



Ulla Tuominen

## **The effect of waiting time on health outcomes and costs of total joint replacement in hip and knee patients**

A prospective randomised controlled trial





# Studies in social security and health 126

KELA, RESEARCH DEPARTMENT | HELSINKI 2013

*Ulla Tuominen*

## ***The effect of waiting time on health outcomes and costs of total joint replacement in hip and knee patients***

*A prospective randomized controlled trial*

*To Tiia and Vili*

**Author**

Ulla Tuominen, M.Sc, Researcher (Health economics)  
The Social Insurance Institution of Finland, Research Department  
firstname.lastname@kela.fi

The publications in this series have undergone a formal referee process.  
© Ulla Tuominen and Kela, Research Department

Layout: Pekka Loiri

ISBN 978-951-669-911-3 (print)  
ISBN 978-951-669-912-0 (pdf)

ISSN-L 1238-5050  
ISSN 1238-5050 (print)  
ISSN 2323-7724 (online)

Printed by Juvenes Print  
Tampere 2013

## Abstract

Tuominen U. **The effect of waiting time on health outcomes and costs of total joint replacement in hip and knee patients. A prospective randomized controlled trial.** Helsinki: The Social Insurance Institution of Finland, Studies in social security and health 126, 2013. 114 pp. ISBN 978-951-669-911-3 (print), ISBN 978-951-669-912-0 (pdf).

Osteoarthritis is the most common cause of musculoskeletal disability and pain. Long waiting times for elective surgery have been a concern in many countries. The aim of this study was to evaluate the effect of waiting time on health outcomes and costs of total joint replacement in hip and knee patients. When placed on the waiting list, patients were randomly allocated to a short waiting time (waiting time  $\leq 3$  months) or a non-fixed waiting time (waiting time not fixed in advance, patients followed the hospitals' routine practice) groups. The outcomes were measured by the 15D questionnaire, modified Harris Hip and Knee Scores and Visual Analogue Scale. The use and costs of medication, health care and social services were measured. The measurement points were, when placed on the waiting list upon admission, three months and twelve months postoperatively. Quality-adjusted life years (QALYs) were calculated to evaluate the effect of waiting time on the cost-utility of total joint replacement. The waiting time difference did not result in a significant difference in health outcomes. After surgery, the mean scores of outcomes improved in both randomized groups by a statistically significant margin. The point estimates from cost-utility analyses showed that hip patients in the short waiting time group compared to the non-fixed group gained more QALYs at lower costs, thus suggesting a strong dominance for the shorter waiting time. In knee patients, the situation was the opposite. There were refusals and dropouts during the follow-up, which might introduce bias and uncertainty into the results. There does not seem to be a significant difference between the cost-utility of short and longer waiting times for total joint replacement, at least given the waiting time difference between our study groups. However, due to quite a lot of uncertainty in the results, one has to be cautious about generalizing the findings.

**Keywords:** total hip replacement, total knee replacement, osteoarthritis, joint replacement operations, waiting time, health-related quality of life, health outcomes, costs, cost-utility analysis

## Tiivistelmä

Tuominen U. **Jonotusajan vaikutukset lonkan ja polven tekonivelleikkaushoidon lopputulokseen ja kustannuksiin. Satunnaistettu kontrolloitu tutkimus.** Helsinki: Kela, Sosiaali- ja terveysturvan tutkimuksia 126, 2013. 114 s. ISBN 978-951-669-911-3 (nid.), ISBN 978-951-669-912-0 (pdf).

Nivelrikko on yksi yleisimmistä liikuntakyvyttömyyttä ja kipua aiheuttavista tuki- ja liikuntaelinsairauksista. Tekonivelkirurgiaan liittyvät pitkät jonotusajat ovat olleet länsimaissa vuosikymmenien ajan huolen aiheena. Tämän tutkimuksen tarkoitus oli selvittää jonotusajan pituuden vaikutusta polven ja lonkan primaariin tekonivelleikkaukseen liittyviin terveysvaikutuksiin sekä kustannuksiin ennen leikkausta ja sen jälkeen. Tekonivelleikkausjonoon asetetut polvi- ja lonkkapotilaat satunnaistettiin kahteen ryhmään: lyhyen jonotusajan ryhmään ja normaalin sairaalakäytännön mukaisesti jonottavaan ryhmään. Terveysvaikutuksia mitattiin 15D-, Harris Hip Score -, Knee Score - ja VAS-mittareilla. Tietoja kerättiin lääkkeiden sekä sosiaali- ja terveystalvelujen käytöstä ja kustannuksista. Vaikuttavuutta arvioitiin kustannusutiliteettianalyysillä. Mittaukset tehtiin jonoon asettamisajankohtana, sairaalaan saapumisajankohtana sekä kolmen ja kahdentoista kuukauden kuluttua leikkauksesta. Tulokset perustuvat satunnaistetun aineiston hoitoaikaan mukaiseen analyysiin, ja tulokset on testattu herkkyysanalyysillä. Jonotusajan pituudella ei ollut satunnaistettujen vertailuryhmien välillä merkittävää vaikutusta elämänlaatuun, kipuun eikä liikuntakykyyn. Leikkauksen jälkeiset vaikuttavuusarvot olivat tilastollisesti merkitsevästi parempia kuin arvot sairaalaan saavuttaessa. Kustannusvaikuttavuusanalyysin mukaan nopean leikkausryhmän lonkkapotilaat saavuttivat enemmän laatupainotettuja lisäelinvuosia pienemmillä kustannuksilla kuin normaalin käytännön mukaan jonottaneet potilaat. Polvipotilailla tilanne oli päinvastainen. Tutkimuksen seuranta-aikana tapahtuneen kadon vuoksi analyysin tuloksiin liittyy jonkin verran epävarmuutta, joten yleistettävää johtopäätöstä jonotusajan vaikutuksesta tekonivelleikkausten kustannusutiliteettiin ei voida varmistaa.

**Avainsanat:** lonkan tekonivelleikkaus, polven tekonivelleikkaus, nivelrikko, tekonivelleikkaukset, jonotusaika, terveyteen liittyvä elämänlaatu, terveysvaikutukset, kustannukset, kustannusutiliteettianalyysi

## Sammandrag

Tuominen U. **Väntetidens inverkan på resultatet och kostnader i samband med protesoperationer i höft- och knäled. En randomiserad, kontrollerad undersökning.** Helsingfors: FPA, Social trygghet och hälsa: Undersökningar 126, 2013. 114 s. ISBN 978-951-669-911-3 (hft.), 978-951-669-912-0 (pdf).

Artros är en av de vanligaste sjukdomarna i rörelseorganen som orsakar rörelseoförmåga och smärta. De långa väntetiderna i samband med ledproteskirurgi har i decennier varit en anledning till bekymmer i västvärlden. Syftet med denna undersökning är att undersöka hur väntetid inverkar på hälsoeffekterna i samband med primär protesoperation i höft- och knäled samt på kostnaderna före och efter operationen. Knä- och höftpatienter som hade placerats i kö för en ledprotesoperation randomiserades i två grupper: en grupp med kort väntetid och en grupp som köade enligt normal sjukhuspraxis. Hälsoeffekterna mättes med hjälp av 15D, Harris Hip Score, Knee Score och VAS. Uppgifter om användningen av och kostnaderna för läkemedel samt social- och hälsotjänster samlades in. Effektiviteten utvärderades med hjälp av kostnadsutilitetanalys. Mätningar gjordes när personen placerades i kö, vid ankomsten till sjukhuset samt 3 och 12 månader efter operationen. Resultaten grundar sig på intention to treat-analys av ett randomiserat material och har testats med sensitivitetsanalyser. Väntetidens längd hade ingen signifikant inverkan på livskvalitet, smärta eller rörelseförmåga mellan de randomiserade jämförelsegrupperna. Efter operationen var effektivitetsvärdena statistiskt signifikant bättre än vid ankomsten till sjukhuset. Enligt kostnadseffektivitetsanalysen uppnådde den grupp av höftpatienter som hade kort väntetid fler kvalitetsjusterade levnadsår med lägre kostnader än de patienter som köade på normalt vis. I fråga om knäpatienterna var situationen den motsatta. På grund av bortfall under observationsperioden var analysens resultat i viss mån osäkra. Således går det inte att göra någon generaliserbar slutsats gällande väntetidens inverkan på kostnadsutiliteten av ledprotesoperationerna.

**Nyckelord:** protesoperationer i höftled, protesoperationer i knäled, artros, protesoperationer, väntetid, livskvalitet i fråga om hälsa, hälsoeffekter, kostnader, kostnadsutilitetanalys

## ACKNOWLEDGEMENTS

The year 2002 was a turning point in my life when I got involved in the research world. During that year, I got a position as a research assistant at Stakes when a large prospective multicentre research project about “Effectiveness in specialized health care” started. My first contact person in this scientific world was Professor Marja Blom, and with her encouragement and as a continuum after this preliminary work, I began to write my doctoral thesis from the very data I had collected for several years. Professor Marja Blom was the key person in the study and my first supervisor. I am deeply grateful to her for giving me the possibility to work with her and for fruitful conversations, feedback and all that endless support during these years. Marja’s expertise as a health economist and her research experience were of great advantage in learning, understanding, questioning and optimizing the main points of this thesis. Thank you, Marja. Without you, this could not be possible. Professor (emeritus) Harri Sintonen has been my other supervisor. He is one of the first health economists in Finland, and I had the honour to work with him. I would like to thank you, Harri, for your invaluable feedback, extensive knowledge in health economics and confidence in my abilities during this process.

I am deeply grateful to the whole Advisor “Polkka” research team and co-authors: my closest colleague, travelling companion and a good friend at Stakes, Johanna Hirvonen, PhD, Professor Seppo Seitso, Professor Matti Lehto, Professor Pekka Paavolainen, Professor Kalevi Hietaniemi and Professor Pekka Rissanen. I also thank the surgical staff of the HUS HUCH Jorvi Hospital and Surgical Hospital, Coxa Hospital for Joint Replacement, and Orton Orthopaedic Hospital. They put this study into practice by recruiting and advising the patients about this research. I also owe my deepest gratitude to all those patients who completed our questionnaires and made this study concrete.

I am deeply grateful to Research Professor Markku Pekurinen, the director of CHES (my workplace at the time I started the thesis), for support, the possibility and his understanding attitude towards my research.

I want to express my sincere gratitude to the official reviewers of this thesis, Associate Professor Arto Ohinmaa and Associate Professor Ville Remes, for their valuable and constructive comments on my thesis. Professor Hannu Valtonen is warmly



acknowledged for agreeing to be my opponent in the public defence of this thesis.

I owe my thanks to Chief of Health Research Hennamari Mikkola and Head of Drug Research Jaana Martikainen for encouragement, interest and mental support during my research years at Kela. I would also like to thank Researcher Hanna Koskinen, my colleague and friend, for sharing numerous moments of conversation and support, and my other wonderful colleagues on Kela's drug research team.

I am grateful to have my thesis published in the Studies in Social Security and Health series. I want to express my gratitude and warm thanks to Research Professor Olli Kangas, Head of Research Publications Tarja Hyvärinen and Publishing Editor Maini Tulokas.

I give my deepest gratitude to the Academy of Finland (no. 51871), HUS HUCH Jorvi Hospital, Coxa Hospital for Joint Replacement, Medical Research of Tampere University Hospital, HUS HUCH Surgical Hospital, Orton Orthopaedic Hospital, and the Yrjö Jahnsson Foundation for awarding funding for this thesis.

I want to thank my mother-in-law Maj, and Kai, and my closest friends at Anttilanmäki for their support and for keeping me in touch with joy and the real world.

Finally, with love, I want to thank Jukka, my beloved husband, who has supported me in all ways, above all with his enthusiasm and wonderful cooking skills (which allowed me to stay out of the kitchen), and our children Tiia and Vili and grandchild Jimi for giving me so much love, happiness and purpose to my life.

