Arthritis Rheum 2008;59(9):1221-1228.

104. Herrmann SD, Barreira TV, Kang M, Ainsworth BE. How many hours are enough? Accelerometer wear time may provide bias in daily activity estimates. J Phys Act Health 2013;10(5):742-749.

105. Hirata S, Ono R, Yamada M et al. Ambulatory physical activity, disease severity, and employment status in adult women with osteoarthritis of the hip. J Rheumatol 2006;33(5):939-945.

106. Hochberg MC, Altman RD, April KT et al. American College of Rheumatology 2012 recommendations for the use of nonpharmacologic and pharmacologic therapies in osteoarthritis of the hand, hip, and knee. Arthritis Care Res (Hoboken) 2012;64(4):465-474.

107. Hoeksma HL, Dekker J, Ronday HK et al. Comparison of manual therapy and exercise therapy in osteoarthritis of the hip: a randomized clinical trial. Arthritis Rheum 2004;51(5):722-729.

108. Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. BMJ 1999;319(7211):670-674.

109. Holm I, Bolstad B, Lutken T, Ervik A, Rokkum M, Steen H. Reliability of goniometric measurements and visual estimates of hip ROM in patients with osteoarthrosis. Physiother Res Int 2000;5(4):241-248.

110. Hunter DJ. Lower extremity osteoarthritis management needs a paradigm shift. Br J Sports Med 2011;45(4):283-288.

111. Hunter DJ, McDougall JJ, Keefe FJ. The symptoms of osteoarthritis and the genesis of pain. Rheum Dis Clin North Am 2008;34(3):623-643.

112. Hunter DJ, Riordan EA. The impact of arthritis on pain and quality of life: an Australian survey. Int J Rheum Dis 2013. [Epub ahead of print]

113. Hurley MV, Walsh N, Bhavnani V, Britten N, Stevenson F. Health beliefs before and after participation on an exercised-based rehabilitation programme for chronic knee pain: doing is believing. BMC Musculoskelet Disord 2010;11:31.

114. Ingvarsson T, Hagglund G, Lindberg H, Lohmander LS. Assessment of primary hip osteoarthritis: comparison of radiographic methods using colon radiographs. Ann Rheum Dis 2000;59(8):650-653.

115. Ingvarsson T, Hagglund G, Lohmander LS. Prevalence of hip osteoarthritis in Iceland. Ann Rheum Dis 1999;58(4):201-207.

116. Jacobsen S, Sonne-Holm S, Soballe K, Gebuhr P, Lund B. Radiographic case definitions and prevalence of osteoarthrosis of the hip: a survey of 4 151 subjects in the Osteoarthritis Substudy of the Copenhagen City Heart Study. Acta Orthop Scand 2004;75(6):713-720.

117. Jakobsen TL, Kehlet H, Bandholm T. Reliability of the 6-min walk test after total knee arthroplasty. Knee Surg Sports Traumatol Arthrosc 2013;21(11):2625-2628.

118. Jessep SA, Walsh NE, Ratcliffe J, Hurley MV. Long-term clinical benefits and costs of an integrated rehabilitation programme compared with outpatient physiotherapy for chronic knee pain. Physiotherapy 2009;95(2):94-102.

119. Johnson SR, Archibald A, Davis AM, Badley E, Wright JG, Hawker GA. Is self-reported improvement in osteoarthritis pain and disability reflected in objective measures? J Rheumatol 2007;34(1):159-164.

120. Johnson SR, Archibald A, Davis AM, Badley E, Wright JG, Hawker GA. Is self-reported improvement in osteoarthritis pain and disability reflected in objective measures? J Rheumatol 2007;34(1):159-164.

121. Jordan JM, Helmick CG, Renner JB et al. Prevalence of hip symptoms and radiographic and symptomatic hip osteoarthritis in African Americans and Caucasians: the Johnston County Osteoarthritis Project. J Rheumatol 2009;36(4):809-815.

122. Judd DL, Thomas AC, Dayton MR, Stevens-Lapsley JE. Strength and functional deficits in individuals with hip osteoarthritis compared to healthy, older adults. Disabil Rehabil 2013. [Epub ahead of print]

123. Judge A, Arden NK, Batra RN et al. The association of patient characteristics and surgical variables on symptoms of pain and function over 5 years following primary hip-replacement surgery: a prospective cohort study. BMJ Open 2013;3(3).

124. Juhakoski R, Tenhonen S, Malmivaara A, Kiviniemi V, Anttonen T, Arokoski JP. A pragmatic randomized controlled study of the effectiveness and cost consequences of exercise therapy in hip osteoarthritis. Clin Rehabil 2010;25(4):370-383.

125. Karlson EW, Mandl LA, Aweh GN, Sangha O, Liang MH, Grodstein F. Total hip replacement due to osteoarthritis: the importance of age, obesity, and other modifiable risk factors. Am J Med 2003;114(2):93-98.

126. Kayes NM, McPherson KM. Measuring what matters: does 'objectivity' mean good science? Disabil Rehabil 2010;32(12):1011-1019.

127. Kean CO, Birmingham TB, Garland SJ, Bryant DM, Giffin JR. Minimal detectable change in quadriceps strength and voluntary muscle activation in patients with knee osteoarthritis. Arch Phys Med Rehabil 2010;91(9):1447-1451.

128. KELLGREN JH, LAWRENCE JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957;16(4):494-502.

129. Kennedy DM, Stratford PW, Wessel J, Gollish JD, Penney D. Assessing stability and change of four performance measures: a longitudinal study evaluating outcome following total hip and knee arthroplasty. BMC Musculoskelet Disord 2005;6:3.

130. Kingsbury SR, Gross HJ, Isherwood G, Conaghan PG. Osteoarthritis in Europe: impact on health status, work productivity and use of pharmacotherapies in five European countries. Rheumatology (Oxford) 2014. [Epub ahead of print]

131. Klassbo M, Harms-Ringdahl K, Larsson G. Examination of passive ROM and capsular patterns in the hip. Physiother Res Int 2003;8(1):1-12.

132. Klassbo M, Larsson G, Harms-Ringdahl K. Promising outcome of a hip school for patients with hip dysfunction. Arthritis Rheum 2003;49(3):321-327.

133. Kraemer WJ, Adams K, Cafarelli E et al. American College of Sports Medicine position stand. Progression models in resistance training for healthy adults. Med Sci Sports Exerc 2002;34(2):364-380.

134. Kraemer WJ, Adams K, Cafarelli E et al. American College of Sports Medicine position stand. Progression models in resistance training for healthy adults. Med Sci Sports Exerc 2002;34(2):364-380.

135. Kroon FP, van der Burg LR, Buchbinder R, Osborne RH, Johnston RV, Pitt V. Self-management education programmes for osteoarthritis. Cochrane Database Syst Rev 2014;1:CD008963.

136. Krueger C, Tian L. A comparison of the general linear mixed model and repeated measures ANOVA using a dataset with multiple missing data points. Biol Res Nurs 2004;6(2):151-157.

137. Kurtz S, Mowat F, Ong K, Chan N, Lau E, Halpern M. Prevalence of primary and revision total hip and knee arthroplasty in the United States from 1990 through 2002. J Bone Joint Surg Am 2005;87(7):1487-1497.

138. Kurtze N, Rangul V, Hustvedt BE. Reliability and validity of the international physical activity questionnaire in the Nord-Trondelag health study (HUNT) population of men. BMC Med Res Methodol 2008;8:63.

139. Lane NE, Nevitt MC, Hochberg MC, Hung YY, Palermo L. Progression of radiographic hip osteoarthritis over eight years in a community sample of elderly white women. Arthritis Rheum 2004;50(5):1477-1486.

140. Lanyon P, Muir K, Doherty S, Doherty M. Assessment of a genetic contribution to osteoarthritis of the hip: sibling study. BMJ 2000;321(7270):1179-1183.

141. Larmer PJ, Reay ND, Aubert ER, Kersten P. Systematic review of guidelines for the physical management of osteoarthritis. Arch Phys Med Rehabil 2014;95(2):375-389.

142. Lawrence RC, Felson DT, Helmick CG et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part II. Arthritis Rheum 2008;58(1):26-35.

143. Ledingham J, Dawson S, Preston B, Milligan G, Doherty M. Radiographic progression of hospital referred osteoarthritis of the hip. Ann Rheum Dis 1993;52(4):263-267.

144. Lee PH, Yu YY, McDowell I, Leung GM, Lam TH, Stewart SM. Performance of the international physical activity questionnaire (short form) in subgroups of the Hong Kong chinese population. Int J Behav Nutr Phys Act 2011;8:81.

145. Leenders NY, Sherman WM, Nagaraja HN. Energy expenditure estimated by accelerometry and doubly labeled water: do they agree? Med Sci Sports Exerc 2006;38(12):2165-2172.

146. Lequesne MG, Mery C, Samson M, Gerard P. Indexes of severity for osteoarthritis of the hip and knee. Validation--value in comparison with other assessment tests. Scand J Rheumatol Suppl 1987;65:85-89.

147. Lexell JE, Downham DY. How to assess the reliability of measurements in rehabilitation. Am J Phys Med Rehabil 2005;84(9):719-723.

148. Lienhard K, Lauermann SP, Schneider D, Item-Glatthorn JF, Casartelli NC, Maffiuletti NA. Validity and reliability of isometric, isokinetic and isoinertial modalities for the assessment of quadriceps muscle strength in patients with total knee arthroplasty. J Electromyogr Kinesiol 2013;23(6):1283-1288.

149. Lievense AM, Koes BW, Verhaar JA, Bohnen AM, Bierma-Zeinstra SM. Prognosis of hip pain in general practice: a prospective followup study. Arthritis Rheum 2007;57(8):1368-1374.

150. Liu RD, Buffart LM, Kersten MJ et al. Psychometric properties of two physical activity questionnaires, the AQuAA and the PASE, in cancer patients. BMC Med Res Methodol 2011;11:30.

151. Lofvendahl S, Bizjajeva S, Ranstam J, Lidgren L. Indications for hip and knee replacement in Sweden. J Eval Clin Pract 2011;17(2):251-260.

152. Lohmander LS, Engesaeter LB, Herberts P, Ingvarsson T, Lucht U, Puolakka TJ. Standardized incidence rates of total hip replacement for primary hip osteoarthritis in the 5 Nordic countries: similarities and differences. Acta Orthop 2006;77(5):733-740.

153. Loland NW. Reliability of the physical activity scale for the elderly (PASE). Eur J Sports Sci 2002;2:1-12.

154. Lorig K, Seleznick M, Lubeck D, Ung E, Chastain RL, Holman HR. The beneficial outcomes of the arthritis self-management course are not adequately explained by behavior change. Arthritis Rheum 1989;32(1):91-95.

155. Macfarlane DJ, Lee CC, Ho EY, Chan KL, Chan DT. Reliability and validity of the Chinese version of IPAQ (short, last 7 days). J Sci Med Sport 2007;10(1):45-51.

156. Madsen OR, Bliddal H, Egsmose C, Sylvest J. Isometric and isokinetic quadriceps strength in gonarthrosis; inter-relations between quadriceps strength, walking ability, radiology, subchondral bone density and pain. Clin Rheumatol 1995;14(3):308-314.

157. Maillefert JF, Hawker GA, Gossec L et al. Concomitant therapy: an outcome variable for musculoskeletal disorders? Part 2: total joint replacement in osteoarthritis trials. J Rheumatol 2005;32(12):2449-2451.

158. Mancuso CA, Ranawat CS, Esdaile JM, Johanson NA, Charlson ME. Indications for total hip and total knee arthroplasties. Results of orthopaedic surveys. J Arthroplasty 1996;11(1):34-46.

159. McConnell S, Kolopack P, Davis AM. The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC): a review of its utility and measurement properties. Arthritis Rheum 2001;45(5):453-461.

160. Meijer EP, Goris AH, Wouters L, Westerterp KR. Physical inactivity as a determinant of the physical activity level in the elderly. Int J Obes Relat Metab Disord 2001;25(7):935-939.

161. Michaelsson K, Byberg L, Ahlbom A, Melhus H, Farahmand BY. Risk of severe knee and hip osteoarthritis in relation to level of physical exercise: a prospective cohort study of long-distance skiers in Sweden. PLoS One 2011;6(3):e18339.

162. Migliore A, Bella A, Bisignani M et al. Total hip replacement rate in a cohort of patients affected by symptomatic hip osteoarthritis following intra-articular sodium hyaluronate (MW 1,500-2,000 kDa) ORTOBRIX study. Clin Rheumatol 2012;31(8):1187-1196.

163. Miyaguchi M, Kobayashi A, Kadoya Y, Ohashi H, Yamano Y, Takaoka K. Biochemical change in joint fluid after isometric quadriceps exercise for patients with osteoarthritis of the knee. Osteoarthritis Cartilage 2003;11(4):252-259.

164. Mokkink LB, Terwee CB, Knol DL et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. BMC Med Res Methodol 2010;10:22.

165. Murphy LB, Helmick CG, Schwartz TA et al. One in four people may develop symptomatic hip osteoarthritis in his or her lifetime. Osteoarthritis Cartilage 2010;18(11):1372-1379.

166. Must A, Spadano J, Coakley EH, Field AE, Colditz G, Dietz WH. The disease burden associated with overweight and obesity. JAMA 1999;282(16):1523-1529.

167. Najonalt kompetansesenter for leddproteser. Årsrapport 2008. 2008

168. Najonalt kompetansesenter for leddproteser. Årsrapport 2013. 2013

169. Nelson AE, Allen KD, Golightly YM, Goode AP, Jordan JM. A systematic review of recommendations and guidelines for the management of osteoarthritis: The Chronic Osteoarthritis Management Initiative of the U.S. Bone and Joint Initiative. Semin Arthritis Rheum 2013. [Epub ahead of print]

170. Ngai SP, Cheung RT, Lam PL, Chiu JK, Fung EY. Validation and reliability of the Physical Activity Scale for the Elderly in Chinese population. J Rehabil Med 2012;44(5):462-465.