*Helsinki, February 2013*

*Ulla Tuominen*

## CONTENTS

List of original publications.....	10
List of Figures.....	11
List of Tables.....	11
Abbreviations.....	12
<b>1 INTRODUCTION</b> .....	<b>13</b>
<b>2 OSTEOARTHRITIS</b> .....	<b>15</b>
2.1 Prevalence of osteoarthritis (OA).....	15
2.2 Diagnosis of OA.....	16
2.3 Treatments of OA.....	16
2.3.1 Medication .....	16
2.3.2 Physiotherapy.....	17
2.3.3 Surgery.....	18
2.4 Co-morbidity and OA.....	20
<b>3 LITERATURE REVIEW OF WAITING TIME AND COST-EFFECTIVENESS OF TOTAL JOINT REPLACEMENT</b> .....	<b>21</b>
3.1 Results from the earlier systematic literature reviews .....	21
3.2 Literature review.....	22
3.2.1 Results of the present literature review .....	23
<b>4 WAITING TIME AND ACTIONS TO SHORTEN IT</b> .....	<b>27</b>
4.1 Access to care .....	27
4.2 Waiting time and health care reforms.....	28
4.2.1 Health care reforms in England and Sweden.....	29
4.2.2 Health care reform in Finland.....	30
<b>5 ECONOMIC EVALUATION</b> .....	<b>31</b>
5.1 Economic evaluation in health care.....	31
5.1.1 Quality-adjusted life year.....	31
5.1.2 Cost-effectiveness analysis .....	32
5.1.3 Cost-utility analysis .....	33
5.1.4 Decision rules of CEA and CUA .....	33
<b>6 AIMS OF THE STUDY</b> .....	<b>35</b>
<b>7 PATIENTS AND METHODS</b> .....	<b>36</b>
7.1 Study design.....	36
7.2 Feasibility of the study.....	36
7.3 Study population .....	36
7.4 Randomization of the study population.....	37
7.5 Outcome measures .....	38
7.5.1 Generic 15D instrument .....	38
7.5.2 Disease-specific outcome measures.....	40
7.5.3 Direct rating scale instrument.....	41

7.6	Direct cost measures.....	41
7.6.1	Social and health service use .....	41
7.6.2	Medication use .....	42
7.7	Statistical analysis.....	42
7.8	Ethical considerations.....	44
<b>8</b>	<b>RESULTS.....</b>	<b>45</b>
8.1	Effect of co-morbidity on HRQoL in patients waiting for total joint replacement (I).....	45
8.2	Characteristics of randomized groups in hip and knee patients and those who were lost to follow-up during the study (II, III).....	46
8.2.1	Waiting times in the randomized groups (II, III).....	47
8.2.2	Effect of waiting time on health outcomes (II, III).....	48
8.2.3	Costs of disease-specific medication (II, III).....	50
8.3	The cost-utility (IV) .....	51
8.3.1	Sensitivity analysis.....	53
<b>9</b>	<b>DISCUSSION.....</b>	<b>56</b>
9.1	HRQoL in patients waiting for Total Joint Replacement.....	56
9.2	Baseline co-morbidity and HRQoL.....	57
9.3	Randomization and waiting time .....	57
9.4	Health outcome measures and costs of osteoarthritis.....	57
9.5	Cost-utility of Total Joint Replacement.....	59
9.6	Limitations of the study.....	59
9.7	Reliability and validity of the study .....	60
<b>10</b>	<b>CONCLUSIONS AND SUGGESTIONS FOR FURTHER STUDIES.....</b>	<b>61</b>
	<b>REFERENCES .....</b>	<b>62</b>
	<b>APPENDICES .....</b>	<b>70</b>
	<b>ORIGINAL PUBLICATIONS .....</b>	<b>75</b>

**LIST OF ORIGINAL PUBLICATIONS**

- I Tuominen U, Blom M, Hirvonen J, Hietaniemi K, Lehto MUK, Paavolainen P, Seitsalo S, Sintonen H, Rissanen P.** Effects of co-morbidities to health-related quality of life in patients awaiting major joint replacement. *Health and Quality of Life Outcomes* 2007; 5: 16; doi:10.1186/1477-7525-5-16.
- II Tuominen U, Sintonen H, Hirvonen J, Hietaniemi K, Lehto MUK, Paavolainen P, Seitsalo S, Blom M.** The effect of waiting time on health outcomes and use and costs of medication in hip replacement patients. A randomised clinical trial. *Osteoarthritis and Cartilage* 2009; 17: 1144–1150.
- III Tuominen U, Sintonen H, Hirvonen J, Hietaniemi K, Lehto MUK, Paavolainen P, Seitsalo S, Blom M.** Is longer waiting time for total knee replacement associated with health outcomes and medication costs? A randomized clinical trial. *Value in Health* 2010; 13 (8): 998–1004.
- IV Tuominen U, Sintonen H, Hirvonen J, Hietaniemi K, Lehto MUK, Seitsalo S, Blom M.** The cost-utility of waiting time in total joint replacements. RCT. 2012 (accepted for publication September 2012). *International Journal of Technology Assessment in Health Care* 2013; 29 (1): 27–34.

The publications are referred to in the text by their Roman numerals and reprinted by permission of the copyright holders.

## LIST OF FIGURES

- Figure 1.** The number of total joint replacements in Finland during years 1991–2010.
- Figure 2.** Flow diagram for study selection.
- Figure 3.** Optimal waiting time.
- Figure 4.** QALYs gained from Intervention I in comparison to Intervention II.
- Figure 5.** Flow diagram of the progress through the phases of the randomized trial.
- Figure 6.** The 15D profiles (mean level values) of patients with (Co-mo) and without co-morbidity (No Co-mo), at baseline.
- Figure 7.** The 15D profile of hip and knee patients placed on the waiting list for total joint replacement.
- Figure 8.** The distribution of waiting times in the randomized groups.
- Figure 9.** The mean 15D scores in the randomized hip and knee patients during the observation time.
- Figure 10.** The mean Modified HHS during the whole observation period.
- Figure 11.** The mean Modified KSS (pain and function) during the whole observation period.
- Figure 12.** The mean total costs of disease-specific medication in the randomized groups from baseline to one year postoperatively.
- Figure 13.** Cost-effectiveness plane in hip replacement.
- Figure 14.** Cost-effectiveness plane in knee replacement.
- Figure 15.** Cost-effectiveness acceptability curves for the short waiting time in hip and knee replacement.

## LIST OF TABLES

- Table 1.** Prevalence (%) of hip and knee osteoarthritis according to the examining physician's diagnosis in Finland 2000.
- Table 2.** Costs and consumption of M01A and one over-the-counter medicine (N02BE01) in 2004–2006.
- Table 3.** Summary of the literature review.
- Table 4.** Measurement of the costs and consequences in economic evaluation.
- Table 5.** The cost-effectiveness decision table.
- Table 6.** Baseline characteristics in patients waiting for total joint replacement and those lost to follow-up.
- Table 7.** Mean 15D scores in hip and knee patients in the randomized groups.
- Table 8.** Mean costs of health care utilization; differences between the randomized groups in hip and knee patients.
- Table 9.** The results of cost-utility analyses in total joint replacement in hip and knee patients (ITT analyses).
- Table 10.** Per-protocol analysis of cost-utility of total joint replacement in hip and knee patients.

**ABBREVIATIONS**

ADL	Activities of daily living
AQoL	Australian Assessment of Quality of Life
BMI	Body mass index
C	Cost
CEA	Cost-effectiveness analysis
CEAC	Cost effectiveness acceptability curve
CUA	Cost-utility analysis
CI	Confidence interval
DDD	Defined Daily Dose
DSM	Disease specific medication
EQ-5D	EuroQol (5 dimension)
FHDR	Finnish Hospital Discharge Register (HILMO)
HHS	Harris Hip Score
HRQoL	Health-Related Quality of Life
HUCH	Helsinki University Hospital
ICER	Incremental Cost Effectiveness Ratio
ICUR	Incremental Cost-Utility Ratio
ITT	Intention To Treat
KSS	Knee Society Rating System
MAU	Multi-Attribute Utility
M01A	Anti-inflammatory and anti-rheumatic products, non-steroids
NfWT	Non-fixed waiting time
NICE	National Institute for Health and Clinical Excellence
NSAID	Non-steroidal anti-inflammatory drug
OA	Osteoarthritis
OECD	Organisation for Economic Co-operation and Development
PRO	Patient-reported outcome
QALY	Quality Adjusted Life Year
QOL	Quality of Life
RCT	Randomized controlled trial
SD	Standard Deviation
SF-6D	Short Form (6 dimension)
STAKES	National Research and Development Centre for Welfare and Health
SWT	Short waiting time
TB	Total benefit
TC	Total cost
THL	National Institute for Health and Welfare
THR	Total Hip Replacement
TJR	Total Joint Replacement
TKR	Total Knee Replacement
U	Utility
WOMAC	Western Ontario and McMaster Universities Osteoarthritis Index
WT	Waiting time

## 1 INTRODUCTION

In many Western countries, long waiting times for elective surgeries are of concern. Major joint replacement is an example of surgery with a high demand and relatively long waiting periods for patients (Siciliani and Hurst 2003). The mean waiting time for elective surgical procedures is more than three months in several countries, and waiting times can stretch out to years. As the population gets older, the prevalence of slowly progressive diseases, such as osteoarthritis in the hip and knee joints, is increasing.

Advances in surgical technology and anaesthesia have improved the range, safety and effectiveness of the surgical procedures. There have been a significant number of attempts to estimate both the adverse and positive consequences of waiting. Generally, the consequences of long waiting times are supposed to be: deterioration in the condition for which treatment is waited, loss of utility from the delay, a significant joint pain or disability, an increase in the costs of surgery and other treatments and health care services pre- and postoperatively. However *The waiting time study* (Siciliani and Hurst 2003) found little evidence from the medical literature of significant deterioration of health or worsening surgical outcomes as a result of waiting for elective surgery in those countries where waiting times are from three months up to six months (Siciliani and Hurst 2003). Furthermore, the conclusions in the systematic review of Hoogbeem et al. (2009) were that osteoarthritis (OA) patients waiting less than 180 days did not experience deterioration in pain or functioning while waiting for total joint replacement (TJR). However, according to a review by Hoogbeem et al. (2009), more quality studies are needed.

Johanna Hirvonen (2007) used data from the same study population as in the present study to evaluate the effect of waiting time on health outcomes (HRQoL, pain, and physical function) during the waiting time, i.e., from baseline to admission and service utilization by the admission. According to Hirvonen (2007), use of health and social services was low in both waiting time groups while waiting, and longer waiting times did not result in poorer HRQoL at admission. The conclusion was that those who waited longer used health and social services for a longer period, and this would represent a negative impact of waiting. The present study continues to report results from the baseline to the end of the follow-up and include results concerning QALYs and costs.

Patients waiting for TJR have poor quality of life and they have difficulties in functioning and daily activities (Derrett et al. 1999; Kelly et al. 2001; Croft et al. 2002; Ackerman et al. 2005; Hirvonen et al. 2006; McHugh et al. 2008). Furthermore, there is a significant improvement in HRQoL and increase in the QALYs gained after surgery, and a number of studies have asserted that TJRs in hip and knee patients are cost-effective (Bourne 1996; Lavernia et al. 1997; Rissanen et al. 1997; March et al. 2002; Räsänen et al. 2007; Xie et al. 2010).

The overall objective of this study was to evaluate the effect of waiting time on health outcomes and costs of total joint replacement in hip and knee patients. Furthermore, the theoretical frame of access to care and the phenomenon of waiting time are discussed and some examples of actions to reduce waiting times in those countries where waiting times to TJR have been of concern (e.g. England, Sweden and Finland) are presented.



## 2 OSTEOARTHRITIS

### 2.1 Prevalence of osteoarthritis (OA)

Osteoarthritis is the most common cause of musculoskeletal disability and pain in the world. Incidence and prevalence will increase in the coming years due to the ageing of the population. The rapid increase in persons 55 years of age and older in Western countries implies that OA is becoming a major public health care problem (van Es et al. 2011).

Osteoarthritis is the most common joint disease in the world. Worldwide estimates are that 10% of men and 18% of women over 60 years old have symptomatic OA. One-third of people over 65 years old have knee OA that can be seen on X-ray. 70% of people over 70 years old have x-ray evidence of OA. Before the age of 50, OA is more prevalent among men than women. However after the age of 50, women are more likely to be affected by OA than men (OECD 2011).

In Finland, about 400,000 people are affected by OA. Furthermore, 5.7% of males and 4.6% of females over the age of 30 suffer from clinically diagnosed hip OA, and the prevalence of clinical knee OA is 6.1% in men and 8.0% in women. The incidence of radiological and clinically diagnosed knee and hip OA increases with age both in men and women. In hip patients, the corresponding OA rates in the earlier Mini-Finland Health Survey (Heliövaara et al. 1993) were 4.6% and 5.5% for males and females, respectively, i.e., there has been hardly any change in its prevalence since 1990. In men under 75 years of age, the prevalence of hip osteoarthritis has remained pretty much unchanged, and in older women, the prevalence of hip osteoarthritis has either remained unchanged or has marginally increased (Aromaa and Koskinen 2004). In knee patients, the corresponding rates in the Mini-Finland Health Survey (Heliövaara et al. 1993) were 5.5% and 14.5% for males and females, respectively. Among women, the prevalence of knee osteoarthritis has thus dropped by more than one-half, which is mainly attributable to changes in the age groups under 75 years (Table 1) (Aromaa and Koskinen 2004).

**Table 1.** Prevalence (%) of hip and knee osteoarthritis according to the examining physician's diagnosis in Finland 2000.

Age	30–44	45–54	55–64	65–74	75–84	85+
<b>Hip OA</b>						
Men	0.5	1.8	5.2	12.1	20.3	41.8
Women	0.4	0.7	3.1	11.6	20.0	24.6
<b>Knee OA</b>						
Men	0.3	2.6	9.2	10.6	16.3	45.8
Women	0.4	2.2	8.1	18.4	31.3	35.3

Source: Aromaa and Koskinen: Health 2000. Health Examination Survey.

## 2.2 Diagnosis of OA

There are no simple and generally acceptable diagnostic criteria for hip (M16) and knee (M17) osteoarthritis (ICD10 classification 1994). According to the American College of Rheumatology's (American College of Rheumatology 2012) clinical/radiological sets of criteria, the diagnosis of OA is based on the symptoms described by the patient and the clinical and radiological findings. Diagnosis is made by reasonable certainty based on a patient's history, clinical examination and x-ray (Zhang et al. 2009 and 2010b).

OA is a common disorder of synovial joints characterized by cartilage degeneration with secondary changes in an adjacent bone. It is a common response to a variety of metabolic, anatomical and physiological conditions (Atkinson 1984). Primary OA is a chronic degenerative disorder related to aging but not caused by it. Osteoarthritis is a disease affecting the entire joint, but its aetiology remains mostly unknown. A characteristic feature of OA is the destruction and later decay of the articular cartilage. These changes appear in a radiograph as narrowing of the intra-articular space. Articular changes generally proceed slowly over the years, and the damaged cartilage does not regenerate. Radiological changes in OA are not always associated with pain (Current Care Guidelines 2007).

The most important risk factors for knee and hip OA are obesity, joint injuries and excessive physical stress. Karlson et al. (2003) studied the relation of potential risk factors leading to hip replacements. According to this study, higher body mass index (BMI > 30) was associated with an increased risk of OA. Because OA is most common in old people, it is often proposed that the disease is part of the aging process. Age had a positive association: women  $\geq 70$  years of age were nine times more likely to have OA than those under 55 years old (Karlson et al. 2003).

## 2.3 Treatments of OA

The aim of treatment for OA is to improve functioning and reduce pain. In its early stages, OA is treated with non-surgical measures. Lifestyle modification, such as weight loss and exercise, support devices, and analgesics are the mainstay of first-line treatment. Simple weight loss can reduce stress on weight-bearing joints and can result in reduced pain and increased function (Jordan et al. 2003; Zhang et al. 2005)

### 2.3.1 Medication

Because OA is the most common and slowly progressive disease among the elderly population, disease-specific medication has a very important role in first-line pharmacological treatment. The purpose of medication is to relieve symptoms. The most common first-line medication is paracetamol (N02BE01). It is effective as or somewhat less effective than non-steroidal anti-inflammatory drugs (NSAIDs) or COX-2

selective NSAIDs in reducing OA pain (Zhang et al. 2010a). Paracetamol is recommended as a first-line drug, and NSAIDs are only recommended as add-on therapy, if there is insufficient pain relief (Flood 2010). If the analgesic effects of paracetamol and NSAIDs are not sufficient or if they cannot be used due to adverse effects, OA pain can be treated with opioids (Avouac et al. 2007). The goal of medication is to reduce the symptoms, namely joint pain, stiffness, and swelling. Drugs are also used postoperatively as pain killers.

During the present study, in 2004 the Finnish Social Insurance Institution paid drug reimbursements for anti-inflammatory and anti-rheumatic products (M01A) to 880,700 patients; and in 2006, the reimbursement was paid to 1.2 million patients (Finnish statistics on medicine 2006).

The consumption of anti-inflammatory and anti-rheumatic products (M01A) was lower in 2006 (68.3 DDD/1,000 inhabitants/day, a decrease of 11%) than in 2005. Their cost decreased even more than their consumption from year 2005 to 2006 (24%) (Table 2). Until 2010, the consumption of anti-inflammatory and anti-rheumatic products (M01A) has increased to 83.3 DDD/1,000 inhabitants per day (Finnish statistics on medicine 2010).

**Table 2.** Costs and consumption of M01A<sup>a</sup> and one over-the-counter medicine (N02BE01<sup>b</sup>) in 2004–2006.

Year	Consumption of M01A <sup>a</sup> DDD <sup>c</sup> /1,000inh/day	Costs/1,000 €	Consumption of N02BE01 <sup>b</sup> DDDC/1000inh /day	Costs/1,000 €
2004	74.75	61,833	3.64	13,217
2005	76.66	56,139	4.62	16,459
2006	68.34	42,584	3.05	11,306

<sup>a</sup> M01A: Anti-inflammatory and anti-rheumatic products, non-steroids.

<sup>b</sup> N02BE01: Paracetamol.

<sup>c</sup> DDD: Defined Daily Dose.

Source: Finnish statistics on medicine 2006.

### 2.3.2 Physiotherapy

The general aims of exercise therapy are to improve functioning, increase levels of activity, and to encourage an adequate way of dealing with complaints (van Es et al. 2011). Appropriate physical exercise represents the basis of OA care before and after surgery. Physiotherapy in knee patients has been shown to improve function and quality of life, to decrease pain, and delay the need for surgical intervention. Exercise by a physical therapist has even been shown to be more effective than medications in treating OA of the knee (van Baar et al. 2001). International guidelines recommend exercise therapy as part of the treatment. According to a review of The Osteoarthritis

Research Society International (OARSI), exercise therapy was shown to reduce pain and improve physical functioning in knee patients (van Es et al. 2011). In hip patients, guided physical exercises as a treatment may reduce pain and improve functional capacity. However, there is less available evidence for the effectiveness of exercise therapy in hip OA than in knee OA (van Es et al. 2011). Postoperative exercise programs are essential procedures when recovering after TJR both in hip and knee patients. Exercises aim at quickly regaining motion following the surgery, preventing muscle loss, which is inevitable after surgery, rebuilding muscle strength and preventing stiffness of the new knee joint. It is important to carefully follow the rehabilitation instructions given by the physical therapists and doctors (Kuster 2002).

Of the other physical treatments, cold therapy may reduce knee-joint swelling and improve thigh muscle strength, but it has not been shown to have any effect on the pain caused by OA of the knee (Adie et al. 2012). Acupuncture treatment appears to decrease pain in OA of the knee, at least in the short term, but reliable evidence of long-term amelioration of pain and an increase in functional capacity is nevertheless lacking (Scharf et al. 2006; Zhang et al. 2010a).

### 2.3.3 Surgery

If the conservative treatment options are ineffective, joint replacement surgery may be required in advanced cases. Primary total hip replacements (THRs) and total knee replacements (TKRs) are surgical procedures to remove the injured part of the joint, replacing it with a new artificial part. THRs or TKRs have become standard surgical procedures when patients have intractable pain that is unresponsive to conservative treatments.

The volume of TJR in Western countries and Finland

The number of THRs and TKRs has increased rapidly over the past decade in most OECD countries. On average, the rate of THRs increased by over 25 % between 2000 and 2009. The growth was even higher for TKRs, nearly doubling over the past decade. There is considerable variation across countries in the rate of hip and knee replacements. Germany, Switzerland and Austria have high rates of both THRs and TKRs. The United States and Germany have the highest rate of TKRs, even though the population structure of the United States is younger than that of Germany (OECD 2011).

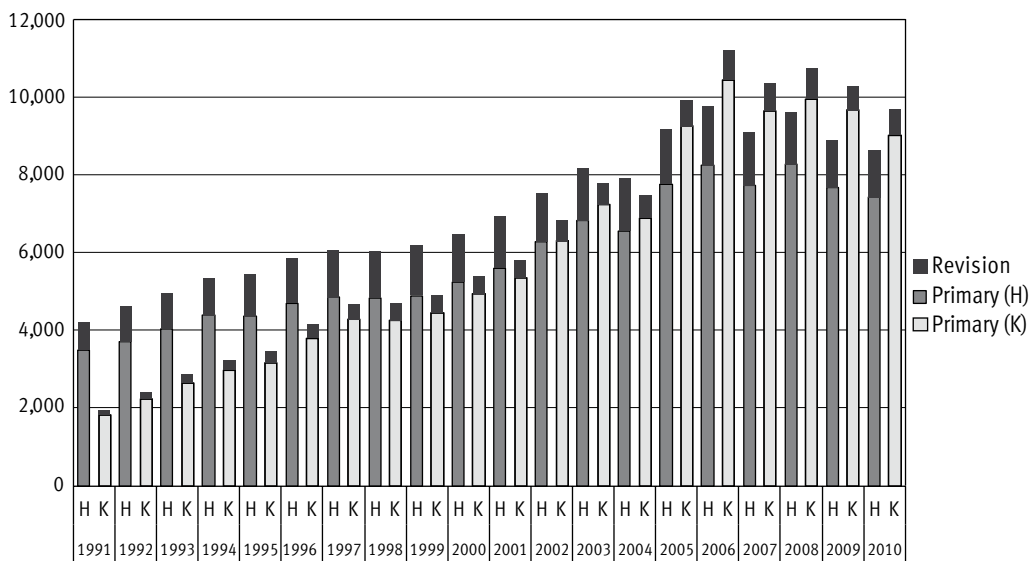
Annually, over three-quarters of a million surgical total hip and knee replacements are carried out in the U.S. (Hoozeboom et al. 2009). In England, the National Annual Report summarizes the data and findings for hip and knee procedures carried out in England and Wales. The development in 2006–2010 in primary THR has been from 58,445 to 77,800, and in TKR from 61,648 to 86,067 surgical procedures (National Joint Registry 2011).

In Finland, a total of 14,219 primary TJRs were performed in 2004 (7,345 hip and 6,874 knee ) with the median waiting time for surgery being 181 days (153 days for hip replacement and 209 days for knee replacement) (National Research and Development Centre for Welfare and Health 2006). During the present study from 2005–2006, the number of TJRs performed increased (9,316 for hip and 10,411 for knee), and in 2009 the number of TJRs has decreased once again to the level of 2005. (THL 2011; Figure 1.)

A total of 18,331 hip and knee replacements were reported to the Finnish Arthroplasty Register in 2010. According to this register, 7,416 were primary THRs and 9,020 were primary TKRs. The total number of operations decreased by 829 in 2009. In 2010, TJRs were performed in 60 different hospitals, on average 40% of the operations were performed in central hospitals, 30% in district hospitals and 10% in private hospitals. (THL 2011.)

A surgical option involving replacement of the hip or knee joint with artificial components has been shown to be a highly effective and cost-effective treatment that results in improvement in patient functioning and quality of life (Bourne 1996; Rissanen et al. 1997; March et al. 2002; Räsänen et al. 2007; Xie et al. 2010).

**Figure 1.** The number of total joint replacements in Finland during 1991–2010.



H = hip, K = knee.  
Source: THL 2011.

## 2.4 Co-morbidity and OA

According to the Finnish Health 2000 Health Examination Survey, over 80% of the inhabitants in the age group of 30 to 44 years regarded their health as good or fairly good. In comparison, among those close to retirement age (55 to 64 years), about half felt their health state at least fairly good. In addition, only one-quarter of persons aged 75 or over said that their health was good or fairly good. Similarly, one-third of the responders in the age group of 30 to 44 reported that they suffered from at least one chronic illness. In the age group of 65 years and over, the prevalence of at least one chronic disease has increased up to 80–90%. The prevalence of chronic illness was about the same among men and women in all age groups (Aromaa and Koskinen 2004).