171. Nilsdotter AK, Lohmander LS, Klassbo M, Roos EM. Hip disability and osteoarthritis outcome score (HOOS)--validity and responsiveness in total hip replacement. BMC Musculoskelet Disord 2003;4:10.

172. Nilsdotter AK, Petersson IF, Roos EM, Lohmander LS. Predictors of patient relevant outcome after total hip replacement for osteoarthritis: a prospective study. Ann Rheum Dis 2003;62(10):923-930.

173. Norkin CC, White DJ. Measurement of Joint Motion: A Guide to Goniometry. Philadelphia, PA: F.A. Davis; 2003.

174. Norwegian Directorate of Health. The action plan on physical activity 2005-2009. Working together for physical activity. 2005

175. Odding E, Valkenburg HA, Algra D, Vandenouweland FA, Grobbee DE, Hofman A. Associations of radiological osteoarthritis of the hip and knee with locomotor disability in the Rotterdam Study. Ann Rheum Dis 1998;57(4):203-208.

176. Olsen IC, Kvien TK, Uhlig T. Consequences of handling missing data for treatment response in osteoarthritis: a simulation study. Osteoarthritis Cartilage 2012;20(8):822-828.

177. Pereira D, Peleteiro B, Araujo J, Branco J, Santos RA, Ramos E. The effect of osteoarthritis definition on prevalence and incidence estimates: a systematic review. Osteoarthritis Cartilage 2011;19(11):1270-1285.

178. Perera S, Mody SH, Woodman RC, Studenski SA. Meaningful change and responsiveness in common physical performance measures in older adults. J Am Geriatr Soc 2006;54(5):743-749.

179. Pincus T, Mitchell JM, Burkhauser RV. Substantial work disability and earnings losses in individuals less than age 65 with osteoarthritis: comparisons with rheumatoid arthritis. J Clin Epidemiol 1989;42(5):449-457.

180. Pisters MF, Veenhof C, de Bakker DH, Schellevis FG, Dekker J. Behavioural graded activity results in better exercise adherence and more physical activity than usual care in people with osteoarthritis: a cluster-randomised trial. J Physiother 2010;56(1):41-47.

181. Pisters MF, Veenhof C, Schellevis FG, de Bakker DH, Dekker J. Long-term effectiveness of exercise therapy in patients with osteoarthritis of the hip or knee: a randomized controlled trial comparing two different physical therapy interventions. Osteoarthritis Cartilage 2010;18(8):1019-1026.

182. Pisters MF, Veenhof C, Schellevis FG, Twisk JW, Dekker J, de Bakker DH. Exercise adherence improving long-term patient outcome in patients with osteoarthritis of the hip and/or knee. Arthritis Care Res (Hoboken) 2010;62(8):1087-1094.

183. Pisters MF, Veenhof C, van Meeteren NL et al. Long-term effectiveness of exercise therapy in patients with osteoarthritis of the hip or knee: a systematic review. Arthritis Rheum 2007;57(7):1245-1253.

184. Poulsen E, Christensen HW, Penny JO, Overgaard S, Vach W, Hartvigsen J. Reproducibility of range of motion and muscle strength measurements in patients with hip osteoarthritis - an inter-rater study. BMC Musculoskelet Disord 2012;13:242.

185. Pua YH, Cowan SM, Wrigley TV, Bennell KL. Discriminant validity of the Western Ontario and McMaster Universities Osteoarthritis index physical functioning subscale in community samples with hip osteoarthritis. Arch Phys Med Rehabil 2009;90(10):1772-1777.

186. Quintana JM, Arostegui I, Azkarate J et al. Evaluation by explicit criteria of the use of total hip joint replacement. Rheumatology (Oxford) 2000;39(11):1234-1241.

187. Quintana JM, Arostegui I, Escobar A, Azkarate J, Goenaga JI, Lafuente I. Prevalence of knee and hip osteoarthritis and the appropriateness of joint replacement in an older population. Arch Intern Med 2008;168(14):1576-1584.

188. Racunica TL, Teichtahl AJ, Wang Y et al. Effect of physical activity on articular knee joint structures in community-based adults. Arthritis Rheum 2007;57(7):1261-1268.

189. Rafamantanantsoa HH, Ebine N, Yoshioka M et al. Validation of three alternative methods to measure total energy expenditure against the doubly labeled water method for older Japanese men. J Nutr Sci Vitaminol (Tokyo) 2002;48(6):517-523.

190. Rasanen P, Paavolainen P, Sintonen H et al. Effectiveness of hip or knee replacement surgery in terms of quality-adjusted life years and costs. Acta Orthop 2007;78(1):108-115.

191. Reiser LM, Schlenk EA. Clinical use of physical activity measures. J Am Acad Nurse Pract 2009;21(2):87-94.

192. Roos EM, Dahlberg L. Positive effects of moderate exercise on glycosaminoglycan content in knee cartilage: a four-month, randomized, controlled trial in patients at risk of osteoarthritis. Arthritis Rheum 2005;52(11):3507-3514.

193. Rothfuss J, Mau W, Zeidler H, Brenner MH. Socioeconomic evaluation of rheumatoid arthritis and osteoarthritis: a literature review. Semin Arthritis Rheum 1997;26(5):771-779.

194. Rothman KJ, Greenland S. Causation and causal inference in epidemiology. Am J Public Health 2005;95 Suppl 1:S144-S150.

195. Rydevik K, Fernandes L, Nordsletten L, Risberg MA. Functioning and disability in patients with hip osteoarthritis with mild to moderate pain. J Orthop Sports Phys Ther 2010;40(10):616-624.

196. Schuit AJ, Schouten EG, Westerterp KR, Saris WH. Validity of the Physical Activity Scale for the Elderly (PASE): according to energy expenditure assessed by the doubly labeled water method. J Clin Epidemiol 1997;50(5):541-546.

197. Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. Lancet 2002;359(9308):781-785.

198. Sellam J, Berenbaum F. Is osteoarthritis a metabolic disease? Joint Bone Spine 2013;80(6):568-573.

199. Shelby RA, Somers TJ, Keefe FJ, Pells JJ, Dixon KE, Blumenthal JA. Domain specific self-efficacy mediates the impact of pain catastrophizing on pain and disability in overweight and obese osteoarthritis patients. J Pain 2008;9(10):912-919.

200. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979;86(2):420-428.

201. Soderman P, Malchau H. Validity and reliability of Swedish WOMAC osteoarthritis index: a self-administered disease-specific questionnaire (WOMAC) versus generic instruments (SF-36 and NHP). Acta Orthop Scand 2000;71(1):39-46.

202. Somers TJ, Keefe FJ, Pells JJ et al. Pain catastrophizing and pain-related fear in osteoarthritis patients: relationships to pain and disability. J Pain Symptom Manage 2009;37(5):863-872.

203. Song J, Semanik P, Sharma L et al. Assessing physical activity in persons with knee osteoarthritis using accelerometers: data from the osteoarthritis initiative. Arthritis Care Res (Hoboken) 2010;62(12):1724-1732.

204. Soucie JM, Wang C, Forsyth A et al. Range of motion measurements: reference values and a database for comparison studies. Haemophilia 2011;17(3):500-507.

205. Steib S, Schoene D, Pfeifer K. Dose-response relationship of resistance training in older adults: a meta-analysis. Med Sci Sports Exerc 2010;42(5):902-914.

206. Steultjens MP, Dekker J, van Baar ME, Oostendorp RA, Bijlsma JW. Range of joint motion and disability in patients with osteoarthritis of the knee or hip. Rheumatology (Oxford) 2000;39(9):955-961.

207. Stratford PW, Kennedy DM. Does parallel item content on WOMAC's pain and function subscales limit its ability to detect change in functional status? BMC Musculoskelet Disord 2004;5:17.

208. Stucki G, Sangha O, Stucki S et al. Comparison of the WOMAC (Western Ontario and McMaster Universities) osteoarthritis index and a self-report format of the self-administered Lequesne-Algofunctional index in patients with knee and hip osteoarthritis. Osteoarthritis Cartilage 1998;6(2):79-86.

209. Tak E, Staats P, Van HA, Hopman-Rock M. The effects of an exercise program for older adults with osteoarthritis of the hip. J Rheumatol 2005;32(6):1106-1113.

210. Taraldsen K, Vereijken B, Thingstad P, Sletvold O, Helbostad JL. Multiple Days of Monitoring are Needed to Obtain a Reliable Estimate of Physical Activity in Hip Fracture Patients. J Aging Phys Act 2013. [Epub ahead of print]

211. Terwee CB, Bouwmeester W, van Elsland SL, de Vet HC, Dekker J. Instruments to assess physical activity in patients with osteoarthritis of the hip or knee: a systematic review of measurement properties. Osteoarthritis Cartilage 2011;19(6):620-633.

212. Terwee CB, Bouwmeester W, van Elsland SL, de Vet HC, Dekker J. Instruments to assess physical activity in patients with osteoarthritis of the hip or knee: a systematic review of measurement properties. Osteoarthritis Cartilage 2011;19(6):620-633.

213. Terwee CB, Mokkink LB, van Poppel MN, Chinapaw MJ, van MW, de Vet HC. Qualitative attributes and measurement properties of physical activity questionnaires: a checklist. Sports Med 2010;40(7):525-537.

214. The Norwegain Directorat of Health. Prioriteringer i helsesektoren - verdigrunnlag, status og utfordringer. 2012

215. Theiler R, Stucki G, Schutz R et al. Parametric and non-parametric measures in the assessment of knee and hip osteoarthritis: interobserver reliability and correlation with radiology. Osteoarthritis Cartilage 1996;4(1):35-42.

216. Thomas JR, Nelson JK. Research Methods in Physical Activity. Leeds, UK: Human Kinetics; 2001.

217. Troiano RP. Large-scale applications of accelerometers: new frontiers and new questions. Med Sci Sports Exerc 2007;39(9):1501.

218. Troosters T, Gosselink R, Decramer M. Six minute walking distance in healthy elderly subjects. Eur Respir J 1999;14(2):270-274.

219. Trost SG, Blair SN, Khan KM. Physical inactivity remains the greatest public health problem of the 21st century: evidence, improved methods and solutions using the '7 investments that work' as a framework. Br J Sports Med 2014;48(3):169-170.

220. Trost SG, O'Neil M. Clinical use of objective measures of physical activity. Br J Sports Med 2014;48(3):178-181.

221. Tubach F, Ravaud P, Baron G et al. Evaluation of clinically relevant changes in patient reported outcomes in knee and hip osteoarthritis: the minimal clinically important improvement. Ann Rheum Dis 2005;64(1):29-33.

222. Urquhart DM, Tobing JF, Hanna FS et al. What is the effect of physical activity on the knee joint? A systematic review. Med Sci Sports Exerc 2011;43(3):432-442.

223. Uthman OA, van der Windt DA, Jordan JL et al. Exercise for lower limb osteoarthritis: systematic review incorporating trial sequential analysis and network meta-analysis. BMJ 2013;347:f5555.

224. van Dijk GM, Dekker J, Veenhof C, van den Ende CH. Course of functional status and pain in osteoarthritis of the hip or knee: a systematic review of the literature. Arthritis Rheum 2006;55(5):779-785.

225. Vaughan K, Miller WC. Validity and reliability of the Chinese translation of the Physical Activity Scale for the Elderly (PASE). Disabil Rehabil 2013;35(3):191-197.

226. Vergara I, Bilbao A, Gonzalez N, Escobar A, Quintana JM. Factors and consequences of waiting times for total hip arthroplasty. Clin Orthop Relat Res 2011;469(5):1413-1420.

227. Vissers MM, Bussmann JB, Verhaar JA, Arends LR, Furlan AD, Reijman M. Recovery of physical functioning after total hip arthroplasty: systematic review and meta-analysis of the literature. Phys Ther 2011;91(5):615-629.

228. Vos T, Flaxman AD, Naghavi M et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012;380(9859):2163-2196.

229. Wang AW, Gilbey HJ, Ackland TR. Perioperative exercise programs improve early return of ambulatory function after total hip arthroplasty: a randomized, controlled trial. Am J Phys Med Rehabil 2002;81(11):801-806.

230. Wang W, Morrison TA, Geller JA, Yoon RS, Macaulay W. Predicting short-term outcome of primary total hip arthroplasty:a prospective multivariate regression analysis of 12 independent factors. J Arthroplasty 2010;25(6):858-864.

231. Wang Y, Wluka AE, Simpson JA et al. Body weight at early and middle adulthood, weight gain and persistent overweight from early adulthood are predictors of the risk of total knee and hip replacement for osteoarthritis. Rheumatology (Oxford) 2013;52(6):1033-1041.

232. Warsi A, LaValley MP, Wang PS, Avorn J, Solomon DH. Arthritis self-management education programs: a meta-analysis of the effect on pain and disability. Arthritis Rheum 2003;48(8):2207-2213.

233. Washburn RA, Ficker JL. Physical Activity Scale for the Elderly (PASE): the relationship with activity measured by a portable accelerometer. J Sports Med Phys Fitness 1999;39(4):336-340.

234. Washburn RA, Smith KW, Jette AM, Janney CA. The Physical Activity Scale for the Elderly (PASE): development and evaluation. J Clin Epidemiol 1993;46(2):153-162.

235. Wesseling J, Bierma-Zeinstra SM, Kloppenburg M, Meijer R, Bijlsma JW. Worsening of pain and function over 5 years in individuals with 'early' OA is related to structural damage: data from the Osteoarthritis Initiative and CHECK (Cohort Hip & Cohort Knee) study. Ann Rheum Dis 2013. [Epub ahead of print]

236. Wood AM, White IR, Hillsdon M, Carpenter J. Comparison of imputation and modelling methods in the analysis of a physical activity trial with missing outcomes. Int J Epidemiol 2005;34(1):89-99.

237. Wood AM, White IR, Thompson SG. Are missing outcome data adequately handled? A review of published randomized controlled trials in major medical journals. Clin Trials 2004;1(4):368-376.