Co-morbid or coexisting disease refers to the occurrence of two or more diseases in the same individual. Each co-morbid disease may have its own effect on health-related quality of life (HRQoL) while also having a clinical effect on a patients' sense of well-being. Co-morbidity has an important cofounder effect on HRQoL of patients waiting for TJR. It is also important for studies of patients with chronic disease, in whom mortality is rare and the goal of medical care is to control the course of the disease and maximize the quality of life (Xuan et al. 1999). The most common chronic diseases in Finland are: cardiovascular disease, musculoskeletal disease, diabetes, and lung diseases (Aromaa and Koskinen 2004).

Evaluations on how different diseases affect patients' HRQoL focus mostly on the index disease, considering the effect of co-morbidity to a lesser extent. However, when the focus is on the consequences on the costs of the medical care – co-morbidity becomes an important factor.

### 3 LITERATURE REVIEW OF WAITING TIME AND COST-EFFECTIVENESS OF TOTAL JOINT REPLACEMENT

The focus of the literature review was to summarize the findings concerning the effect of waiting time on HRQoL and other health outcomes and the costs of total joint replacement (i.e., cost-utility) in randomized clinical trials where hip and knee patients were followed from baseline to one year after the TJR. Several studies have been published about the effect of waiting time on HRQoL in patients waiting for TJR, and use of quality-adjusted life years (QALYs) for estimating the effectiveness of health care. Earlier systematic literature reviews from Ethgen et al. (2004) and Hoozeboom et al. (2009) have been published relating to these outcomes.

#### 3.1 Results from the earlier systematic literature reviews

The objective in a systematic literature review of Ethgen et al. (2004) was to review the literature regarding the outcomes of total hip and knee arthroplasty as evaluated by HRQoL instruments. In this review, the Medline and EMBASE medical literature databases were searched from 1980 to 2003, with the following terms: hip and knee arthroplasty/replacement, quality of life and outcomes. The findings of this systematic review, related to our study were:

Total hip replacements (THRs) and total knee replacements (TKRs) were found to be effective in terms of HRQoL improvement. When improvement was found to be modest, the role of co-morbidities was highlighted. Further, when patients reported poorer preoperative HRQoL, they were more likely to experience greater improvement. A longer waiting time did not necessarily diminish the potential improvements that patients could achieve with surgery. Shortening the waiting times and prioritizing patients in terms of severity would serve to reduce the burden of waiting for surgery (Ethgen et al. 2004).

Most of the studies of the review of Ethgen et al. were observational studies rather than randomized controlled studies. According to this review, no studies have fully investigated the effectiveness of THR and TKR compared to non-operative care. Yet there were studies where patients were randomized with respect to different prosthetic types.

The objective in the systematic review of Hoozeboom et al. (2009) was to describe changes in pain and functioning in patients with OA awaiting TJR and to assess determinants of this change. In this review, the Medline, Embase, Cinahl and Cochrane databases were searched from 1999 through June 2008.

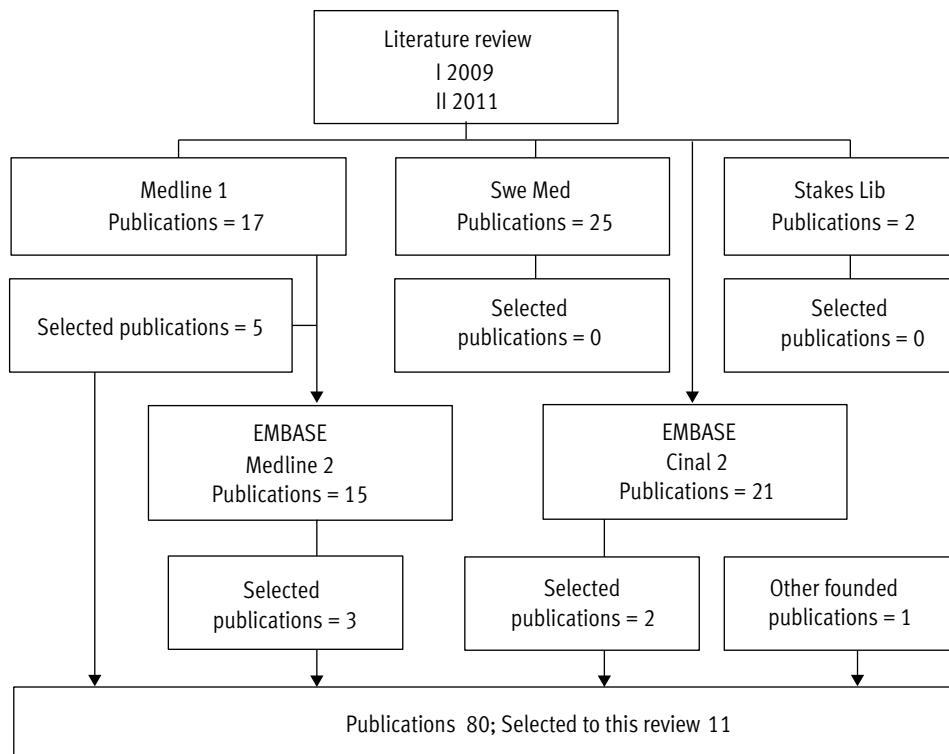
The main results of this review indicated that there was strong evidence that pain and perceived functional status did not deteriorate in patients waiting for major joint replacement during a period less than 6 months. Conflicting evidence was found for

the change in self-reported functional status while waiting less than 6 months for TKR. Hoozeboom et al. concluded that there were only a few high-quality studies examining the effect of long waiting times ( $\geq 180$  days) on pain and functioning. Only six randomized clinical trials were reported including Hirvonen et al. (2007a), using data from this study and five other studies from preoperative exercises and rehabilitation.

### 3.2 Literature review

The present review of the literature published from 2002 to 2009 was carried out first in 2009 and re-run in 2011 using the PubMed, Cinahl, Medic, Arto and Linda and SweMed databases. The following search terms were used: randomization, waiting time, waiting list, hip, knee, HRQoL, cost-effectiveness, cost-utility, health care costs, arthroplasty/total joint replacement. The first selection was based on the contents of the abstract. Result of the present literature review is presented in Figure 2.

Figure 2. Flow diagram for study selection.





Studies were eligible for inclusion in the review if: (1) the data were prospective in nature and the waiting time was measured, (2) generic and disease-specific instruments were used to measure HRQoL, (3) the cost-effectiveness or costs related to waiting time or TJR were calculated, and (4) follow-up time was at least three months. A total of 80 studies were found, of which 69 were excluded because they did not contain all the main criteria, which were: randomization, waiting time, HRQoL and costs related to total joint replacement of hip and knee patients. In addition, six of these were our own studies from the same data and were excluded (Hirvonen et al. 2006, 2007a, 2007b and 2009; Tuominen et al. 2009 and 2010; articles II and III included in this thesis). The present literature review contains 11 studies.

### *3.2.1 Results of the present literature review*

There were no randomized studies about cost-effectiveness/cost-utility related to different waiting times. The costs related to osteoarthritis were calculated in 11 studies, the direct costs of hospital stay and prosthesis in two studies (Fielden et al. 2005; Núñez et al. 2007), and seven evaluated the cost-effectiveness of the surgery (Segal et al. 2004; Brauer et al. 2005; Brunenberg et al. 2005; Räsänen et al. 2007; Navarro Espigares et al. 2008; Xie et al. 2009; Higashi et al. 2011). The effect of waiting time on HRQoL and costs were addressed in 2/11 studies. The follow-up time from baseline, when patients were placed on the waiting list up to 12 months, was in the studies of March et al. 2002; Brunenberg et al. 2005; Xie et al. 2009; and Higashi et al. 2011. The main findings of the present literature review are summarized in Table 3 (pp. 24–25).

#### Outcome measurements

Both generic and disease-specific quality of life instruments were used. The EQ-5D (4/10), SF-36 (4/10), AQoL (1/11) and 15D (2/11) were used to measure HRQoL. Pain and function of hip or knee patients were measured by the WOMAC (4/11), Harris Hip Score (HHS), and Knee Society Rating System (KSS score) used in the four studies.

The EQ-5D, AQoL and 15D are generic Multi-Attribute Utility (MAU) instruments, designed to generate a single index value for each health state (Hawthorne et al. 2000). The aim of these instruments is to provide a standardized, non-disease-specific survey instrument and generate a cardinal index of health. The fully scaled MAU instrument may be used to estimate the utility of all possible health states (Hawthorne et al. 2000; Bowling 2005). The Short Form-36 health survey questionnaire (SF-36) is the most frequently used HRQoL instrument. It consists of 36 health-related questions. The SF-36 is a profile instrument and thus does not belong to the category of instruments producing a valuation-based single index number (Bowling 2005), although a valuation algorithm has been elicited to also generate a single index number (SF-6D) (Brazier et al. 2007).

Table 3. Summary of the literature review.

Authors, year Country	Design	Aim of the study	Subjects	Measurement Instrument	Costs	Main results
March et al. 2002 Australia	Before and after trial questionnaires	To address costs of TJR from patients' perspective by determining patients out-of-pocket costs during the first year after TJR.	98 TKR patients 76 THR patients	SF-36 WOMAC	Prospective cost	WT was 3 months. Poorer pre-surgery health status predicts higher expenditure during the first postoperative year.
Segal et al. 2004 Australia	Evidence-based priority setting model Literature review	To evaluate the selected interventions, using evidence on outcomes and costs and comparing marginal cost-effectiveness ratios of interventions.	THR and THR patients	AQoL, SF-36 EQ-5D Variety of disease-specific instruments	Cost, QALYs	The most effective intervention was THR. The estimated utility gain from THR was 0.305. Estimated QALY gain over lifetime from THR was 3.52 per hip and 2.086 per knee TKR.
Fielden et al. 2005 New Zealand	Prospective cohort study	To determine the economic and health costs of waiting for THA.	153 patients waiting for THA	WOMAC, EQ-5D	Monthly cost	Mean WT was 5.1 months and mean total cost of waiting for surgery was NZ 4305 dollars per person. Longer waits for THA incur higher economic costs and deterioration in physical function while waiting.
Brunenberg et al. 2005 Netherlands	Before and after trial	To determine the ICER of a clinical pathway for joint replacement surgery (JRP), as compared with usual care.	160 patients undergoing THR and TK	HHS, AKS, EQ-5D	Hospital costs	The clinical pathway dominates usual care and is a highly cost-effective approach to contain costs related to JRP without adverse consequences for patients. JRP resulted in significant cost savings as a result of reduction in hospital stay. Functional level and QoL were higher in JRP group.
Navarro Espigares et al. 2008 Spain	Prospective cohort pre-test/ post-test study	To determine the ICER of THR and TKR and to compare the results with the costs/QALY for other medical procedure	80 patients waiting for and undergoing TJR	Womac, EQ-5D	Direct costs of intervention, with length of hospital stay and prosthesis	The costs of both TKR and THR were lower than the threshold of 30,000€/QALY considered acceptable in Spain and compared favourably with other medical and surgical procedures.
Räsänen et al. 2007 Finland	Head-to-head comparison	To evaluate CE of THR or TKR in terms of QALYs and costs.	223 patients who were enrolled for THR or TKR	15D	Direct hospital costs	Hip and knee replacement both improve HRQoL. The cost per QALY gained from knee replacement is twice that gained from hip replacement.

Authors, year Country	Design	Aim of the study	Subjects	Measurement Instrument	Costs	Main results
Núñez et al. 2007 Spain	Cross-sectional study	To determine HRQoL in patients with severe OA on waiting list for TKR and compare it with general Spanish reference population values, the influence of socio-demographic and clinical variables on HRQoL dimensions and the use costs of resources related to knee OA.	100 consecutive outpatient	WOMAC, SF-36	Pharmacological treatment and of economic resources related to knee OA	HRQoL of patients on a waiting list is worse than that of the reference population. The main costs of these patients were non-medical resources, mainly due to function limitations and loss of autonomy.
Xie et al. 2009 Canada	Prospective observational cohort study	To evaluate (ICUR) incremental cost-utility ratio of TKR versus UKA (unicompartmental knee arthroplasty), in patients with knee OA.	431 TKR patients 102 UKA patients	OKS (Oxford knee score) SF-36	Costs related to surgery	From the societal perspective, the ICUR was US\$65,245/QALY; from the patient perspective, the ICUR was US\$60,382/QALY. TKR gained more QALYs at higher costs compared to UKA.
Hawker et al. 2008 Canada	Population-based nested case-control study of the costs of THR and TKR	To examine changes in direct health care costs and arthritis severity after TJA for hip/ knee arthritis compared with contemporaneous changes in matched controls.	185 TJA patients of these 183 cases and controls were matched	Arthritis severity, general health status	Direct health care costs	Compared with matched controls, arthroplasty is associated with significant reduction in pain, disability and direct costs related to arthritis.
Montin et al. 2009 Finland	Follow-up study	To determine the economic outcomes related to THR from the perspective of HRQoL.	100 THR patients	15D	Out-of-pocket costs	Health care costs were over 90% of the total out-of-pocket costs. HRQoL was worse before the surgery and improved after surgery. There was no significant correlation between costs and HRQoL. When deciding on the timing of surgery, patients' characteristics (especially the level of pain and HRQoL) should be evaluated, as they may predict patients' service use and ability to manage at home after surgery.
Higashi and Barendregt 2011 Australia		To evaluate CE of THR and TKR in modelling the severities of right and left joint separately.	Australian OA population			Both THR and TKR were cost-effective under the AUD 50,000/DALY (disability adjusted life years) threshold level.

The Western Ontario and McMaster Universities OA index (WOMAC) is the most used proprietary set of standardized condition-specific instruments to evaluate patients with OA of hip or knee, including pain, stiffness, and physical functioning of the joints. When used in clinical studies, the WOMAC pain and function subscales perform comparably or better than other tests in being responsive to change from experimental interventions, but this varies for the different subscales and types of intervention. The Western Ontario and McMaster Universities OA index (WOMAC) is the most used condition-specific instruments for the assessment of hip or knee OA and is recommended by OMERACT (Outcome Measures in Rheumatology Clinical Trials) (American College of Rheumatology 2012).

The effect of waiting time on HRQoL and costs of TJR

In the studies of March et al. (2002) and Fielden et al. (2005, where the length of waiting time was measured, the effect of waiting more than six months was associated with higher total mean costs than waiting less than 6 months and led to poorer physical function before the operation. Furthermore, patients with poor health status showed greater improvements in pain and physical function and HRQoL after the intervention (Fielden et al. 2005 Räsänen et al. 2007; Navarro Espigares et al. 2008; Hawker et al. 2008; Montin et al. 2009; Higashi and Barendregt 2011). Six studies also determined the economic and health outcomes of waiting in hip or knee patients. The incremental cost-effectiveness ratio (ICER) was calculated in five of these studies and suggested that TJR was a highly cost-effective procedure. None of these studies evaluated the differences between short (< 3 months) and longer (> 3 months) waiting times.

The interest in evaluating cost-effectiveness and the effect of waiting time is increasing. Most of the studies are before-after studies and the follow-up times are short. For clinical outcomes, randomized controlled trials are the standard and accepted approach for evaluating interventions. Pragmatic randomized trials provide a suitable environment not only for assessing clinical effectiveness but also for comparing costs. Carrying out an economic evaluation alongside a randomized controlled trial allows detailed information to be collected about the amount of resources used by each patient. This information allows an estimate of the cost of treatment for each individual patient, producing a set of cost values, which will be referred to as patient-specific cost data (Barber and Thompson 1998).

Randomizing patients in clinical trials is not novel. But according to earlier systematic literature reviews, no earlier studies have been published with randomization of patients waiting for surgery, one year postoperative follow-up, and analysis based on the intention-to-treat (ITT) principle.

## 4 WAITING TIME AND ACTIONS TO SHORTEN IT

### 4.1 Access to care

It is a common assumption that long waiting times for elective surgery are a heavy burden both to patients and those who deliver health care services. The municipalities are under pressure from both sides (supply and demand) because of the lack of health care resources to reduce long waiting lists.

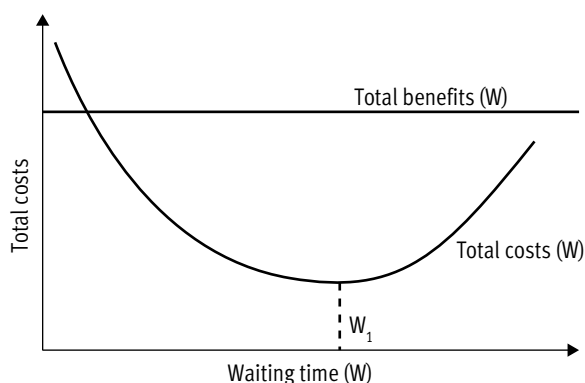
Access to care at hospitals is usually managed by scheduling demands for service. In the scheduling of surgical demand, the patients available for the next service period are identified, and hospital resources are reserved to ensure appropriate care before and after an instance of surgery. Within services, patients are selected from waiting lists and scheduled for surgery on the basis of urgency, best use of allocated operating time, and availability of hospital resources (Sobolev and Kuramoto 2008).

Patients are placed on a wait list after the decision to perform surgery is made. Before being added to the operating room schedule, each patient is assessed by his or her surgeon to determine suitability for the surgery. If a patient is deemed unfit, scheduling of the operation may be postponed. Scheduling the operation may be delayed for a number of reasons: the patient decides to postpone the surgery, the hospital ward or operating room is unavailable at the scheduled time, or the doctor decides to send the patient for additional preoperative investigation. Also the patient may die or their condition deteriorates or the operation becomes unnecessary (Sobolev and Kuramoto 2008).

What is the optimal waiting time? There is inconsistent scientific evidence on whether there is a relationship between the length of waiting time and the outcomes of the surgery in terms of cost-utility. In agreement with Siciliani and Hurst (2003), the optimum waiting time will not be zero. It can be cost-effective to maintain short queues of elective patients because the adverse health consequences of short delays are small and because there are savings in hospital capacity from allowing queues to form (Siciliani and Hurst 2003). A simple criterion has been proposed that waits should be reduced until the costs of doing so exceed the benefits (Schaafsma 2006), and Siciliani and Hurst defined an optimal waiting time where the costs and benefits of making further reductions were equal (Harrison and Appleby 2010). Figure 3 (p. 28) describes the hypothesis of Siciliani and Hurst on the effect of different waiting times on the health benefits and costs of surgery. The mean waiting time ( $W$ ) is shown on the horizontal axis; the total benefits and costs of surgery are on the vertical axis. The total benefits of surgery will be constant at a given rate of surgery, which might, or might not, be the optimum rate. That is because waiting times can only be varied for a given rate of surgery if surgeons vary their clinical thresholds for admitting patients on lists. This will not in itself affect benefits because changes in formal waiting must be offset exactly by equal and opposite changes in hidden waiting. Only if the surgery rate changes will the benefit curve depicted in Figure 3 change, shifting upwards or downwards. The total costs of supplying surgery for different waiting times are assumed to be roughly

in a U-shape. Costs fall as waiting time increases, and there can be savings in surgical capacity if a waiting list is formed and additional elective patients are called in for treatment when there are lulls in the flow of emergency patients. This reduces the probability that supply exceeds demand, leaving capacity unused. Furthermore, the administrative costs of a waiting list will increase with the length of the waiting time and waiting lists as will the diversion of clinical resources to a regular reassessment of patients on the list. Beyond a certain point, rising administrative and clinical costs are likely to outweigh falling capacity costs as the length of the list increases (Iversen 1993; Siciliani and Hurst 2004).

**Figure 3.** Optimal waiting time  $W_1$ .



Source: Siciliani and Hurst 2004.

However, there is little scientific evidence about the relationship of length of the waiting time and cost-effectiveness of elective surgery.

#### 4.2 Waiting time and health care reforms

When comparing surgical procedures, Finland, England and Sweden have been the countries with the longest waiting times (Siciliani and Hurst 2003). Because of the long waiting times for elective surgery, health care reforms have come into force in recent years. The primary aims of these reforms were to reduce and equalize waiting times for the selected treatments (Hanning and Winblad Spångberg 2000). Some countries have explicit waiting time prioritization, recommended admission within a certain time period, or priority scoring systems, points for patients, for elective surgery (Gravelle and Siciliani 2008).

#### *4.2.1 Health care reforms in England and Sweden*

The health care reforms have a long history. Since the establishment of the National Health Service (NHS), the British Government has introduced a large number of health care reforms. The main objectives have been to: reduce waiting times, increase productivity, improve indicators of health outcome, improve quality and reduce inequalities in health across socio-economic groups and across regions with the worst and best health outcomes indicators (Hagen and Kaarbøe 2006; Oliver 2006).

One of the latest implementations in health care in England is benchmarking for acceptable waiting time. There is a timeframe to represent a goal, target or standard for the length of waiting time for a specified service. There are financial incentives based on performance, further increasing capacity of hospitals, equipment's and staff, booking systems, nationally consistent prioritization of patients, improving communication between specialists and hospitals, and giving patients certainty of treatment and choices if the benchmarking is not met. There is consistency across benchmarks specifically for hip and knee replacements for waiting times to be no longer than 6 months (Noseworthy et al. 2005).

Between 2000 and 2005, the main focus in English government policy was to reduce waiting times for inpatient services and the first outpatient appointment with explicit targets: the major emphasis being on more choice by patients. Extra funding was provided, and hospitals were directly managed towards published targets. The aggressive deployment of robust performance management system alongside targets, and increasing funding appeared to be a success in reducing waiting times in England (Wilcox et al. 2007).

The health care system in Sweden, as in Finland, is financed primarily through taxes levied by county councils and municipalities (Swedish Association of Local Authorities and Regions 2005; Albin et al. 2010). In 1992, the Swedish government issued a 3-month treatment guarantee for ten elective treatments with long waiting times (e.g. cataract surgery and hip replacements). If the county council could not provide treatment within 3 months, the patient was to be offered treatment at a hospital in some other county council or at a private hospital. The national government supported the guarantee with additional funding, which made the task easier and waiting lists shortened temporarily.

In 1997 a new reform came into force, where health care should give priority to patients with greater needs (Anell 2005). According to Anell (2005), there was an obvious difficulty in instituting a reform that takes into account all relevant health care objectives simultaneously, which may be linked to a pattern of political decision-making. New policies in Swedish health care have tended to focus on one objective at a time, reflecting the most urgent problem. Furthermore, every new reform thereafter carries a seed of new problems, generating demand for additional change.

#### *4.2.2 Health care reform in Finland*

A national project to ‘Ensure the Future of Health Care in Finland’ was commenced in 2001 by the Finnish government to reform and develop health care services in the country. This induced the Finnish government and local authorities to adopt a new plan and funding arrangements for the Finnish health care system. New health care reform was instigated in 2005, the year of the present study. The target of this reform was to reduce waiting times and guarantee equal access and to develop criteria for non-urgent care (e.g. hip and knee replacements) in a reasonable time period for all citizens. The following recommendations for access to care were made: 1) patients contacting health centres should be assessed by a doctor or other health professional within three days; 2) patients requiring assessment by a specialist should be offered an appointment within three weeks; 3) recommended hospital treatment, including elective surgery, should be offered within six months (Ministry of Social Affairs and Health 2004).

Before the reform came into force, the state together with municipalities gave extra financing to hospital districts in the sum of EUR 50 million during 2003–2004 to decrease long waiting times in specialised medical care. Despite the extra money, the waiting times were still long in 2004 and 2005. The information on waiting times and queues for surgery was not transparent before the waiting time reform. When the reform came into force on 1 March 2005, the state started an efficient monitoring of queues in the hospital districts (Mikkola et al. 2008).

The Finnish health care sector invested EUR 382 million (EUR 72 per capita) in 2002–2007 to reach the waiting time targets. A peak year was 2005, when around 2.3% of health spending by municipalities was devoted to waiting-time reform. Investment declined gradually to 0.9% in 2007. Around 69% of extra money due to waiting time reform went to hospitals and 31% went to health centres (Mikkola et al. 2011). The waiting times for most queued operations decreased slightly, but in 2006 the decrease was more extensive, and the number of elective patients that had waited longer than 6 months decreased by 42% (Mikkola et al. 2008) compared to earlier years.