238. Woolf AD, Pfleger B. Burden of major musculoskeletal conditions. Bull World Health Organ 2003;81(9):646-656.

239. World Health Organization. Global recommendations on physical activity for health. Geneva, Switzerland: World Health Organization; 2010.

240. Wright AA, Cook C, Abbott JH. Variables associated with the progression of hip osteoarthritis: a systematic review. Arthritis Rheum 2009;61(7):925-936.

241. Wright AA, Cook CE, Baxter GD, Garcia J, Abbott JH. Relationship between the Western Ontario and McMaster Universities Osteoarthritis Index Physical Function Subscale and physical performance measures in patients with hip osteoarthritis. Arch Phys Med Rehabil 2010;91(10):1558-1564.

242. Yesil H, Hepguler S, Ozturk C, Capaci K, Yesil M. Prevalence of symptomatic knee, hand and hip osteoarthritis among individuals 40 years or older: a study conducted in Izmir city. Acta Orthop Traumatol Turc 2013;47(4):231-235.

243. Zhang W, Doherty M, Arden N et al. EULAR evidence based recommendations for the management of hip osteoarthritis: report of a task force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). Ann Rheum Dis 2005;64(5):669-681.

244. Zhang W, Nuki G, Moskowitz RW et al. OARSI recommendations for the management of hip and knee osteoarthritis: part III: Changes in evidence following systematic cumulative update of research published through January 2009. Osteoarthritis Cartilage 2010;18(4):476-499.

245. Zhang Y, Jordan JM. Epidemiology of osteoarthritis. Rheum Dis Clin North Am 2008;34(3):515-529.

PAPERS I-III

I

II

ARD Online First, published on November 20, 2013 as 10.1136/annrheumdis-2013-203628

OPEN ACCESS

EXTENDED REPORT

Exercise therapy may postpone total hip replacement surgery in patients with hip osteoarthritis: a long-term follow-up of a randomised trial

Ida Svege,[1] Lars Nordsletten,[2,3] Linda Fernandes,[1,4] May Arna Risberg[1]

Handling editor Tore K Kvien

[1]Department of Orthopaedics, Norwegian Research Center for Active Rehabilitation (NAR), Oslo University Hospital, Oslo, Norway
[2]Department of Orthopaedics, Oslo University Hospital, Oslo, Norway
[3]Faculty of Medicine, University of Oslo, Oslo, Norway
[4]Department of Orthopaedic Surgery and Traumatology, Odense University Hospital, Odense C, Denmark

Correspondence to
Ida Svege, Department of Orthopaedics, Oslo University Hospital, Seksjon for forskning og kompetansesenter, Kirkeveien 166, bygg 72 (2. etg), Oslo 0407, Norway; ida.svege@nimi.no

Received 15 March 2013
Revised 4 September 2013
Accepted 26 October 2013

To cite: Svege I, Nordsletten L, Fernandes L, *et al*. *Ann Rheum Dis* Published Online First: [*please include* Day Month Year] doi:10.1136/annrheumdis-2013-203628

ABSTRACT

Background Exercise treatment is recommended for all patients with hip osteoarthritis (OA), but its effect on the long-term need for total hip replacement (THR) is unknown.

Methods We conducted a long-term follow-up of a randomised trial investigating the efficacy of exercise therapy and patient education versus patient education only on the 6-year cumulative survival of the native hip to THR in 109 patients with symptomatic and radiographic hip OA. Results regarding the primary outcome measure of the trial, self-reported pain at 16 months follow-up, have been reported previously.

Results There were no group differences at baseline. The response rate at follow-up was 94%. 22 patients in the group receiving both exercise therapy and patient education and 31 patients in the group receiving patient education only underwent THR during the follow-up period, giving a 6-year cumulative survival of the native hip of 41% and 25%, respectively (p=0.034). The HR for survival of the native hip was 0.56 (CI 0.32 to 0.96) for the exercise therapy group compared with the control group. Median time to THR was 5.4 and 3.5 years, respectively. The exercise therapy group had better self-reported hip function prior to THR or end of study, but no significant differences were found for pain and stiffness.

Conclusions Our findings in this explanatory study suggest that exercise therapy in addition to patient education can reduce the need for THR by 44% in patients with hip OA. ClinicalTrials.gov number NCT00319423 (original project protocol) and NCT01338532 (additional protocol for long-term follow-up).

INTRODUCTION

Physical activity and patient information is recommended for all patients with osteoarthritis (OA) of the hip and knee as first-line treatment. Total joint replacement surgery is to be considered in cases of advanced disease with severe pain and functional limitations where other treatment options have failed.[1 2] Exercise therapy is found to be beneficial in reducing pain and improving function in lower limb OA,[3–6] but evidence for this is primarily based on studies including patients with knee OA. In hip OA, exercise interventions have shown promising results,[7–9] but the need for high-quality clinical trials with sufficient follow-up time is emphasised.[3 4 6] Based on the general consensus that total joint replacement surgery is appropriate only in advanced stages of the disease, joint replacement surgery may be used as an endpoint to evaluate disease progression.[10–14] It is unknown whether exercise therapy can influence the progression of OA and thereby reduce the need for total joint replacement.

The main objective of this study was therefore to evaluate the long-term effect of exercise therapy in addition to patient education on the patient's need for total hip replacement (THR). Our null hypothesis was that there would be no difference in cumulative survival of the native hip to THR in patients with hip OA going through exercise therapy *and* patient education compared with patient education *only*.

METHODS

Study design and patients

This is a long-term follow-up of a randomised, controlled trial evaluating the effect of exercise therapy and patient education in patients with hip OA.[9] Inclusion criteria were age between 40 and 80 years, hip pain for at least 3 months, radiographically verified minimum joint space according to Danielsson's criterion[15] (<4 mm for patients <70 years, <3 mm for patients >70 years) and Harris Hip Score between 60 and 95 points.[16] In patients with bilateral hip OA, the most painful hip was defined as the index joint. Night pain and Harris Hip Score below 60 are used as criteria for THR at our institution.[9] Thus, the patients included in the study were not candidates for THR at the time of inclusion, and none of them were on waiting lists for THR. Exclusion criteria were THR in the index joint, knee pain or knee OA, low back pain, rheumatoid arthritis, osteoporosis, cancer, cardiovascular disease unable to tolerate exercise, dysfunction in lower extremities due to accident or disease, pregnancy and not understanding Norwegian. Patient recruitment and screening for inclusion has been described previously, together with the results of the primary outcome measure for this trial.[9]

Randomisation and treatment groups

All included patients were given three group sessions of a patient education programme developed for patients with hip OA.[17] Thereafter they were randomised to either an exercise therapy group or a control group.[9] A computer-generated randomisation list (block length 10, allocation ratio 1:1) was conducted by a statistician prior to inclusion. Sequentially numbered, sealed envelopes were used to assign treatment for patients consecutively by a

research coordinator not involved in the patient assessment or interventions. Allocation concealment was maintained until written informed consent was obtained, and baseline assessments and patient education sessions were completed. The randomisation sequence was concealed from the study collaborators until treatment was assigned. The exercise therapy programme was specifically designed for patients with hip OA[18] and consisted of strengthening, flexibility and functional exercises. Patients in the exercise therapy group performed the exercise programme two to three times per week for 12 weeks, supervised by a physical therapist at least once weekly. Compliance was based on training diaries filled in weekly by the patients in the exercise therapy group during the 12-week intervention period. Attending at least 20 of a total of 24 sessions was defined as satisfactory adherence. Patients in the control group attended a 2-month follow-up visit at the physiotherapy clinic as part of the patient education programme. They did not have access to the exercise therapy programme during the intervention period.

Outcome measures and follow-up

Characteristics of the patients' included age, gender, height, weight, work status, education level, unilateral or bilateral hip pain, pain duration, minimum joint space and Harris Hip Score.

The main outcome measure for this long-term follow-up was survival of the native hip to THR in the index joint. At inclusion all patients were instructed to report if and when they went through THR surgery during the project period. Additionally, data on THR were recorded at follow-ups 4, 10, 16 and 29 months after inclusion and by contacting all patients by telephone in April and May 2011 (figure 1). The outcome assessor was blinded to group allocation. The mean time from inclusion till the end of study at 15 May 2011 was 4.8 years, ranging from 3.6 to 6.1 years.

The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC)[19] and the Physical Activity Scale for the Elderly (PASE)[20] were filled in at baseline and at the 4-, 10-, 16- and 29-month follow-up. In this long-term follow-up study, WOMAC was used to assess symptoms and functional limitations prior to THR surgery or end of study. PASE is a brief, self-administered, 7-day recall questionnaire to assess physical activity in older adults. The Norwegian version was used, which consisted of 24 questions giving a total score ranging from 0 to 315.[21] Data on training sessions per week were collected at baseline and at 4 months, data on engagement in strength training and flexibility training were collected at 16 and 29 months, and data on physical therapy treatment were collected at 10, 16 and 29 months.

Statistical analysis

Patients were followed until time of THR in the index joint or until death, drop-out or end of study. A Kaplan–Meier survival analysis was constructed to evaluate cumulative 6-year survival, and group difference was tested by the log rank test. THR in the index joint was defined as event, while patients who were lost to follow-up, were dead or were followed until the end of study were treated as censored in the analysis. Time to THR is reported as median and 95% CI. A Cox proportional hazard model was used to calculate HR and 95% CI between groups. No adjusted analysis was conducted due to equality of groups at baseline. Baseline comparisons were performed with Student t tests and χ^2 tests. A linear mixed model (variance component model), with time and the interaction of time and group as fixed effects and time as random effect intercept and slope, was used to compare WOMAC scores between the exercise therapy group and the control group over the 29-month follow-up period. A linear mixed model was also applied to compare WOMAC scores prior to THR surgery or end of study between patients who went through THR and patients who did not. The analyses were based on the intention to treat principle. For the outcome measures of physical activity and exercise, mean (SD) or number was calculated, and a linear mixed model was used to compare PASE scores between the exercise therapy group and the control group. p Values below 0.05 were considered statistically significant.

Analyses were performed by IBM SPSS Statistics, V.19.0 (IBM Corp., Somers, New York, USA).

RESULTS

Characteristics of the patients

Two hundred and twenty patients were screened for eligibility between April 2005 and October 2007. One hundred and nine patients were included in the trial and randomised to the exercise therapy group or the control group (figure 1). Baseline data were similar in the two intervention groups (table 1). The patients completed a median of 20 (IQR 16–24) exercise sessions over the 12-week period, with 53% completing ≥20 exercise sessions. One patient discontinued exercise after three sessions due to increasing hip pain. No other adverse events were registered.

Data on whether THR had been performed were obtained from 102 patients. One patient died and was treated as censored at the time of death. The remaining six patients were treated as censored at the time of last follow-up or contact during the follow-up period. Patients who were censored before the end of study did not differ at baseline from those attending the long-term follow-up.

A total of 41 patients in the exercise therapy group and 30 patients in the control group completed WOMAC at the 29-month follow-up (figure 1). Also, 27 patients had gone through THR prior to the 29-month follow-up and 11 patients were lost to follow-up at the 29-month follow-up.

THRs and cumulative survival of native hip

A total of –22 patients in the exercise therapy group and 31 patients in the control group went through THR within the 3.6–6.1 years follow-up period. Estimated median time to THR was 5.4 (CI 4.5 to 6.2) years in the exercise therapy group and 3.5 (CI 2.3 to 4.6) years in the control group. The Kaplan–Maier analysis showed that the cumulative 6-year survival of the native hip to THR was 0.41 in the exercise therapy group compared with 0.25 in the control group (p=0.034) (figure 2). Cox proportional hazard analysis showed that participating in both exercise therapy and patient education had a protective effect against THR compared with patient education only (HR=0.56, CI 0.32 to 0.96, p=0.036). Thirty-five per cent of the patients went through THR surgery at the Oslo University Hospital, and the remaining 65% went through surgery at 1 of 11 other hospitals in the southern parts of Norway. None of the non-operated patients reported to be on waiting list for THR at the end of study.

Self-reported pain, stiffness and function

Over the 29-month WOMAC follow-up period, the exercise therapy group had significantly better WOMAC physical function scores compared with the control group (p=0.004), but the between-group differences in the WOMAC pain (p=0.083) and WOMAC stiffness (p=0.112) scores did not reach statistical significance (table 2).