According to international and national experiences, the waiting time reforms have been more or less successful, and a lot of effort is still needed to solve the problems of long waiting times and differences in treatment practices. It is important to identify cost-effective strategies for managing OA patients, i.e., to maximise health benefits given the limited health care resources.



## 5 ECONOMIC EVALUATION

### 5.1 Economic evaluation in health care

The goal of treatment is not only to prolong life, but also to contribute to well-being. The rising burden of chronic diseases puts pressure on societies to find more effective and cost-effective ways to improve citizens' health in circumstances of strict budgeting and cost control. For allocating resources, national political planning and international organisations need evidence of the impacts of interventions (Bowden and Fox-Rushby 2003). To facilitate the comparison of different interventions, measures have been developed to bring together clinical, quality-of-life, and economic outcomes in summary measures such as quality-adjusted life-years (QALYs), cost-effectiveness, and cost-utility ratios (Lyman 2001).

Economic evaluation is the comparative assessment of the costs and benefits of alternative health care interventions (Drummond et al. 2005; Brazier et al. 2007). Interventions that are associated with large or uncertain resource consequences and small or unclear efficacy are most likely to be candidates for economic analysis. There are four main forms of economic evaluations, each dealing with costs, but differing in the way that the consequences of health care programmes are measured and valued (Table 4).

**Table 4.** Measurement of costs and consequences in economic evaluation.

Type of study	Measurement / valuation of costs in both alternatives	Identifications of consequences	Measurement / valuation of consequences
Cost analysis	Monetary units	None	None
Cost-effectiveness analysis	Monetary units	Single effect of interest, common to both alternatives, but achieved to different degrees	Natural units (e.g. life-years gained)
Cost-utility analysis	Monetary units	Single or multiple effects, not necessarily common to both alternatives	Healthy years (quality-adjusted life years, QALYs)
Cost-benefit analysis	Monetary units	Single or multiple effects, not necessarily common to both alternatives	Monetary units

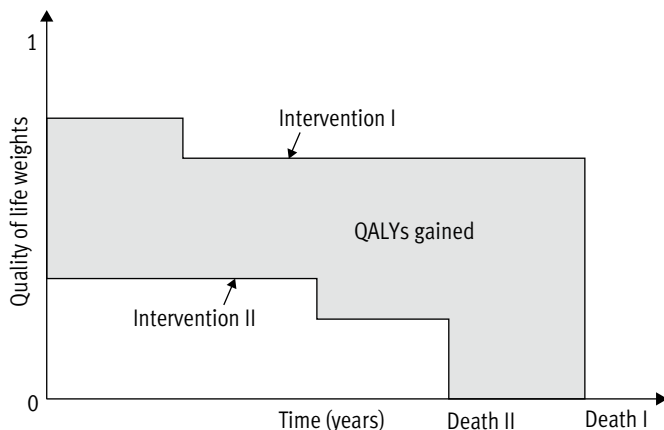
Source: Drummond et al., 2005.

#### 5.1.1 Quality-adjusted life year

A quality-adjusted life-year (QALY) takes into account both the quantity and quality of life generated by health care interventions. It is the arithmetic product of life expectancy (a measure of the remaining life-years), and it reflects the change in survival with a weighting factor for quality of life. A QALY places a weight on time in different health states. A year of perfect health is worth 1 and death is considered to be equivalent to 0. The QALY provides a common currency for measuring the extent

of health gain that results from health care interventions and when combined with the costs associated with their relative worth from an economic perspective (Bowling 2005; NICE 2009). Figure 4 presents a situation where Intervention I provides a consistently greater area under the QALY time curve than Intervention II.

**Figure 4.** QALYs gained from Intervention I in comparison to Intervention II (area between the curves).



Source: Drummond et al. 2005.

### 5.1.2 Cost-effectiveness analysis

Cost-effectiveness analysis (CEA) is a method designed to assess the comparative inputs of expenditures and effectiveness from different health interventions. It is based on the premise that “for any given level of resources available, society wishes to maximize the total aggregate health benefits conferred” (Gold et al. 1996; Drummond et al. 2005; Brazier et al. 2007).

The central measure used in CEA for comparing two alternatives is the incremental cost-effectiveness ratio (ICER), where the difference in their costs is divided by the difference in their effectiveness. When one of the alternatives (intervention) is both more effective and less costly than the comparator, then it strongly dominates the other alternative (Gold et al. 1996; Drummond et al. 2005). Where the benefits of competing alternatives can be measured along a single dimension, then CEA can be used to rank interventions in terms of their ratio of cost per unit of effect. Effects are usually measured in natural units (Gold et al. 1996).

### 5.1.3 Cost-utility analysis

Cost-utility analysis (CUA) is a form of evaluation that focuses on the quality of the health outcome produced by health programmes or treatments. It has many similarities to CEA. In CUA, the incremental cost of a programme from a particular viewpoint is compared to the incremental health improvement attributable to the programme, where the health improvement is measured in QALYs gained. The comparison is possible if the measure of effectiveness is general enough to capture all of the important health dimensions of the effects of the interventions. The results are expressed as a cost per QALY gained. In the QALY approach, the quality adjustment is based on a set of values or weights called utilities (one for each possible health state), which reflect the relative desirability of the health state (Drummond et al. 2005.) CUA distinguishes between those studies that use a generic outcome measure and are potentially comparable across studies. It highlights the crucial role of consumer preferences in valuing the outcomes (Drummond et al. 2005.)

### 5.1.4 Decision rules of CEA and CUA

To decide which of the evaluated interventions is cost-effective, both costs and consequences have to be compared. There are nine different situations to take into account when making this decision about cost-effectiveness (Table 5).

**Table 5.** The cost-effectiveness decision table.

A new intervention compared with an old	Less effective	Same effectiveness	More effective
Less costly	No clear decision (no dominance)  Incremental analysis needed	The new intervention, the new dominates the old (weak dominance)	The new intervention, the new dominates the old (strong dominance).
Same costs	Keep the old intervention, the old dominates the new (weak dominance)	The interventions are equal	Opt for the new intervention, the new dominates the old (weak dominance)
More costly	Keep the old intervention, the old dominates the new (strong dominance)	Keep the old intervention, the old dominates the new (weak dominance).	No clear decision (no dominance)  Incremental analysis needed

Source: Drummond et al. 2005.

In two of the nine situations described in Table 5, there is no dominance and incremental analysis is needed, i.e., the ICER needs to be calculated. The new intervention is considered cost-effective if the society is willing to pay for the additional benefits or if the society considers that the cost savings compensate for the lower effectiveness.

The concept of the value of a QALY has reached prominence in policy and in empirical research due to the creation of national-level health technology assessment agencies. When assessing particular interventions in terms of health gains against the costs of provision, such agencies must in effect put a monetary value on those health gains. In the context of England, where the National Institute for Health and Clinical Excellence (NICE) uses the QALY as its health metric, NICE must decide what value(s) of a QALY to use (Brazier et al. 2007; Donaldson et al. 2005). Since the inception of NICE, the threshold value of the QALY gained has been £20–30,000 (EUR 24–33,000) (Rawlins and Culyer 2004). In Finland, there is no common agreement on the monetary value for QALY gained.

## 6 AIMS OF THE STUDY

The overall objective of this study was to evaluate the effect of waiting time on health outcomes and costs of total joint replacement in hip and knee patients before and after the operation. In more detail, the main questions addressed were:

- 1) Does co-morbidity affect the HRQoL in patients awaiting major joint replacement? (Paper I).
- 2) Does the length of waiting time have an effect on health outcomes, use and costs of medication in hip replacement patients during the follow-up time from baseline to 12 months postoperatively? (Paper II).
- 3) Is the longer waiting time associated with quality of life outcomes and costs of medication in knee replacement patients during the follow-up time from baseline to 12 months after the operation? (Paper III).
- 4) Is it possible to improve the cost-utility of major joint replacement by shortening the waiting time? (Paper IV)

## 7 PATIENTS AND METHODS

### 7.1 Study design

This study is a prospective randomized controlled trial (RCT), where effects of the different lengths of waiting time on health outcomes and costs were evaluated. Patients placed on the waiting list for TJR were randomly allocated to different waiting time groups: an intervention group (patients in short WT group, SWT, waiting time  $\leq 3$  months) and a control group (patients in a non-fixed WT group, NFWT, waiting time  $> 3$  months). Patients were followed in the randomized groups according to the ITT principle through the whole follow-up time.

### 7.2 Feasibility of the study

Four surgical hospitals were chosen from southern Finland. However, one of the hospitals was neither included in the randomization procedure nor in the final analyses because in that hospital patients waited less than three months anyway. Several arrangements were made to create a functional and equal procedure for all patients. In surgery, randomization is not novel, but when stepping in to the hospital routine procedure by selecting (randomizing) patients who were placed on the waiting list, the situation was different. The key persons to recruit the patients after the TJR decision were the practise surgeon and the nurse. In each hospital, we had a named nurse who was a person in charge. Common guidelines for administering the questionnaires were provided in each hospital, and the procedure was weekly under supervision by a researcher. We had some difficulties at the beginning. Some surgeons did not allow their patients to participate in the study, some nurses were unwilling to recruit the patients, and some patients were unwilling to fill in the questionnaires four times during the observation period.

### 7.3 Study population

Between August 2002 and November 2003, 1,236 consecutive osteoarthritis patents were informed about the study in four Finnish hospitals and were invited to participate. Two of the hospitals (the Surgical Hospital and Jorvi Hospital) are part of the Helsinki University Central Hospital, the third is the Coxa Hospital for Joint Replacement in Tampere, and the fourth is Orton Orthopaedic Hospital in Helsinki. Of these eligible patients, 168 were from Orton Orthopaedic Hospital and 1,068 patients were from the other three hospitals.

The patients came for an outpatient surgical assessment by referral from a health care centre, a local central hospital, or a private physician and were recruited to the study when placed on the waiting list through contact with orthopaedic and practice staff. They were recruited to the study in three recruitment periods (for the Coxa Hospital for Joint Replacement) or four (for the other two hospitals), each period lasting three

months. In Orton Orthopaedic Hospital, there was only one recruitment period, and patients were not randomly assigned to either of these groups because surgery performed according to the hospital's routine procedure was done in less than 3 months. The last patient was admitted to the hospital in May 2005.

The key inclusion criteria were: (1) need for a primary total joint replacement due to severe osteoarthritis of hip or knee joint as evaluated by the hospital surgeon; (2) patient was an adult, aged 16 or older; (3) patient was placed on the waiting list in a research hospital; and (4) patient was willing and mentally able to participate in the study. The key exclusion criteria were patients with rheumatoid arthritis, fractures, and congenital haemophilia or congenital deformities.

According to the randomization protocol, a total of 1,068 osteoarthritis patients was informed about the study and asked to participate. Of the eligible patients, 235 (22%) were excluded, 206 patients refused to participate, 4 did not understand Finnish or Swedish, 2 were operated on the contralateral side, one was an inmate at an institution, and 22 for unknown reasons. A total of 833 patients were randomized, and of these, 550 (66%) patients completed all four questionnaires and were included in the final cost-utility analyses. At baseline, 893 non-randomized patients from all four research hospitals were included in the co-morbidity study (Study I). This selection is not shown in the patient flow. (Figure 5, p. 39.)

#### 7.4 Randomization of the study population

Patients were recruited to the study in three recruitment periods (for one of the hospitals) or four (for the other two hospitals), each period lasting three months to avoid the waiting time for the SWT group exceeding three months. Patients randomized into the SWT group were operated within two weeks following the end of each recruitment period. As only half of the hospital's one-month surgical capacity could be allocated to the SWT group, the number of SWTs was restricted and determined specifically for each hospital. The size of the NFWT group was not restricted so as to ensure that all eligible patients placed on the waiting list had an opportunity to be recruited to the study. The two patient groups therefore differed in size.

Computer-generated randomization sequences were produced by the National Research and Development Centre for Welfare and Health and supplied to the hospitals using consecutively numbered and sealed opaque envelopes. The patient's named nurse assigned participants to their groups after the decision for surgery had been made. The randomization envelopes contained information on whether the patient belonged to the SWT group or the NFWT group. Surgeons were blind to patient allocation. For ethical reasons, double-blinding was not possible.