Mean minimum joint space at baseline was 1.5±0.9 mm in patients who went through THR compared with 2.5±1.0 mm

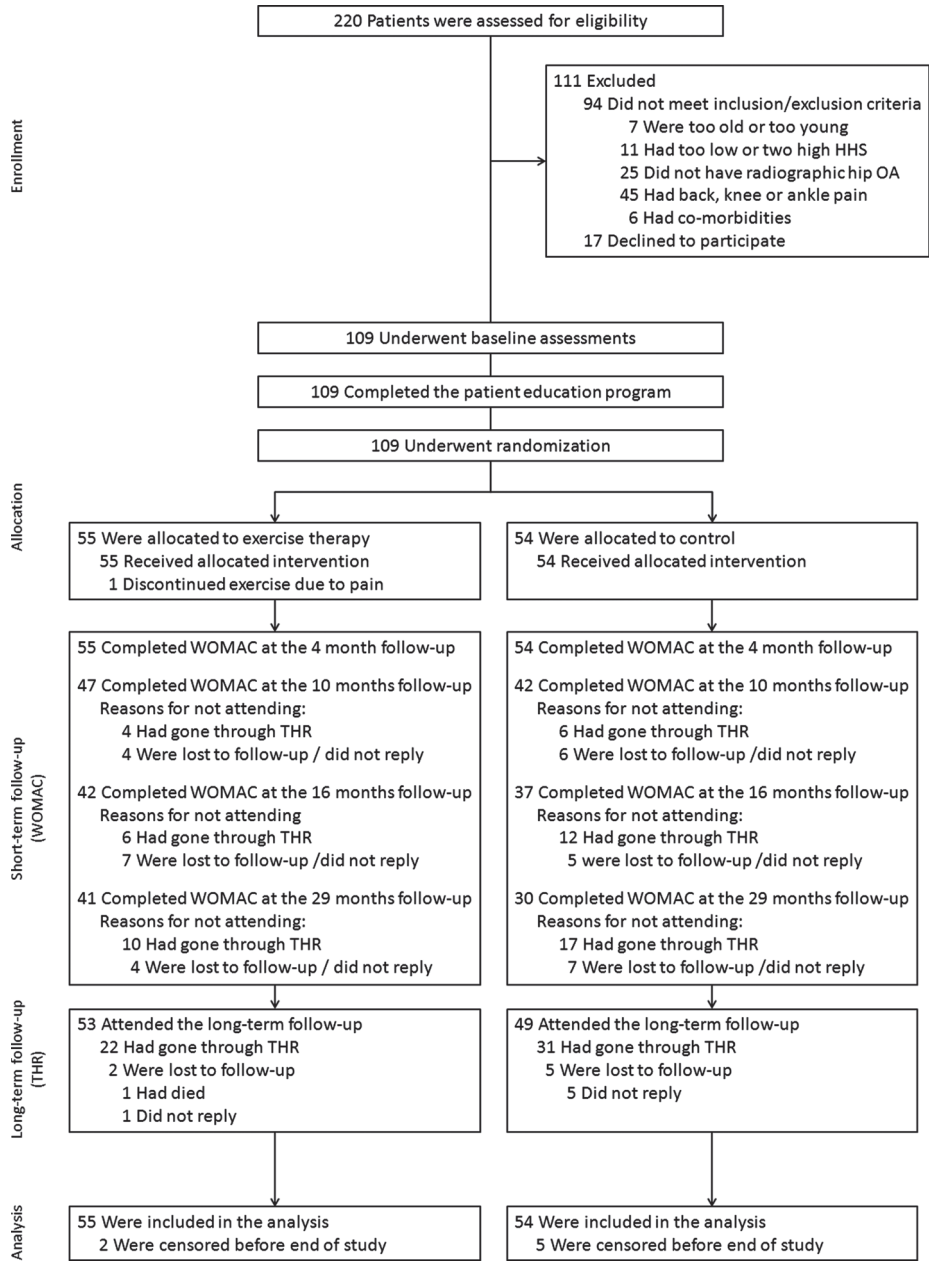

Figure 1 Enrolment, randomisation and follow-up of patients.

in the patients who did not ($p<0.01$). At baseline there were no significant differences between patients who went through THR and patients who did not in neither WOMAC pain ($p=0.967$), WOMAC stiffness ($p=0.333$) nor WOMAC physical function ($p=0.092$). The 53 patients who underwent THR before the end of study had worse preoperative score in all WOMAC subscales over the 29-month WOMAC follow-up period compared with the patients who did not go through THR or were censored at the end of study ($p<0.01$) (table 2).

Self-reported physical activity and exercise

The number of self-reported exercise sessions per week was similar in the two groups. At the 16-month follow-up, 75 patients replied to the questions on exercise and physical

Table 1 Baseline characteristics of the study patients*

	Exercise therapy group (n=55)	Control group (n=54)
Age (years)	58.4±10.0	57.2±9.8
Female sex, no. (%)	31 (56.4)	28 (51.9)
Body mass index (kg/m^2)	24.6±3.2	24.9±3.8
Minimum joint space in target joint† (mm)	2.1±1.0	1.9±1.1
Pain duration (months)	47.3±53.3	49.5±50.9
Harris Hip Score‡	79.6±7.7	76.9±8.2
Bilateral radiographic hip OA, no. (%)	38 (69.1)	38 (70.3)
THR in contralateral hip at inclusion, no. (%)	4 (7.3)	2 (3.7)
Hereditary OA/known OA in family, no. (%)	17 (33.3)	21 (38.9)
>12 years of education, no. (%)	43 (78.2)	35 (67.3)
Work status		
Employed, no. (%)	35 (63.6)	36 (66.7)
Sick leave, no. (%)	8 (14.5)	5 (9.3)
Retired, no. (%)	12 (21.8)	9 (16.7)
WOMAC score§		
Pain subscale	26.0±16.1	27.3±17.9
Stiffness subscale	34.8±23.7	34.3±20.5
Physical function subscale	21.1±15.3	23.6±15.7

*Plus-minus values are mean±SD. The body mass index is the weight in kilograms divided by the square of the height in metres.
†The minimum joint space in the hip joint was assessed according to Danielsson's criterion.[15] For patients older than 70 years, a minimum joint space below 3 mm was characterised as radiographic hip OA. For patients younger than 70 years, a minimum joint space below 4 mm was characterised as radiographic hip OA.
‡The Harris Hip Score is a clinician-administered tool to evaluate hip pain, hip function and hip range of motion.[16] An overall score is calculated ranging from 0 to 100, with a lower score indicating more severe disease.
§The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) comprise three subscales (pain, stiffness and physical function) composed of 24 questions. Scores range from 0 to 100, with a higher score indicating more severe disease.[19]
OA, osteoarthritis; THR, total hip replacement.

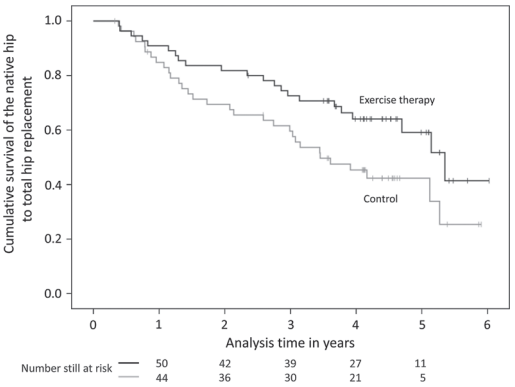

Figure 2 Kaplan–Meier survival estimates over the 6-year follow-up period. The black line represents the exercise therapy group, and the grey line represents the control group. Censored data are marked at each line. The number of patients at risk is given for each year for each group.

therapy, and at the 29-month follow-up 70 patients replied (table 3). There was no significant difference in PASE scores between the exercise therapy group and the control group over the 29-month follow-up period (p=0.397).

DISCUSSION

Participating in both exercise therapy and patient education resulted in significantly higher 6-year cumulative survival of the native hip to THR compared with patient education only. Thus, the null hypothesis was rejected. The cumulative survival of the native hip was higher in the exercise therapy group from 1 year and throughout the follow-up period.

This is the first study to evaluate whether exercise therapy affects the need for THR in patients with isolated hip OA. One previous study has used total joint replacement as an outcome to compare the effect of individually tailored exercises and usual care in knee and/or hip OA.[22] They found that 20% underwent THR in the individually tailored exercise therapy group compared with 45% in the usual care group for those with hip OA. The probability for THR within 5 years was 2.87 (95% CI 1.1 to 7.3) times higher in the usual care group.[22] In our study, 40% in the exercise therapy group and 57% in the control group underwent THR, with the control group having 1.80 times higher probability of THR. The somewhat smaller protective effect of exercise in our study may be due to the patients having both symptomatic and radiographic hip OA. Pisters *et al*[22] based inclusion on the clinical criteria of the American College of Rheumatology alone, which does not include radiographic evidence of OA.

Previous studies have reported that 24–53% of patients with symptomatic and radiographic hip OA undergo THR during follow-up ranging from 14 months to 10 years.[15 23 24] THR rates have increased steadily during the past four decades,[25] which in turn has enlarged healthcare costs substantially.[26] Our finding, that exercise therapy enhances the survival of the native hip to THR, is therefore important for healthcare consumption and for patients who may avoid surgery and its potential complications. Some studies have recommended and used total joint replacement as a hard endpoint in OA,[10 27–29] but it is debatable whether it can be interpreted as an expression for OA progression. Attempts are requested[26] and have been made,[30] but still no clearly defined criteria for THR exist. Worse self-reported pain and functional limitations are associated with a higher THR rate,[31] but cannot be used to discriminate between patients who are or are not in need of a THR as clinical severity varies widely.[30 32] In our study, the patients who went through THR had poorer scores in the WOMAC subscales for pain, stiffness and physical function prior to THR compared with the patients who did not undergo THR. This supports the assumption that the patients who undergo THR surgery have more severe symptoms and functional limitations. Also, the patients who went through THR had smaller minimum joint space at baseline. Abadie *et al*[13] stated that THR is probably the most relevant clinical endpoint for evaluating effect of disease-modifying treatment, but it is potentially biased by non-disease-related factors such as economic factors, availability and geographical differences, comorbidities and contraindications for surgery, and willingness to undergo surgery.[12 13] However, in a randomised design study, equal distribution of potential confounding factors is assumed.

Other studies have found beneficial short-term effects of exercise therapy.[7 8] No significant difference in self-reported pain was demonstrated in the 16-month follow-up of our trial, but the patients in the exercise therapy group had better self-reported physical function compared with the control group.[9] This was supported by the findings in our study, with the exercise therapy group demonstrating better results in WOMAC

Table 2 Difference in self-reported pain, stiffness and function at baseline and at the 4-, 10-, 16- and 29-month follow-up between the exercise therapy group and the control group, and between the patients who went through THR surgery and the patients who did not*

	Baseline	4 months	10 months	16 months	29 months
Mean difference (95% CI) between the exercise therapy group and the control group					
WOMAC†					
Pain	−1.3 (−8.0 to 5.3)	−4.7 (−11.4 to 1.9)	−6.6 (−13.9 to 0.8)	−6.5 (−14.3 to 1.3)	−5.9 (−14.2 to 2.4)
Stiffness	0.5 (−8.0 to 9.1)	−3.5 (−12.0 to 5.0)	−6.3 (−15.8 to 3.2)	−12.5 (−22.5 to −2.5)	−3.9 (−14.6 to 6.7)
Physical function	−2.5 (−8.7 to 3.7)	−4.6 (−10.7 to 1.6)	−8.4 (−15.2 to −1.6)	−9.2 (−16.5 to −1.9)	−6.4 (−14.1 to −1.3)
Mean difference (95% CI) between the patients who underwent THR‡ (n=53) and the patients who did not (n=56)					
WOMAC†					
Pain	0.1 (−6.4 to 6.7)	5.6 (−0.9 to 12.1)	11.9 (4.7 to 19.1)	9.3 (1.5 to 17.1)	13.2 (4.6 to 21.8)
Stiffness	4.1 (−4.3 to 12.5)	9.5 (1.1 to 17.8)	10.6 (1.3 to 19.9)	15.2 (5.2 to 25.2)	12.7 (1.7 to 23.8)
Physical function	5.0 (−0.9 to 11.0)	8.9 (2.9 to 14.8)	11.9 (5.3 to 18.4)	13.3 (6.2 to 20.4)	15.1 (7.3 to 23.0)

*Plus-minus values are mean±SD.
†The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) comprise three subscales (pain, stiffness and physical function) composed of 24 questions. Scores range from 0 to 100, with a higher score indicating more severe disease.[19]
‡Results for patients who went through THR are preoperative results up until time of surgery.
THR, total hip replacement.

physical function compared with the control group over the complete 29-month follow-up period (p=0.004), The differences in WOMAC pain (p=0.083) and WOMAC stiffness (p=0.112) did not reach statistical significance. This may indicate that the lower rate and longer time to THR in the exercise therapy group are due to better hip function, with or without the presence of pain. Ten patients in the exercise therapy group and 17 patients in the control group had gone through THR surgery prior to the 29-month follow-up, and it is not unlikely that this uneven distribution of performed THRs has biased the WOMAC results, giving an underestimation of the treatment effect of exercise therapy. Pisters *et al*[22] found no long-term differences in pain and function when comparing individually tailored exercises and usual care, but suggested that patients who underwent THR may have biased the results. Fifty-three per cent of the patients in the exercise therapy group completed ≥20 exercise sessions and were thus regarded as compliant. Data on continuation of the exercise therapy programme after the 12-week intervention period were not obtained, and this must be regarded as a limitation of the study. However, the data on physical activity, exercise and physical therapy treatment suggest that no major between-group differences were present. Self-reported outcome measures lack validity for measuring physical activity and exercise due to recall bias and overestimation of time, frequency and intensity,[33] and these data should therefore be interpreted with caution. Better adherence to exercises has been shown to improve long-term results,[34] and higher leisure time physical activity may have a protective effect against THR.[35]

Our study had some limitations. The criteria for when THR surgery was indicated were not specified prior to the start of the study. The criteria used for THR at our institution (night pain and Harris Hip Score below 60 points) are not necessarily used at other hospitals, and the symptom state may differ at time of surgery. Preoperative assessment was not conducted, but pain and physical function were assessed with a mean time of 0.7±0.8 years prior to THR. Calculation of statistical power for this study was not based on survival of the native hip to THR, but rather the WOMAC pain subscale, which was the primary outcome measure of this trial.[9]

Some caution should be taken when interpreting these results. Our findings are applicable for patients with symptomatic and radiographic hip OA, with mild to moderate symptoms. Patients with severe symptoms and patients with knee or back pain were excluded. Patients recruited to non-surgical treatment trials may have a stronger desire to avoid surgery compared with the general OA population.[12] It is debatable whether postponing

Table 3 Self-reported physical activity in the exercise therapy group and the control group at baseline and at the 4-, 10-, 16- and 29-month follow-up*

	Baseline	4 months	10 months	16 months	29 months
Exercise therapy group					
PASE score†	114±43.5	115±52.9	118±48.6	123±50.7	120±46.8
Exercise sessions per week	3.2±2.0	3.7±1.9			
Engaged in strength training—*no*				22	21
Engaged in flexibility training—*no*				29	27
Physical therapy treatment—*no*			14	16	14
Control group					
PASE score†	123±50.6	121±45.4	126±57.3	133±57.3	139±59.2
Exercise sessions per week	3.2±2.1	3.7±2.0			
Engaged in strength training—*no*				24	18
Engaged in flexibility training—*no*				25	22
Physical therapy treatment—*no*			18	13	20

*Plus-minus values are mean±SD.
†The Physical Activity Scale for the Elderly (PASE) consists of 24 questions on physical activity and the total score expresses the overall physical activity level. Scores range from 0 to 315, with 0 indicating complete inactivity and 315 indicating extremely high level of activity.[20 21]

surgery is beneficial for the patients in the long term.[36 37] We argue that for patients with tolerable pain who are able to maintain their desired activity level and who are relatively young postponing surgery is appropriate and may reduce the future need for THR or repetitive THR revision surgery.

CONCLUSIONS

Our findings in this explanatory study show that participating in a 12-week exercise therapy programme in addition to patient education can reduce the need for THR or postpone surgery in patients with hip OA. This supports the recommendations stating that exercise therapy should be offered to patients with hip OA as first-line treatment.

Correction notice This article has been corrected since it was published Online First. The affiliation of the last author has been corrected.

Acknowledgements We thank all the patients who participated in this trial; the physiotherapists Karin Rydevik, Sigmund Ruud, Emilie Jul-Larsen Aas, Marie Moltubakk and Marte Lund for their contribution in leading the patient education programme, supervising the patients during the exercise therapy intervention or conducting the outcome assessments; the research coordinator Kristin Bølstad for study management; and Leiv Sandvik and Ingar Holme for help and advice on statistical procedure and analysis. The Norwegian Sports Medicine Clinic (NIMI), Oslo, Norway, is acknowledged for supporting the Norwegian Research Center for Active Rehabilitation with rehabilitation facilities and research staff.

Contributors All authors contributed to the development of the protocol and were involved in the conduct of the study. MAR was responsible for the project funding. LN and LF contributed to clinical screening of participants and LN assessed all radiographs. LF and IS carried out the outcome assessments. IS performed the statistical analysis and drafted the manuscript. All authors revised and approved the final manuscript. MAR acts as a guarantor for this study.

Funding This study was supported by grants from the former science council at Ullevaal University Hospital, Oslo, and the EXTRA funds from the Norwegian Foundation for Health and Rehabilitation, through the Norwegian Rheumatism Association.

Competing interests None.

Patient consent Obtained

Ethics approval The study was approved by the regional medical research ethics committee and was carried out in compliance with the Helsinki Declaration.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Dieppe PA, Lohmander LS. Pathogenesis and management of pain in osteoarthritis. *Lancet* 2005;365:965–73.
2 Lohmander LS, Roos EM. Clinical update: treating osteoarthritis. *Lancet* 2007;370:2082–4.
3 Zhang W, Nuki G, Moskowitz RW, *et al*. OARSI recommendations for the management of hip and knee osteoarthritis: part III: Changes in evidence following systematic cumulative update of research published through January 2009. *Osteoarthritis Cartilage* 2010;18:476–99.
4 Fransen M, McConnell S, Hernandez-Molina G, *et al*. Exercise for osteoarthritis of the hip. *Cochrane Database Syst Rev* 2009;3:CD007912.
5 Fransen M, McConnell S. Exercise for osteoarthritis of the knee. *Cochrane Database Syst Rev* 2008;4:CD004376.
6 Hernandez-Molina G, Reichenbach S, Zhang B, *et al*. Effect of therapeutic exercise for hip osteoarthritis pain: results of a meta-analysis. *Arthritis Rheum* 2008;59:1221–8.
7 Tak E, Staats P, Van HA, *et al*. The effects of an exercise program for older adults with osteoarthritis of the hip. *J Rheumatol* 2005;32:1106–13.
8 Juhakoski R, Tenhonen S, Malmivaara A, *et al*. A pragmatic randomized controlled study of the effectiveness and cost consequences of exercise therapy in hip osteoarthritis. *Clin Rehabil* 2010;25:370–83.
9 Fernandes L, Storheim K, Sandvik L, *et al*. Efficacy of patient education and supervised exercise vs patient education alone in patients with hip osteoarthritis: a single blind randomized clinical trial. *Osteoarthritis Cartilage* 2010;18:1237–43.
10 Dougados M, Gueguen A, Nguyen M, *et al*. Requirement for total hip arthroplasty: an outcome measure of hip osteoarthritis? *J Rheumatol* 1999;26:855–61.
11 Maillefert JF, Hawker GA, Gossec L, *et al*. Concomitant therapy: an outcome variable for musculoskeletal disorders? Part 2: total joint replacement in osteoarthritis trials. *J Rheumatol* 2005;32:2449–51.
12 Altman RD, Abadie E, Avouac B, *et al*. Total joint replacement of hip or knee as an outcome measure for structure modifying trials in osteoarthritis. *Osteoarthritis Cartilage* 2005;13:13–19.
13 Abadie E, Ethgen D, Avouac B, *et al*. Recommendations for the use of new methods to assess the efficacy of disease-modifying drugs in the treatment of osteoarthritis. *Osteoarthritis Cartilage* 2004;12:263–8.
14 Dieppe P. Osteoarthritis: time to shift the paradigm. This includes distinguishing between severe disease and common minor disability. *BMJ* 1999;318:1299–300.
15 Danielsson LG. Incidence and prognosis of coxarthrosis. 1964. *Acta Orthop Scandi* 1964;66(Suppl):61–114.
16 Harris WH. Traumatic arthritis of the hip after dislocation and acetabular fractures: treatment by mold arthroplasty. An end-result study using a new method of result evaluation. *J Bone Joint Surg Am* 1969;51:737–55.
17 Klassbo M, Larsson G, Harms-Ringdahl K. Promising outcome of a hip school for patients with hip dysfunction. *Arthritis Rheum* 2003;49:321–7.
18 Fernandes L, Storheim K, Nordsletten L, *et al*. Development of a therapeutic exercise program for patients with osteoarthritis of the hip. *Phys Ther* 2010;90:592–601.
19 Bellamy N, Buchanan WW, Goldsmith CH, *et al*. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 1988;15:1833–40.
20 Washburn RA, Smith KW, Jette AM, *et al*. The Physical Activity Scale for the Elderly (PASE): development and evaluation. *J Clin Epidemiol* 1993;46:153–62.
21 Loland NW. Reliability of the physical activity scale for the elderly (PASE). *Eur J Sports Sci* 2002;2:1–12.
22 Pisters MF, Veenhof C, Schellevis FG, *et al*. Long-term effectiveness of exercise therapy in patients with osteoarthritis of the hip or knee: a randomized controlled trial comparing two different physical therapy interventions. *Osteoarthritis Cartilage* 2010;18:1019–26.
23 Ledingham J, Dawson S, Preston B, *et al*. Radiographic progression of hospital referred osteoarthritis of the hip. *Ann Rheum Dis* 1993;52:263–7.
24 Lane NE, Nevitt MC, Hochberg MC, *et al*. Progression of radiographic hip osteoarthritis over eight years in a community sample of elderly white women. *Arthritis Rheum* 2004;50:1477–86.
25 Singh JA, Vessely MB, Harmsen WS, *et al*. A population-based study of trends in the use of total hip and total knee arthroplasty, 1969–2008. *Mayo Clin Proc* 2010;85:898–904.
26 Ghomrawi HM, Schackman BR, Mushlin AI. Appropriateness criteria and elective procedures—total joint arthroplasty. *N Engl J Med* 2012;367:2467–9.
27 Dougados M, Nguyen M, Berdah L, *et al*. Evaluation of the structure-modifying effects of diacerein in hip osteoarthritis: ECHODIAH, a three-year, placebo-controlled trial. Evaluation of the Chondromodulating Effect of Diacerein in OA of the Hip. *Arthritis Rheum* 2001;44:2539–47.
28 Bruyere O, Honore A, Ethgen O, *et al*. Correlation between radiographic severity of knee osteoarthritis and future disease progression. Results from a 3-year prospective, placebo-controlled study evaluating the effect of glucosamine sulfate. *Osteoarthritis Cartilage* 2003;11:1–5.
29 Migliore A, Bella A, Bisignani M, *et al*. Total hip replacement rate in a cohort of patients affected by symptomatic hip osteoarthritis following intra-articular sodium hyaluronate (MW 1,500–2,000 kDa) ORTOBRIX study. *Clin Rheumatol* 2012;31:1187–96.
30 Gossec L, Paternotte S, Bingham CO III, *et al*. OARSI/OMERACT initiative to define states of severity and indication for joint replacement in hip and knee osteoarthritis. An OMERACT 10 Special Interest Group. *J Rheumatol* 2011;38:1765–9.
31 Lofvendahl S, Bizjajeva S, Ranstam J, *et al*. Indications for hip and knee replacement in Sweden. *J Eval Clin Pract* 2011;17:251–60.
32 Dieppe P, Judge A, Williams S, *et al*. Variations in the pre-operative status of patients coming to primary hip replacement for osteoarthritis in European orthopaedic centres. *BMC Musculoskelet Disord* 2009;10:19.
33 Svege I, Kolle E, Risberg MA. Reliability and validity of the Physical Activity Scale for the Elderly (PASE) in patients with hip osteoarthritis. *BMC Musculoskelet Disord* 2012;13:26.
34 Pisters MF, Veenhof C, Schellevis FG, *et al*. Exercise adherence improving long-term patient outcome in patients with osteoarthritis of the hip and/or knee. *Arthritis Care Res (Hoboken)* 2010;62:1087–94.
35 Ageberg E, Engstrom G, Gerhardsson de Verdier M, *et al*. Effect of leisure time physical activity on severe knee or hip osteoarthritis leading to total joint replacement: a population-based prospective cohort study. *BMC Musculoskelet Disord* 2012;13:73.
36 Fortin PR, Penrod JR, Clarke AE, *et al*. Timing of total joint replacement affects clinical outcomes among patients with osteoarthritis of the hip or knee. *Arthritis Rheum* 2002;46:3327–30.
37 Vergara I, Bilbao A, Gonzalez N, *et al*. Factors and consequences of waiting times for total hip arthroplasty. *Clin Orthop Relat Res* 2011;469:1413–20.

III

Svege et al. BMC Musculoskeletal Disorders 2012, 13:26
http://www.biomedcentral.com/1471-2474/13/26

BMC Musculoskeletal Disorders

RESEARCH ARTICLE Open Access

Reliability and validity of the Physical Activity Scale for the Elderly (PASE) in patients with hip osteoarthritis

Ida Svege[1*†], Elin Kolle[2†] and May Arna Risberg[1,2†]

Abstract

Background: Physical activity (PA) is beneficial in reducing pain and improving function in lower limb osteoarthritis (OA), and is recommended as a first line treatment. Self-administered questionnaires are used to assess PA, but knowledge about reliability and validity of these PA questionnaires are limited, in particular for patients with OA. The purpose of this study was to evaluate the reliability and validity of the Physical Activity Scale for the Elderly (PASE) in patients with hip OA.

Methods: Forty patients with hip OA (20 men and 20 women, mean age 61.3 ± 10 years) were included. For test-retest reliability PASE was administered twice with a mean time between tests of 9 ± 4 days. Intraclass correlation coefficient (ICC), standard error of measurement (SEM) and minimal detectable change (MDC) were calculated for the total score and for the particular items assessing different PA intensity levels. In addition a Bland-Altman analysis for the total PASE score was performed. Construct validity was evaluated by comparing the PASE results with the Actigraph GT1M accelerometer and the International Physical Activity Questionnaire (IPAQ), using the Spearman rank correlation coefficient.

Results: ICC for the total PASE score was 0.78, with relatively large error of measurement; SEM = 31 and MDC = 87. ICC for the intensity items was 0.20 for moderate PA intensity, 0.46 for light PA intensity and to 0.68 for vigorous PA intensity. The Spearman rank correlation coefficient between the Actigraph GT1M total counts per minute and the total PASE score was 0.30 (p = 0.089), and ranging from 0.20-0.38 for the different PA intensity categories. The Spearman rank correlation between IPAQ and PASE was 0.61 (p = 0.001) for the total scores.

Conclusions: In patients with hip OA the test-retest reliability of the total PASE score was moderate, with acceptable ICC, but with large measurement errors. The construct validity of the PASE was poor when compared to the Actigraph GT1M accelerometer. Test-retest reliability and construct validity revealed that the PASE was unable to assess PA intensity levels. PASE is not recommended as a valid tool to examine PA level for patients with hip OA.

Background

Physical inactivity is considered to be a risk factor for many life-threatening diseases and regarded as a major burden on general public health, therefore international and national guidelines recommend that all adults engage in moderate to vigorous physical activity (MVPA) for at least 30 minutes per day[1-3]. Patients with OA are found to be less physically active than the general adult population, and fewer fulfill the recommendations of 30 minutes MVPA per day[4,5]. Being physically active according to the recommended guidelines is beneficial in preserving function and reduce symptoms[6], and PA is recommended as a first line treatment that should be offered to all individuals with hip or knee OA[7,8]. The efficacy and importance of PA and exercise for patients with OA of the lower limbs have been emphasized in several studies[9-12].

* Correspondence: ida.svege@hjelp24.no
† Contributed equally
[1]Norwegian Research Center for Active Rehabilitation, Department of Orthopedics, Oslo University Hospital, Ullevaal and Hjelp 24 NIMI, Oslo, Norway
Full list of author information is available at the end of the article

Valid and reliable methods for PA assessment are essential for studying its health effects. Frequency, duration and intensity are important factors when evaluating PA as a protective factor against OA progression and functional decline[13]. Numerous methods for assessing PA are available, and can be categorized into three main groups; self-reported assessments (questionnaires, rating scales, diaries), activity monitors (accelerometers, pedometers, heart rate monitors) and direct assessment of energy expenditure (doubly labelled water, indirect calorimetry). Self-administered questionnaires, including the Physical Activity Scale for the Elderly (PASE), can potentially capture all types of activities and allow grading by intensity. They are widely used, due to being inexpensive and easy to administer, and are considered particularly useful in large epidemiological and longitudinal studies. However, questionnaires have obvious weaknesses considering recall and reporting bias. In contrast, accelerometers offer a method for measuring body acceleration, and thereby quantify amount and intensity of movement[14]. Accelerometers often serve as a comparator when validity of questionnaires is evaluated, as they are expected to measure the same construct[15].