Each patient provided informed consent and completed a self-administered questionnaire when placed on the waiting list, at admission and at three and twelve months

postoperatively. The questionnaires were either distributed to the patients at the hospital or in some cases mailed to patients, as happened with one hospital for the third and fourth questionnaires. All questionnaires were returned by post. The number of patients placed on the waiting list varied from one month to another, being specific to each hospital. Therefore, no advance estimate could be made of the number of patients to be placed on the list. The data collection was successfully completed. The flow of patients is presented in Figure 5.

In this study, standard analysis for randomised groups were made according to the intention-to-treat (ITT) principal. The procedure gives a valid estimate of the effect of intervention assignment on outcome (Mealli and Li 2011). This form of data analysis provides a strategy for analysing data, in which all participants are included in the treatment group to which they were originally assigned at randomization.

As a secondary analysis, we used per-protocol analysis. It is an alternative approach that compares those who were assigned to and received treatment with those who were assigned to and received control, i.e., compares those who appeared to comply with the protocol and completed the treatment as originally allocated (Mealli and Li 2011; Shah 2011).

## 7.5 Outcome measures

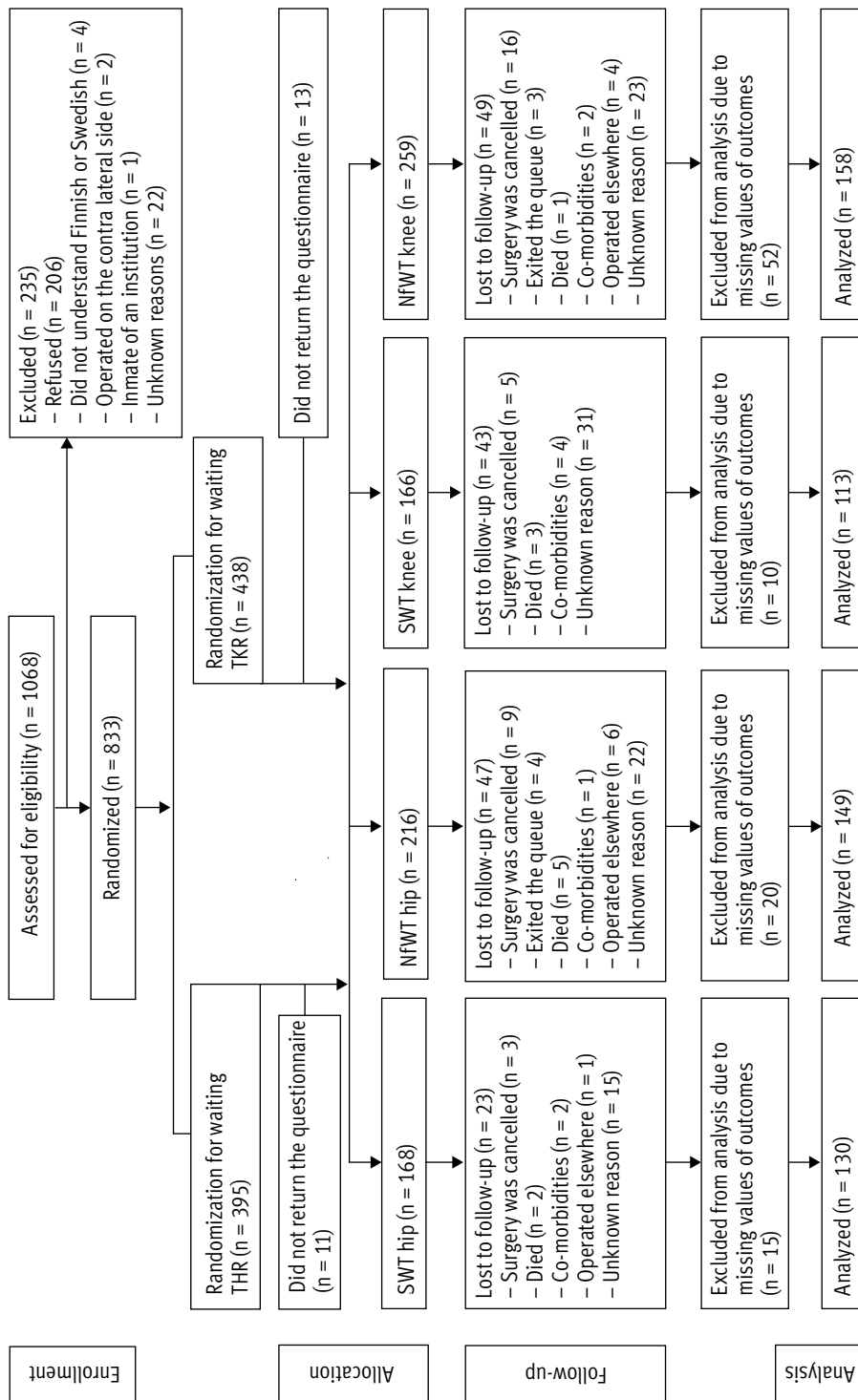
### 7.5.1 *Generic 15D instrument*

HRQoL was measured by the generic 15D instrument. The 15D is composed of 15 dimensions: moving, vision, hearing, breathing, sleeping, eating, speech, excretion, vitality, usual activities, mental function, discomfort and symptoms, depression, distress, and sexual activity. Each dimension has 5 ordinal levels to choose from. The 15D can be used as a profile measure or to give a single index score by means of population-based preference weights. The index score (15D score) ranges from 0 (dead) to 1 (completely healthy). Completing the 15D questionnaire takes 5 to 10 minutes, and it describes the HRQoL of the respondent at present. A difference of  $> |0.03|$  in the 15D score is clinically important in the sense that on average people can feel the difference (Sintonen 1994a and 2001).

The 15D has been used successfully in earlier studies dealing with hip and knee replacement and thus facilitates the comparison to the pre-surgery scores in these studies (Rissanen et al. 1995; Räsänen et al. 2005; Hirvonen et al. 2006). There is earlier evidence that in most of the important properties (reliability, content validity, sensitivity in terms of discriminatory power, and responsiveness to change), the 15D compares at least equally with other similar instruments that produce a valuation-based single index number (Stavem 1999; Sintonen 2001; Hawthorne et al. 2001; Drummond et al. 2005; Brazier et al. 2007; Mook and Kohlmann 2008).



Figure 5. Flow diagram of the progress through the phases of the randomized trial.



Although the SF-36 (Short Form 36) has been widely used in the OA area, it is basically a profile instrument and thus does not belong to the category of instruments producing a valuation-based single index number. However, in 1998, the first algorithm was published to generate a valuation-based single index score (the SF-6D score) from the SF-36. The paper by Hawthorne et al. (2001) showed that the 15D performed at least as well as the SF-6D, if not better. Recent research with that algorithm indicates that especially in the rehabilitation of musculoskeletal disorders, the 15D was at least as responsive as the SF-6D, if not better. These two instruments were anyway the most responsive ones (Moock and Kohlmann 2008).

The scores of 15D are suitable for QALY calculation. In this study, QALYs during the equally long follow-up time was calculated by the following formula (knee patient):

$$QALY_{SWT} = [0.5 (a + b) * h_1 + 0.5 (b + c) * h_2 + 0.5 (c + d) * h_3] / 12$$

were a, b, c and d are the mean values of the 15D score at different measurement points (a = baseline, b = at admission, c = three, and d = 12 months after the surgery) and  $h_1$  is the mean waiting time in months,  $h_2$  is the time from admission (three months) and  $h_3$  is the total follow-up time (in SWT 14 months, and in NfWT nine months, up to 20 months in both groups).

### 7.5.2 Disease-specific outcome measures

Disease-specific quality of life instruments contain a list of symptoms relevant to the condition under study. One of the disease-specific instruments in hip patients is the Harris Hip Score (HHS). In this study, HHS was modified and used to measure hip pain and function as a supplement to generic 15D. The self-report HHS consists of two sections: pain (0–44 points) and functional activities of daily living and gait (0–47). The total score ranges from 0 to 91, with higher scores representing better health states (Harris 1969). A further 9 points of the total Harris Hip Score would normally be assigned to describing deformity and range of motion, but were excluded because these items could not be measured in a patient self-report questionnaire. The performance of self-reported HHS is comparable to that of surgeon-assessed HHS and has shown to be less burdensome to patients than a physician-administered HHS (Mahomed et al. 2001).

For knee patients, the self-report Knee Society Clinical Rating System (KSS) was chosen. KSS has been shown to be more responsive as a measurement of standardized response in pain and function domains than WOMAC and SF-36 (Lingard et al. 2001). In this study, KSS was used as modified to measure knee pain and physical function (Insall et al. 1989; Lingard et al. 2001). Functional performances were assessed using walking distance and stair climbing. The pain score (0–50 points) and function score (0–100 points) are presented separately; clinical dimensions (range of motion, stability, flexion contracture, extension lag and alignment) was 30% of the

total score) could not be measured in a patients' self-administered questionnaire, and thus the total score could not be derived.

### 7.5.3 *Direct rating scale instrument*

To evaluate the patients' sense of well-being, a Visual Analogue Scale (VAS – a health-state rating scale) was used. The measurement consists of a line on a page with clearly defined endpoints. The distance between intervals on a VAS should reflect a person's understanding of the relative difference between the concepts being measured. VAS has been used in the context of health as a measure of symptoms and domain of health, and to provide a single index measure of HRQoL. The most preferred health state is placed at one end of the line and the least preferred at the other end (McDowell and Newell 1996; Drummond et al. 2005; Brazier et al. 2007). In this study, VAS was a horizontal line (100 mm = 100%) illustrating the total of the patients' health state deficiency at that moment. It was used to value the effect of arthritis on this deficiency of health. Patients were asked to mark on the line which part of the deficiency of health is due to arthritis. The higher the number was, the more powerful was the effect of arthritis. In addition, the patients' self-reported state of health was described on a five-point scale, representing health states from excellent to bad.

## 7.6 Direct cost measures

### 7.6.1 *Social and health service use*

It is a common assumption that patients having OA use more health and social services than the same-age population without OA. Data on the use of social and health services due to OA were collected using a self-administered questionnaire delivered to patients at the hospital and via the post. Participants were asked whether they had had home visits from a nurse, chiropodist or physician and whether they had used rehabilitation services within three months due to OA. In addition, patients were asked whether they had used any support services of home help, including regular or temporary meals-on-wheels, housework, laundry services, bathing or transportation.

In addition, patients were asked whether they had physician visits to a university hospital, central hospital, regional hospital, health care centre, private physician or occupational health care visits, and how many hospital days they have had due to OA. The use of services in the previous three months due to OA was measured when placed on the waiting list, on admission, three months and one year after operation.

The costs related to TJR consist of the use of hospital care costs (total costs of hospital services, including TJR, paid by purchaser). The total direct costs during the waiting time and after the surgery include the following items: outpatient visits (doctor, nurse, and chiropodist), costs of the surgery including radiology, laboratory services, hospital days and rehabilitation services. The use and costs of regular social services

due to OA including meal-on-wheels, home help, laundry services, bathing services and transportation during the waiting time were included.

The mean number of service use was calculated and the mean costs of use were calculated. Costs of social and health services were valued in Finnish unit costs in 2006 (Hujanen et al. 2008.) The costs of TJR were obtained from FHDR (HILMO) and cost data were linked to the collected patients' data.

### 7.6.2 Medication use

The goal of medication to OA is to reduce the symptoms, namely joint pain, stiffness, and swelling. Drugs are also used postoperatively as pain killers. To get this medication information, we asked patients to report the name and dosage of the medication, which they used due to OA. Because of the circumstances mentioned before, reporting was very heterogeneous. However, the main stream of medication was observed. The unit costs of medication per tablet were obtained from CD-Pharmacy (Pharmaceutical Information Centre in Finland 2004). In this study we did not separate different medication groups related to OA; disease-specific medication (DSM) was used as a keyword. The costs of DSM during the observation period were calculated as product use per week. The calculations were made at each of the four measurement points.