Despite the fact that many self-administered questionnaires are available, evidence for validity and reliability is limited [13]. PASE has been found to significantly correlate in expected directions with physical performance, knee pain and knee functioning in patients with knee pain[6,16], and previous studies have reported correlation coefficients of 0.16, 0.43 and 0.49 when compared to an accelerometer in the general, elderly population[17-19]. However, the validity of PASE has not been evaluated in patients with hip OA by comparing it to an accelerometer. The purpose of this study was therefore to evaluate the construct validity and the test-retest reliability of the Norwegian version of the Physical Activity Scale for the Elderly (PASE) in patients with hip OA.

Methods

Subjects

Forty patients with hip OA from a larger ongoing randomized controlled trial (RCT), evaluating the effect of patient education and supervised exercise in patients with hip OA[20], were included. Inclusion criteria were age between 40 and 80 years, uni- or bilateral hip pain for more than three months, Harris Hip Score[21] between 60 and 95, and radiographically verified hip OA according to Danielsson's criteria[22]. Patients with low back pain or knee pain, trauma or functional impairments, or diseases that might interfere with participation were excluded. Patients who had gone through total hip replacement surgery (THR) since inclusion in the RCT were also excluded. During September 2010, 61 patients who had been included in the original RCT between 2006 and 2008, were re-contacted and requested to participate in this validation study. Twelve patients did not respond, eight had gone through THR surgery and one lived abroad. The remaining 40 patients agreed to participation and were included in the study.

Anthropometrical (age, gender, height, weight) and sociodemographic data (work status, educational level), as well as data on Harris Hip Score, minimal joint space width, bilateral hip pain and pain duration was recorded at time of inclusion in the original RCT. Data on age has been altered to reflect the actual age at the time of data collection in this validation study.

The study was approved by The Regional Committee for Medical Research Ethics for South-Eastern Norway. All participants received both oral and written information and signed a written informed consent, before inclusion. The data collection was carried out in accordance with the directives given in the Declaration of Helsinki.

Outcome measurements

PASE is a brief, self-administered, 7-day recall questionnaire designed to assess PA in older adults[23]. It has also been used in studies assessing PA in patients with OA[24,25]. In this study we used the Norwegian version of the PASE, which was slightly adapted when translated due to cultural differences[26], i.e. the question in the original version addressing walking activities was incorporated in the three questions addressing light, moderate and vigorous PA activity. It consists of 24 questions in total and the overall PASE score ranges from 0-315 (and above). The instructions for use and scoring given in the PASE Administration and Scoring Manual were followed (http://www.neri.org). The questions included in PASE address leisure-time, household and work-related PA, with the different items weighted differently. Participation in leisure-time PA, including light, moderate and vigorous PA intensity, and strengthening activities, is recorded as never, seldom (1-2 days per week), sometimes (3-4 days per week), and often (5-7 days per week). Duration is categorized as less than 1 hour, 1-2 hours, 2-4 hours and more than 4 hours. Housework activities are recorded as yes or no, and paid or unpaid work, requiring some PA, is recorded in hours/week. The total PASE score is computed by multiplying time spent in each activity (hours per day) (for leisure and work-related activities) or participation (yes/no) in an activity (for household-related activities), by empirically derived weighting, and then summarizing all items[26]. From the PASE recordings we calculated the total PASE score, representing the overall activity level. In addition we calculated the PASE score for household-/work-

related activities and the PASE score for leisure-time PA, as well as the PASE score from the items addressing light, moderate and vigorous PA intensity.

Construct validity of the PASE was evaluated by comparing it to the Actigraph GT1M accelerometer (ActiGraph, LLC, Pensacola, FL, USA) and to the short form of the International Physical Activity Questionnaire (IPAQ). The Actigraph GT1M is an electronic motion sensor comprising a single plane (vertical) accelerometer. Movement in the vertical plane is detected as a combined function of the frequency and intensity of the movement. Counts are summed over 10 second epochs and downloaded to memory. All sequences of 60 minutes or more of consecutive zero counts were excluded from each individuals recording. For the analyses, a valid day was defined as having 10 or more hours of monitor wear. Six or more valid days of registration were considered sufficient. Accelerometers were initialized and downloaded using the software program ActiLife (ActiGraph, LLC, Pensacola, FL, US). Data were reduced using the SAS-based software program (SAS Institute Inc., Cary, North Carolina, USA) called CSA Analyzer (csa.svenssonsport.dk). From the Actigraph GT1M registrations we calculated average counts per minute representing the overall activity level. In addition we calculated total minutes spent in 0-99 counts per minute, 100-2019 counts per minute, 2020-5999 counts per minute and above 6000 counts per minute, representing minutes spent inactive, and in light, moderate and vigorous PA intensity, respectively[27,28]. The proportion of patients who achieved the recommended 30 minutes of daily MVPA was established by dividing total time in MVPA by the number of valid days of recording, giving an average (minutes per day) across the assessment period.

The development of the IPAQ was initiated in 1996, and conducted by an International Consensus Group, with the intention to develop a measure suitable for assessing population levels of PA across countries[29]. IPAQ is a short, self-administered, 7-day recall questionnaire designed for assessing PA in adults. It consists of seven questions which include PA in all contexts of everyday life, and addresses days, hours and minutes spent on vigorous PA, moderate PA and walking. A question on sitting hours per day is also included. The IPAQ is scored by using the Metabolic Equivalent of Task (MET) method, where different activities and levels of intensity are given different MET estimates. In this study the Norwegian version of the IPAQ short form was used, as well as instructions given in the IPAQ Scoring Protocol, both described at http://www.ipaq.ki.se. For the IPAQ we calculated the total MET-minutes per week, representing the overall activity level. In addition we calculated MET-minutes per week for walking activities, moderate activities and vigorous activities.

Procedures

Data collection for the evaluation of test-retest reliability and construct validity was carried out during October 2010. The Actigraph GT1M was administered by postal mail to all included patients, and it was worn in an elastic belt placed on the right hip. All participants were instructed to wear the accelerometer during all waking hours, except during bathing and swimming, over a period of seven consecutive days (1^{th} -7^{th} day), se Figure 1. The questionnaires, PASE and IPAQ, was administered to the participants by mail on the 7^{th} day, and filled in on the 8^{th} day, the day after finishing the accelerometry registration period, and returned by mail. For evaluation of test-retest reliability PASE was also filled out seven days later (on the 15^{th} day).

Analysis

Baseline characteristics and descriptive data for the Actigraph GT1M, the PASE and the IPAQ calculations are presented as mean and standard deviation (SD) or number and percentage (%). To evaluate the test-retest reliability for the total PASE score the intraclass correlation coefficient ($ICC_{2.1}$ - two-way random effect model, absolute agreement) was calculated. In addition, $ICC_{2.1}$ was calculated for the sub-score for household/work-related PA, the sub-score for leisure-time PA, and for the PASE score of the items for light, moderate and vigorous PA intensity. Measurement error was assessed by estimating the standard error of measurement (SEM), minimal detectable change (MDC) and limits of agreement (LoA). SEM was calculated as the square root of the within-subject total variance of an ANOVA analysis, $SEM = \sqrt{var_{tot}}$, and the MDC was calculated as $MDC = 1.96 \times \sqrt{2} \times SEM$ [30]. LoA were calculated according to the Bland-Altman method and a Bland Altman plot for visual judgment of the relationship between the individual mean total PASE score of the test and retest, and the difference in total PASE score between test and retest was made[31].

The construct validity of the PASE was evaluated by calculating the Spearmans rank correlation coefficients (ρ) for the total PASE score and the Actigraph GT1M (total counts per minute), and for the total PASE score and the total IPAQ score (total MET-minutes per week). A priori hypotheses were made based on previous studies comparing PA questionnaires and PA measured by accelerometry. As recommended by Terwee et al.[15], the most similar constructs of the PASE and the Actigraph GT1M were compared. We hypothesized a low to moderate positive correlation (ρ between 0.15 and 0.5) between the total PASE score and the Actigraph GT1M counts per min. We hypothesized a moderate to strong positive correlation (ρ between 0.6 and 0.9) between the total PASE score and the IPAQ total

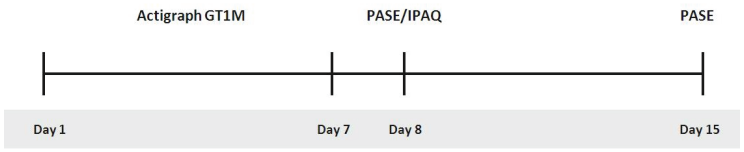

Figure 1 Schematic view of the timeline of the study. *PASE*: Physical Activity Scale for the Elderly; *IPAQ*: International Physical Activity Questionnaire.

MET-minutes per week. Terwee et al.[15] suggested that the correlation between a PA questionnaire (total score) and accelerometry (counts per minute) should exceed 0.5. We therefore interpreted this as a cut-off for acceptable validity.

In addition, Spearmans ρ were calculated for the PASE items for light, moderate and vigorous PA intensity and the different intensity levels/categories assessed by the Actigraph GT1M and IPAQ. For these comparisons the approach was more explorative, but the PASE score for the different intensity items were hypothesized to correlate most strongly with the respective categories of the Actigraph GT1M and the IPAQ as follows: 1) the PASE light PA intensity with the Actigraph GT1M minutes of light PA intensity and the IPAQ walking MET-minutes per week, 2) the PASE moderate PA intensity with the Actigraph GT1M minutes of moderate PA intensity and the IPAQ walking MET-minutes per week and IPAQ moderate MET-minutes per week, and 3) the PASE vigorous PA intensity with the Actigraph GT1M minutes of vigorous PA intensity and the IPAQ vigorous MET-minutes per week.

All statistical analyses were performed using the PASW Statistics 18 for Windows (IBM Corporation, Route, Somers, NY, USA).

Results

All 40 patients completed PASE at day 8, but at day 15 PASE were missing or inadequately filled out for seven patients. Calculation of the test-retest reliability was therefore based on the 33 patients with complete PASE questionnaires both at test and retest. Thirty-six patients had completed the Actigraph GT1M recording period and had readable files. Two patients returned the Actigraph GT1M unused, and data from two patients were not successfully downloaded. Six or more days of registration were considered to be sufficient. Three patients had less than six days of registration and were thus excluded from the analysis. In total, recordings from 33 patients were included to calculate correlation coefficients between the PASE and the Actigraph GT1M. The average days of registration were 7.0 (0.6). For the IPAQ, 15 patients had missing or incomplete questionnaires, leaving 25 patients to be included to calculate correlation coefficients between the PASE and the IPAQ. This was mainly due to inability to calculate the IPAQ score because the response alternative "don't know" was chosen.

Demographic and clinical characteristics of the patients are shown in Table 1. Based on the Actigraph GT1M measurements 67% fulfilled the recommendations of at least 30 minutes of accumulated MVPA per day, and 30% fulfilled the recommendations of at least 30 minutes of MVPA per day in blocks of minimum 10 minutes. At average the patients spent 45 (32) minutes per day on MVPA.

Test-retest reliability

Mean days between test and retest was nine days (SD 4.0), ranging from six to 25 days. Mean PASE score at test (n = 33) was 143 (SD 71) and at retest 125 (SD 56). The decline in the total PASE score from test to retest was significant (p = 0.02), but no significant differences was revealed for any of the sub scores/items. $ICC_{2,1}$ for the total PASE score was 0.77, SEM was 31 and MDC was 87 (Table 2). Test-retest values for the different sub

Table 1 Demographics and clinical characteristics of the 40 patients

Variables	
Age in years, mean (SD)	61.3 (10.0)
Men, n (%)	20 (50)
Body mass index, kg/m^2, mean (SD)	24.5 (3.6)
Years of education, n (%)	
7-9 years	11 (28.2)
10-12 years	13 (33.3)
> 12 years	15 (38.5)
Work status, n (%)	
At work	26 (66.7%)
Retired	10 (25.6%)
Sick-leave	3 (7.8%)
Bilateral hip-pain, n (%)	30 (75)
Minimal joint space in most painful hip, mm, mean (SD)	2.5 (0.9)
Pain duration, months, mean (SD)	49.8 (55.4)
Harris Hip Score, mean (SD)	80.7 (7.9)

Table 2 Test-retest reliability of the PASE

PASE score	Test, mean (SD)	Retest, mean (SD)	*d*, mean (95% CI)	$ICC_{2,1}$ (95% CI)	SEM	MDC
Total score (n = 33)	143 (71)	125 (56)	18 (-3,-32)	0.77 (0.56, 0.88)	31	87
Household/Work activities	114 (63)	84 (59)	15 (-1, 30)	0.69 (0.46, 0.84)	32	89
Leisure time PA	29 (24)	26 (19)	3 (-4, 10)	0.53 (0.24, 0.74)	15	40
Light PA intensity	13 (22)	10 (11)	3 (-4, 9)	0.46 (0.15, 0.69)	13	35
Moderate PA intensity	9 (10)	9 (13)	1 (-5, 6)	0.20 (-0.16, 0.51)	10	28
Vigorous PA intensity	4 (6)	5 (7)	-1 (-2, 1)	0.68 (0.44, 0.83	4	10

PA physical activity, *PASE* Physical Activity Scale for the Elderly, *d*: difference between test and retest, $ICC_{2,1}$ intraclass correlation coefficient, two-way random effects ANOVA, *SEM* standard error of measurement, *MDC* minimal detectable change

scores/items are also shown in Table 2. The Bland Altman plot for the total PASE score is shown in Figure 2. The lower LoA was -65 and the upper LoA was 100. One out of 33 values (3%) was outside the LoA.

Construct validity

The Spearman's rank correlation coefficient (ρ) between the PASE score and the Actigraph GT1M, and the PASE score and the IPAQ score is shown in Table 3. The correlation between the total PASE score and the Actigraph GT1M mean counts per minute was 0.30(p = 0.089). When comparing the total PASE score with the IPAQ total MET-minutes per week the correlation coefficient was 0.61 (p = 0.001).

For the different PA intensity items of the PASE we expected higher correlation coefficients with the respective categories of the Actigraph and the IPAQ. These comparisons are highlighted in Table 3. The correlation coefficients ranged from 0.10 to 0.35 between PASE and the Actigraph for the comparisons with the expected highest correlation, with only the correlation between the PASE item for moderate PA intensity and the

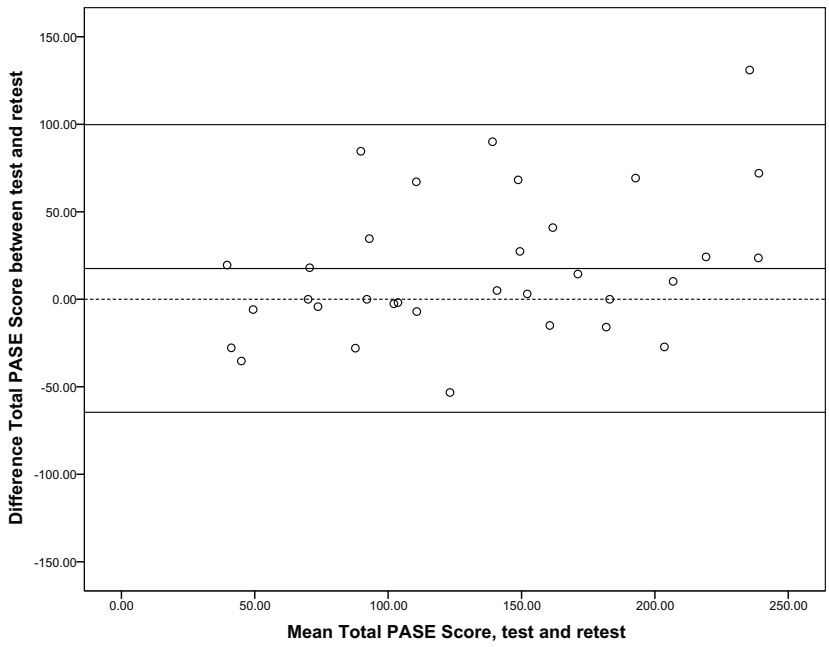

Figure 2 Bland-Altman plot for total PASE score. Intra-individual differences (n = 33) plotted against the difference between test and retest scores for the total PASE score. The central horizontal line represents the mean difference, while the flanking lines represent the 95% limits of agreement. The dotted line represents no difference between test and retest.

Table 3 Construct validity of the total PASE score, and the scores for light, moderate and vigorous PA intensity

	Mean (SD)	Total score	Score for Light PA intensity	Score for Moderate PA intensity	Score for Vigorous PA intensity
Actigraph GT1M					
Average counts per minute, counts/min	370 (199))	**0.30**			
Total minutes in interval counts 0-99, min	4015 (736)				
Total minutes in interval counts 100-2019, min	1989 (669)		**0.20**	0.21	0.21
Total minutes in interval counts 2020-5999, min	294 (194)		0.46**	**0.38***	0.11
Total minutes in interval counts > 6000, min	25 (50)		0.20	0.08	**0.29**
IPAQ					
Total IPAQ, MET-min/week	3476 (3609)	**0.61****			
Walking MET-min/week	2098 (3145)		**0.31**	**0.58****	0.05
Moderate Intensity, MET-min/week	707 (678)		0.20	**0.29**	0.16
Vigorous Intensity, MET-min/week	526 (869)		0.02	0.06	**0.75****

PA physical activity, *PASE* Physical Activity Scale for the Elderly, *IPAQ* International Physical Activity Questionnaire, The categories of PASE and Actigraph GT1M and of PASE and IPAQ with the highest expected correlation coefficients are highlighted by bold text

*: significant at 0.05-level; **: significant at 0.01-level

respective Actigraph category reaching statistical significance. The correlation coefficients ranged from 0.29 to 0.75 between PASE and IPAQ for the comparisons with the expected highest correlation. Of these, the correlation between the PASE score for moderate PA intensity and the IPAQ score for walking, and the PASE score for vigorous PA intensity and the IPAQ score for vigorous PA intensity reached statistical significance.

Discussion

This is the first study to address the test-retest reliability and the construct validity of the PASE in patients with hip OA, and the first study to evaluate the validity of the Norwegian version of the PASE. It is also one of relatively few studies evaluating the construct validity of a self-administered instrument for assessing PA by comparing it to an accelerometer, a method for direct measurement of PA, in patients with OA[13].

In our study we found that 67% of patients with hip OA fulfilled the recommendations of achieving at least 30 minutes of accumulated MVPA per day, but only 30% fulfilled the recommendations of achieving at least 30 minutes of MVPA per day in blocks of minimum 10 minutes. However, a larger percentage of the hip OA patients did fulfill the recommendations compared to the general Norwegian population. Only 20% of the general adult Norwegian population fulfill these recommendations, and a decline in the amount of PA was present after the age of 64 years. Mean counts per minute was 338, compared to 370 in our study[28]. The patients in our study were found to have high levels of PA when compared to other studies investigating levels of PA by accelerometers in OA patients[4,5,32]. Hirata et al.[32] found that women with hip OA were engaged in MVPA for 17 minutes per day, and only 14% met the recommendations of more than 30 minutes accumulated MVPA per day[32]. For patients with knee OA mean time spent on MVPA was 14-25 minutes per day[4,5] and 30% met the recommendations[4]. However, studies on PA levels in patients with knee OA may not be a valid comparison for the patients in our study. These previous studies[4,5,32] may have included patients with more progressive and severe OA than we did in our study, where patients with a Harris Hip Score below 60 points were excluded from participation. It is also important to stress that the hip OA patients in our study originally participated in a RCT where the importance of PA was emphasized through a patient education program, and this may have altered their PA levels. However, no changes in total PASE score was found for the 16 months follow-up of the RCT[20]. In addition, the possibility for selection bias is present, i.e. patients with a more positive attitude to PA might have been more likely to participate, and the education level was high. Thirty-nine percent of the patients in our study had more than 12 years of education, compared to 28% in the general Norwegian population (http://www.ssb.no/utniv). The levels of PA found in this study may therefore not be representative for the hip OA population in general.

PA has also been estimated in a representative sample of elderly Norwegians using PASE to assess physical activity[26]. The mean total PASE score was 127, quite consistent with the findings in our study on hip OA patients, where total PASE score was 143 and 125 at test and retest, respectively.

Measurement properties of an instrument are related to the population and context in which it is being used. In this study we evaluated the test-retest reliability of the PASE in patients with hip OA by calculating the $ICC_{2.1}$, and in addition estimating the standard error of measurement (SEM) and the minimal detectable change (MDC). There are no absolute consensus regarding limits for what should be considered an acceptable ICC value. When instruments for assessing PA is evaluated, Terwee et al.[13,15] and Forsèn et al.[33] have suggested, and used, 0.70 as a cut-off for acceptable test-retest reliability. Based on this the test-retest reliability for the total PASE score was considered to be acceptable, with an $ICC_{2.1}$ of 0.77. However, Terwee et al.[34] also suggested that the lower limit of the 96% CI of the ICC should exceed 0.60, and for the total PASE score the lower 95% CI was slightly lower than this, 0.56. The Norwegian version of PASE has previously been found to have acceptable reliability when tested in the general, elderly population, with an internal consistency of items (Cronbach's alpha) of 0.73, and test-retest reliability coefficient (Pearson's) of 0.93-0.99[26].

The SEM and MDC of the total PASE score were 31 and 87, respectively, indicating that 87 represents the smallest within-person change in score that can be interpreted as a real change, exceeding measurement error. However, a change exceeding the measurement error is not necessarily clinically relevant, which can be evaluated by estimating the Minimal Clinically Important Difference (MCID). It is advised that the MCID is estimated by using an anchor-based approach [35-37]. However, distribution-based approaches for estimating the MCID are also proposed, and the MCID has been found to equal approximately 0.5 SD at baseline[38] or approximately one SEM[39]. To be able to distinguish important changes from measurement error and to measure changes over time, the MCID should exceed the MDC[15], but by the smallest possible limit. The LoA indicates that if a subject completes a questionnaire twice, the second score could be as much as these limits smaller or larger than the first score, due to measurement error. Thus, the MCID should also lie outside the LoA[15]. Despite an acceptable test-retest ICC of the total PASE score, we consider the reliability to be moderate, due to large measurement error and wide LoA when compared to the mean total PASE score.

In our study, a significant decline in total PASE score of 18 points was present from test to retest, indicating a systematic error. We may therefore question whether the situation or the subjects actually were stable. When systematic error is present, this is often believed to occur due to a learning effect. However, this is not likely to be the case when the instrument of interest is a self-administered questionnaire. A more plausible explanation may be that wearing the Actigraph GT1M encouraged the patients to increase their activity levels, during the week the PASE referred to. According to Reiser and Schlenk[40] direct observations of PA by accelerometry may modify the pattern and level of PA among the participants, and may therefore bias the results.

Furthermore, this study evaluated the construct validity of the PASE by comparing it to an accelerometer, the Actigraph GT1M, and with another PA questionnaire, the IPAQ. As proposed by Terwee et al.[30] we tested predefined specific hypotheses including the expected direction and magnitude of correlations. In this study we found no significant correlation between the total PASE score and the Actigraph GT1M mean total counts per minute. The correlation coefficient was 0.30, in line with our a priori hypothesis. It was comparable to previous studies investigating the correlation between PASE and accelerometers in different populations, where correlations between 0.16-0.52 have been reported[17-19,41,42]. The correlation did not reach the cut-off for what we considered satisfactory correlation, above 0.50, as suggested by Terwee et al.[15]. Whereas self-reporting PA questionnaires is found to over-report levels of PA compared to accerelometers[43,44], Leenders et al.[45] found that accelerometers significantly underestimated PA related energy expenditure when compared to the doubly labelled water method. This may be due to some of its limitations. Accelerometers can of course only provide measurements for the particular time it is observed and recorded, cannot measure water exercises, and also fails to measure activities such as cycling and upper limb exercise correctly. Overestimation of total PA levels when using questionnaires and underestimation when using accelerometers, may to some degree explain the discrepancy between the two methods for measuring PA.

The correlation between total PASE score and IPAQ MET-minutes per week was moderate, with a correlation of 0.61, and barely within our a priori hypothesize of correlation between 0.6 and 0.9. Both PASE and IPAQ are self-administered with a seven day recall period, but household- and work activities is included in the PASE and weighed quite highly, whereas the IPAQ mainly captures leisure-time PA. This may, at least partly, explain the discrepancy between the two questionnaires. Both questionnaires were originally developed for use in a general population (generic), with PASE being specifically designed for an elderly population.

The PASE is not designed to be used to measure and report different PA intensity levels separately. One might therefore argue that acceptable test-retest reliability for the overall score is what is important. However, assessment of intensity seems valuable when investigating the effect of exercise and PA, especially for evaluating the dose-response relationship and to establish recommendations for patients with OA regarding amount and intensity. We therefore wanted to evaluate these specific items, to evaluate whether a PA questionnaire is able to provide reliable and valid data for PA intensity. The $ICC_{2.1}$ for the sub-scores for household/work-related PA and for leisure-time PA was 0.69 and 0.53, respectively, and the $ICC_{2.1}$ for the items for light, moderate and vigorous PA intensity was 0.46, 0.20 and 0.68, respectively. None of the ICC's for the sub-scores or the single item scores exceeded 0.7, which we interpreted as a cut-off for acceptable reliability, and the 95% CI were wide for all the sub-scores and items. The SEM and the MDC were also large compared to the mean values of the sub-scores and items, indicating moderate to low reliability.

Our a priori hypothesis; that the respective intensity categories of the PASE would correlate strongest with the respective intensity categories of the Actigraph GT1M, was confirmed for moderate PA intensity and vigorous PA intensity, but not for light PA intensity. However, all correlation coefficients were below 0.46. This indicated that the intensity items of the PASE were not able to distinguish between light, moderate and vigorous PA intensity, and we therefore consider the PASE not to be valid or reliable for assessing PA intensity. The item for moderate PA intensity of PASE correlated stronger with the IPAQ category for walking than the IPAQ category for moderate PA intensity. This may be due to the fact that the IPAQ includes a specific item for assessing walking activities, whereas walking activities are included in the items for light, moderate and vigorous PA intensity in the Norwegian version of PASE. Walking is a widespread leisure time activity in Norway, and is likely to be scored in the item for moderate PA intensity of the PASE, giving a higher correlation with the IPAQ walking compared to the IPAQ moderate PA intensity.

This study has some limitations. Both analysis of test-retest reliability and construct validity by comparing PASE to the Actigraph GT1M were based on data obtained from 33 patients. After referring a statistician, and based on that other studies have used similar sample sizes[19,33], we decided to include 40 patients in this study. According to the statistician a sample size between 30 and 40 is usually sufficient when evaluating outcome measurements that uses a continuous scale. According to Terwee et al.[15] sample size in reliability and/or validity studies evaluating PA assessment tools should exceed 50. A recently developed scoring system for rating methodological quality of measurement properties suggests that a sample size of 100 should be considered excellent, 50 as good, 30 as fair and under 30 as poor[46]. Correlation between PASE and IPAQ was only based on data from 25 patients. The Norwegian version of IPAQ has been validated for the Norwegian population, but has included an item "don't know" as an option for duration of activity which challenge the interpretation and the score calculations.

The use of Actigraph GT1M and the IPAQ to evaluate construct validity have some weaknesses. The doubly labeled water method is often considered to be the gold standard for measuring PA[15], but is seldom used to evaluate validity of PA questionnaires, as it is expensive, time-consuming and relies on access to both technical expertise and equipment. Only two studies have validated the PASE by comparing it to doubly labelled water, and found correlation coefficients of 0.28[47] and 0.68[48]. However, the doubly labelled water method is affected by the basal metabolic rate, and it cannot capture frequency, duration and intensity of activity. Accelerometers may therefore represent a more appropriate comparator because it can provide information on amount, pattern and intensity of PA, and therefore seem to measure the same construct as most PA questionnaires[15]. There is evidence for reasonable correlation between waist-worn accelerometers and the doubly labelled water method in adults, with correlations ranging from 0.30-0.83[49]. IPAQ was also included as a comparator because it is a widely used PA questionnaire, but like other questionnaires it is vulnerable to recall and reporting bias. Previous studies comparing IPAQ and accelerometers/activity monitors have reported correlation coefficients between 0.29 to 0.35 [50-52]. However, Ainsworth[53] states that questionnaires may be suitable for assessing PA for most patients. More sophisticated methods, like accelerometers, provide more precise measurements, but are less practical for use in clinical settings. Kayes and McPherson[54] emphasize that PA questionnaires and accelerometers both have weaknesses, but that both methods are likely to assess important aspects of the PA construct. Use of both tools may therefore be appropriate to capture all aspects of PA.

Conclusions

The test-retest reliability of the total PASE score in patients with hip OA was found to be moderate, based on an acceptable $ICC_{2.1}$, but the large SEM, SDC and LoA indicate large measurement errors. The construct validity of the total PASE score was found to be poor when compared to the Actigraph GT1M accelerometer.

These findings suggest that PASE is not sufficient for assessing PA levels and intensity in patients with hip OA. Accelerometers provide a more precise tool of assessing amount and intensity of PA, and should preferably be included if feasible in studies where these dimensions are considered important.

Acknowledgements
The authors thank all participants, and the Department of Sports Medicine, Norwegian School of Sport Sciences, for providing the accelerometers used in this study.
The study was financially supported by the former Ullevaal University Hospital, Oslo, Norway and the Norwegian Foundation for Health and Rehabilitation, through the Norwegian Rheumatism Association.

Author details
[1]Norwegian Research Center for Active Rehabilitation, Department of Orthopedics, Oslo University Hospital, Ullevaal and Hjelp 24 NIMI, Oslo, Norway. [2]Department of Sports Medicine, Norwegian School of Sport Sciences, Oslo, Norway.

Authors' contributions
All authors participated in the design of the study, contributed in drafting the article, and read and approved the final manuscript. IS carried out the patient inclusion, handled the administration of questionnaires and accelerometers, and carried out the statistical analysis. EK carried out the processing of the Actigraph GT1M data.

Competing interests
The authors declare that they have no competing interests.

Received: 16 September 2011 Accepted: 21 February 2012
Published: 21 February 2012

References

1. Norwegian Directorate of Health: *The action plan on physical activity 2005-2009: Working together for physical activity* Oslo, Norway: Norwegian Directorate of Health; 2005.
2. Haskell WL, Lee IM, Pate RR, Powell KE, Blair SN, Franklin BA, *et al*: **Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association.** *Circulation* 2007, **116**:1081-1093.
3. World Health Organization: *Global recommendations on physical activity for health* Geneva, Switzerland: World Health Organization; 2010.
4. Farr JN, Going SB, Lohman TG, Rankin L, Kasle S, Cornett M, *et al*: **Physical activity levels in patients with early knee osteoarthritis measured by accelerometry.** *Arthritis Rheum* 2008, **59**:1229-1236.
5. Song J, Semanik P, Sharma L, Chang RW, Hochberg MC, Mysiw WJ, *et al*: **Assessing physical activity in persons with knee osteoarthritis using accelerometers: data from the osteoarthritis initiative.** *Arthritis Care Res (Hoboken)* 2010, **62**:1724-1732.
6. Dunlop DD, Song J, Semanik PA, Sharma L, Chang RW: **Physical activity levels and functional performance in the osteoarthritis initiative: a graded relationship.** *Arthritis Rheum* 2011, **63**:127-136.
7. Dieppe PA, Lohmander LS: **Pathogenesis and management of pain in osteoarthritis.** *Lancet* 2005, **365**:965-973.
8. Lohmander LS, Roos EM: **Clinical update: treating osteoarthritis.** *Lancet* 2007, **370**:2082-2084.
9. Zhang W, Nuki G, Moskowitz RW, Abramson S, Altman RD, Arden NK, *et al*: **OARSI recommendations for the management of hip and knee osteoarthritis: part III: Changes in evidence following systematic cumulative update of research published through January 2009.** *Osteoarthr Cartil* 2010, **18**:476-499.
10. Fransen M, McConnell S, Hernandez-Molina G, Reichenbach S: **Exercise for osteoarthritis of the hip.** *Cochrane Database Syst Rev* 2009, CD007912.
11. Fransen M, McConnell S: **Exercise for osteoarthritis of the knee.** *Cochrane Database Syst Rev* 2008, CD004376.
12. Hernandez-Molina G, Reichenbach S, Zhang B, Lavalley M, Felson DT: **Effect of therapeutic exercise for hip osteoarthritis pain: results of a meta-analysis.** *Arthritis Rheum* 2008, **59**:1221-1228.
13. Terwee CB, Bouwmeester W, van Elsland SL, de Vet HC, Dekker J: **Instruments to assess physical activity in patients with osteoarthritis of the hip or knee: a systematic review of measurement properties.** *Osteoarthr Cartil* 2011, **19**:620-633.
14. Freedson P, Pober D, Janz KF: **Calibration of accelerometer output for children.** *Med Sci Sports Exerc* 2005, **37**:S523-S530.
15. Terwee CB, Mokkink LB, van Poppel MN, Chinapaw MJ, van MW, de Vet HC: **Qualitative attributes and measurement properties of physical activity questionnaires: a checklist.** *Sports Med* 2010, **40**:525-537.
16. Martin KA, Rejeski WJ, Miller ME, James MK, Ettinger WH Jr, Messier SP: **Validation of the PASE in older adults with knee pain and physical disability.** *Med Sci Sports Exerc* 1999, **31**:627-633.
17. Hagiwara A, Ito N, Sawai K, Kazuma K: **Validity and reliability of the Physical Activity Scale for the Elderly (PASE) in Japanese elderly people.** *Geriatr Gerontol Int* 2008, **8**:143-151.
18. Dinger MK, Oman RF, Taylor EL, Vesely SK, Able J: **Stability and convergent validity of the Physical Activity Scale for the Elderly (PASE).** *J Sports Med Phys Fitness* 2004, **44**:186-192.
19. Washburn RA, Ficker JL: **Physical Activity Scale for the Elderly (PASE): the relationship with activity measured by a portable accelerometer.** *J Sports Med Phys Fitness* 1999, **39**:336-340.
20. Fernandes L, Storheim K, Sandvik L, Nordsletten L, Risberg MA: **Efficacy of patient education and supervised exercise vs patient education alone in patients with hip osteoarthritis: a single blind randomized clinical trial.** *Osteoarthr Cartil* 2010, **18**:1237-1243.
21. Harris WH: **Traumatic arthritis of the hip after dislocation and acetabular fractures: treatment by mold arthroplasty. An end-result study using a new method of result evaluation.** *J Bone Joint Surg Am* 1969, **51**:737-755.
22. Danielsson LG: **Incidence and prognosis of coxarthrosis.** *Clin Orthop Relat Res* 1964, **1993**:13-18.
23. Washburn RA, Smith KW, Jette AM, Janney CA: **The Physical Activity Scale for the Elderly (PASE): development and evaluation.** *J Clin Epidemiol* 1993, **46**:153-162.
24. Petrella RJ, Bartha C: **Home based exercise therapy for older patients with knee osteoarthritis: a randomized clinical trial.** *J Rheumatol* 2000, **27**:2215-2221.
25. Dunlop DD, Semanik P, Song J, Sharma L, Nevitt M, Jackson R, *et al*: **Moving to maintain function in knee osteoarthritis: evidence from the osteoarthritis initiative.** *Arch Phys Med Rehabil* 2010, **91**:714-721.
26. Loland NW: **Reliability of the physical activity scale for the elderly (PASE).** *Eur J Sports Sci* 2002, **2**:1-12.
27. Troiano RP: **Large-scale applications of accelerometers: new frontiers and new questions.** *Med Sci Sports Exerc* 2007, **39**:1501.
28. Hansen BH, Kolle E, Dyrstad SM: *Holme I* Anderssen SA: Accelerometer-Determined Physical Activity in Adults and Older People. Med Sci Sports Exerc; 2011.
29. Craig CL, Marshall AL, Sjostrom M, Bauman AE, Booth ML, Ainsworth BE, *et al*: **International physical activity questionnaire: 12-country reliability and validity.** *Med Sci Sports Exerc* 2003, **35**:1381-1395.
30. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, *et al*: **Quality criteria were proposed for measurement properties of health status questionnaires.** *J Clin Epidemiol* 2007, **60**:34-42.
31. Bland JM, Altman DG: **Measuring agreement in method comparison studies.** *Stat Methods Med Res* 1999, **8**:135-160.
32. Hirata S, Ono R, Yamada M, Takikawa S, Nishiyama T, Hasuda K, *et al*: **Ambulatory physical activity, disease severity, and employment status in adult women with osteoarthritis of the hip.** *J Rheumatol* 2006, **33**:939-945.
33. Forsen L, Loland NW, Vuillemin A, Chinapaw MJ, van Poppel MN, Mokkink LB, *et al*: **Self-administered physical activity questionnaires for the elderly: a systematic review of measurement properties.** *Sports Med* 2010, **40**:601-623.
34. Terwee CB, Mokkink LB, Steultjens MP, Dekker J: **Performance-based methods for measuring the physical function of patients with osteoarthritis of the hip or knee: a systematic review of measurement properties.** *Rheumatology (Oxford)* 2006, **45**:890-902.
35. Turner D, Schunemann HJ, Griffith LE, Beaton DE, Griffiths AM, Critch JN, *et al*: **The minimal detectable change cannot reliably replace the minimal important difference.** *J Clin Epidemiol* 2010, **63**:28-36.

36. de Vet HC, Terwee CB, Ostelo RW, Beckerman H, Knol DL, Bouter LM: **Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change.** *Health Qual Life Outcomes* 2006, **4**:54.
37. de Vet HC, Terwee CB: **The minimal detectable change should not replace the minimal important difference.** *J Clin Epidemiol* 2010, **63**:804-805.
38. Norman GR, Sloan JA, Wyrwich KW: **Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation.** *Med Care* 2003, **41**:582-592.
39. Wyrwich KW: **Minimal important difference thresholds and the standard error of measurement: is there a connection?** *J Biopharm Stat* 2004, **14**:97-110.
40. Reiser LM, Schlenk EA: **Clinical use of physical activity measures.** *J Am Acad Nurse Pract* 2009, **21**:87-94.
41. Liu RD, Buffart LM, Kersten MJ, Spiering M, Brug J, van MW, *et al*: **Psychometric properties of two physical activity questionnaires, the AQuAA and the PASE, in cancer patients.** *BMC Med Res Methodol* 2011, **11**:30.
42. Harada ND, Chiu V, King AC, Stewart AL: **An evaluation of three self-report physical activity instruments for older adults.** *Med Sci Sports Exerc* 2001, **33**:962-970.
43. Emaus A, Degerstrom J, Wilsgaard T, Hansen BH, Dieli-Conwright CM, Furberg AS, *et al*: **Does a variation in self-reported physical activity reflect variation in objectively measured physical activity, resting heart rate, and physical fitness? Results from the Tromso study.** *Scand J Public Health* 2010, **38**:105-118.
44. Leenders NYJM, Sherman WM, Nagaraja HN: **Comparisons of four methods of estimating physical activity in adult women.** *Med Sci Sports Exerc* 2000, **32**:1320-1326.
45. Leenders NY, Sherman WM, Nagaraja HN, Kien CL: **Evaluation of methods to assess physical activity in free-living conditions.** *Med Sci Sports Exerc* 2001, **33**:1233-1240.
46. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC: **Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist.** *Qual Life Res* 2011.
47. Bonnefoy M, Normand S, Pachiaudi C, Lacour JR, Laville M, Kostka T: **Simultaneous validation of ten physical activity questionnaires in older men: a doubly labeled water study.** *J Am Geriatr Soc* 2001, **49**:28-35.
48. Schuit AJ, Schouten EG, Westerterp KR, Saris WH: **Validity of the Physical Activity Scale for the Elderly (PASE): according to energy expenditure assessed by the doubly labeled water method.** *J Clin Epidemiol* 1997, **50**:541-546.
49. Plasqui G, Westerterp KR: **Physical activity assessment with accelerometers: an evaluation against doubly labeled water.** *Obesity (Silver Spring)* 2007, **15**:2371-2379.
50. Kurtze N, Rangul V, Hustvedt BE: **Reliability and validity of the international physical activity questionnaire in the Nord-Trondelag health study (HUNT) population of men.** *BMC Med Res Methodol* 2008, **8**:63.
51. Hagstromer M, Ainsworth BE, Oja P, Sjostrom M: **Comparison of a subjective and an objective measure of physical activity in a population sample.** *J Phys Act Health* 2010, **7**:541-550.
52. Macfarlane D, Chan A, Cerin E: **Examining the validity and reliability of the Chinese version of the International Physical Activity Questionnaire, long form (IPAQ-LC).** *Public Health Nutr* 2010, **13**:1-8.
53. Ainsworth BE: **How do I measure physical activity in my patients? Questionnaires and objective methods.** *Br J Sports Med* 2009, **43**:6-9.
54. Kayes NM, McPherson KM: **Measuring what matters: does 'objectivity' mean good science?** *Disabil Rehabil* 2010, **32**:1011-1019.

Pre-publication history
The pre-publication history for this paper can be accessed here:
http://www.biomedcentral.com/1471-2474/13/26/prepub

doi:10.1186/1471-2474-13-26
Cite this article as: Svege *et al.*: Reliability and validity of the Physical Activity Scale for the Elderly (PASE) in patients with hip osteoarthritis. *BMC Musculoskeletal Disorders* 2012 **13**:26.