## 7.7 Statistical analysis

In this study the sample size estimate was based on the primary outcome variable of the 15D score. A subgroup of 177 patients would provide the 80% power (two-tailed  $\alpha$  error 5%) to detect a clinically important difference ( $\Delta 0.03$ ) in the 15D score between the randomized groups (Sintonen 1994b). At baseline (when placed on the waiting list), descriptive statistics were used to describe the socio-demographic and clinical characteristics and differences in the co-morbidity groups and randomized patient groups. Comparative variance analyses of socio-demographic and clinical characteristics were carried out using either an independent samples t-test or a Chi-square test depending on whether the variable was on a continuous or a nominal scale. The differences on each of the 15D dimensions and in the overall 15D single index score for patients were calculated.

Primary analyses were conducted with an ITT principle (Mealli and Li 2011) so that patients were followed in the groups to which they were randomly allocated. Also a supplementary per-protocol analysis was carried out first at admission by excluding the patients in SWT group who were admitted beyond the short waiting time (waiting time > three months), and secondly when the cost-utility of waiting time was evaluated with patients who were treated according to the time limits of the group to which they were randomized.

When comparing the effect of waiting time on the mean scores of 15D, total HHS, pain, function in knee patients and costs of DSM at the various points of follow-up between SWT and NfWT, a general linear model for repeated measures was used. Univariate (tests of within-subject effects) p-values were calculated, while a p-value < 0.05 was considered statistically significant. The results for estimated effects were given as mean differences, standard deviations and 95% confidence intervals (CI) for the score of 15D, total HHS, pain, function and costs of DSM.

Missing data on the 15D questionnaire were predicted by means of a regression model with the patient's responses for other dimensions, and also with other data from the patients, as age and gender as explanatory variables (Sintonen 2001). The missing data were replaced if a minimum of 80% of the dimensions had been completed. Furthermore, the incidence of co-morbid diagnoses and also the mean 15D score were calculated for each co-morbidity group in the model.

To assess the degree of uncertainty in the results of cost-utility analyses, probabilistic sensitivity analyses (bootstrapping with 1,000 replicates) were carried out. Results are given as mean incremental costs and effects with their 95% confidence intervals, incremental cost-effectiveness ratio (ICER), cost-effectiveness plane (CE plane) and cost-effectiveness acceptability curve (CEAC).

The CE plane was presented to show how decisions can be related to both costs and QALYs. The plane is divided into four quadrants indicating four possible situations relating to additional costs and additional QALYs associated with the SWT compared with the NfWT (Figures 13 and 14). When one group is clearly less costly and more effective than the other, it is said to be dominant (Brazier et al. 2007).

The CEAC was presented to estimate confidence intervals around ICERs to represent the uncertainty of cost-utility results. The CEAC represents the probability that an intervention (here, one waiting time group) is cost-effective at each value of willingness to pay (WTP) for a QALY. Stochastic analysis by bootstrapping was applied to generate a distribution of costs and QALYs for both interventions. CEACs provided a graphical representation of the probability that a particular intervention was cost-effective over a range of maximum WTP for a QALY value (Brazier et al. 2007).

All analyses were carried out from the Finnish societal perspective, excluding production losses and value-added taxes. While the longest observation period was less than two years, no discounting was applied. The data analyses were performed using SPSS versions 14 and 16 for Windows<sup>1</sup>.

---

1 SPSS Inc., Chicago, IL, USA

## 7.8 Ethical considerations

Each patient provided a written informed consent. The study protocol was approved by the HUCH Surgery Ethics Committee (registration number 134/E6/02). Surgeons were unaware of the assigned intervention. For ethical reasons, double blinding was not possible. The trial was registered in the U.S. National Institutes of Health (NIH) ClinicalTrials.gov Register<sup>2</sup> under trial number NCT00294424.

---

<sup>2</sup> See <http://www.clinicaltrials.gov>.

## 8 RESULTS

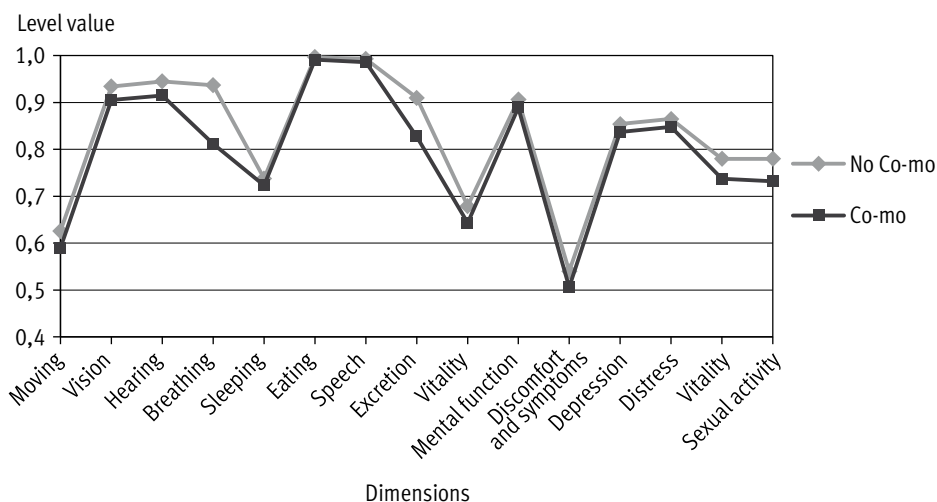
### 8.1 Effect of co-morbidity on HRQoL in patients waiting for total joint replacement (I)

The effect of co-morbidity was assessed as the difference in HRQoL between the patients (n = 893) with and without co-morbidity. Of the patients, 73% had one or more co-morbidities. At baseline, the mean 15D scores in patients with and without co-morbidity were 0.778 and 0.816, respectively. The difference in the score ( $\Delta 0.038$ ) was clinically important and statistically significant ( $p < 0.001$ ). Of the dimensions related to OA, patients with co-morbidity scored lower on the dimensions of moving ( $\Delta 0.036$ ,  $p < 0.001$ ), vitality ( $\Delta 0.052$ ,  $p < 0.001$ ) and sexual activity ( $\Delta 0.051$ ,  $p < 0.008$ ). The deterioration of HRQoL was significantly associated with co-morbidity (Figure 6). Among knee patients, there were more patients with co-morbidities (79%) than among hip patients (67%) ( $p < 0.001$ ).

The self-assessed health status of patients with co-morbidity, measured on a five-point scale, was worse than in the group without co-morbidity ( $p < 0.001$ ). The effect of OA on the health-state deficiency as measured by VAS was 62% among patients with co-morbidity versus 76% ( $p < 0.001$ ) among patients without co-morbidity.

In this analysis, severity of OA and co-morbidity were the main observational criteria of patients who were placed on the waiting list. In this analysis, we found that the HRQoL of all patients waiting for surgery was poor, but significantly worse in patients with co-morbidity.

**Figure 6.** The 15D profiles (mean values) of patients with (Co-mo) and without co-morbidity (No Co-mo), at baseline.



## 8.2 Characteristics of randomized groups in hip and knee patients and those who were lost to follow-up during the study (II, III)

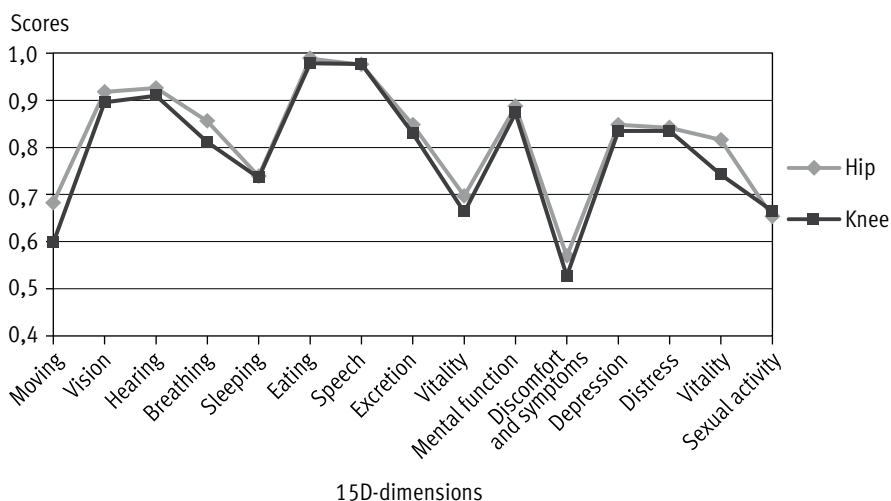
At baseline, after providing informed consent, 833 patients in three research hospitals were randomly allocated to either the SWT (n = 346) or NfWT (n = 487) group. Of the 833 randomized patients, 24 did not return the questionnaire at baseline, although they had signed informed consent forms and had been randomized. The only available information on these dropouts was gender and age. The mean age in dropouts was 68 years in the SWT group (n = 12) and 72 years in the NfWT group (n = 12); 75% of these were women. Of the remaining 809 patients, the mean age in both randomized groups was 66 years; and of these, 65% and 61% (respectively) were women.

Of the patients, 162 (20%) dropped out after randomization at different stages of the follow-up for various reasons. Among these patients the mean age was 68 years and 62% of these patients were women. The only statistically significant differences in the baseline characteristics between the dropouts and those who remained in the study to the end of follow-up were found in the mean age (dropouts slightly older) and in the proportion of living alone (slightly higher among dropouts).

These lost to follow-ups were not included in the final analyses. The baseline characteristics and comparison of the randomized groups and those who were lost to follow-up are presented in Table 6.

At baseline in both hip and knee patients' groups, the most significant 15D dimensions affecting patients' everyday life before TJR were moving, sleeping, vitality, discomfort and symptoms, and sexual activity (Figure 7).

**Figure 7.** The 15D profile of the hip and knee patients placed on the waiting list for TJR.





**Table 6.** Baseline characteristics in patients waiting for Total Joint Replacement and those lost to follow-up.

Characteristics at baseline	SWT n = 268	NfWT n = 379	Lost n = 162	P-value <sup>a</sup>	P-value <sup>b</sup>
Age, years (mean ± SD)	66 (9.5)	66 (9.7)	68 (9.9)	0.687	0.023*
Females [n, (%)]	185 (65.1)	231 (60.5)	89 (62.2)	0.152	0.538
Housing [n, (%)]					
Living alone	95 (33.9)	104 (28.1)	59 (41.3)	0.066	0.010*
Basic education [n, (%)]					
Lower level	227 (81.1)	309 (83.3)	119 (83.8)	0.264	0.390
Co-morbidity, yes [n, (%)]	207 (73.9)	270 (72.9)	108 (75.5)	0.406	0.330
Medication to arthritis, yes [n, (%)]	248 (88.6)	329 (88.7)	124 (86.7)	0.531	0.302
Medication to co-morbidity yes [n, (%)]	215 (76.8)	273 (73.6)	112 (78.3)	0.200	0.232
Hip patients [n, (%)]	145 (51.1)	169 (47.1)	70 (41.3)	0.126	0.520
HHS (mean ± SD) <sup>c</sup>	43.9 (13.5)	44.09 (14.4)	44.3 (12.9)	0.239	0.853
Knee					
Pain score (mean ± SD) <sup>d</sup>	19.9(11.8)	20.9 (12.4)	19.96 (11.0)	0.355	0.030*
Function score (mean ± SD) <sup>e</sup>	48.4 (22.1)	46.9 (23.3)	52.4 (22.6)	0.567	0.171
15D score (mean ± SD) <sup>f</sup>	0.765 (0.105)	0.774 (0.102)	0.761 (0.120)	0.222	0.360
BMI (mean ± SD) <sup>g</sup>	28.8 (5.7)	28.7 (5.2)	27.7 (7.1)	0.809	0.099
VAS (mean ± SD) <sup>h</sup>	67 (26)	65 (27)	62 (28)	0.540	0.089

\*p < 0.05.

<sup>a</sup> Difference between the randomized groups.

<sup>b</sup> Difference between the patients who remained in the study to the end of follow-up and those lost to follow-up.

<sup>c</sup> HHS (scale 0 = worst, 91 = best).

<sup>d</sup> KSS pain (score 0 = worst, 50 = best).

<sup>e</sup> KSS function (scale 0 = worst, 100 = best).

<sup>f</sup> 15D score (scale 0 = worst, 1 = best).

<sup>g</sup> Body mass index (kg/m<sup>2</sup>).

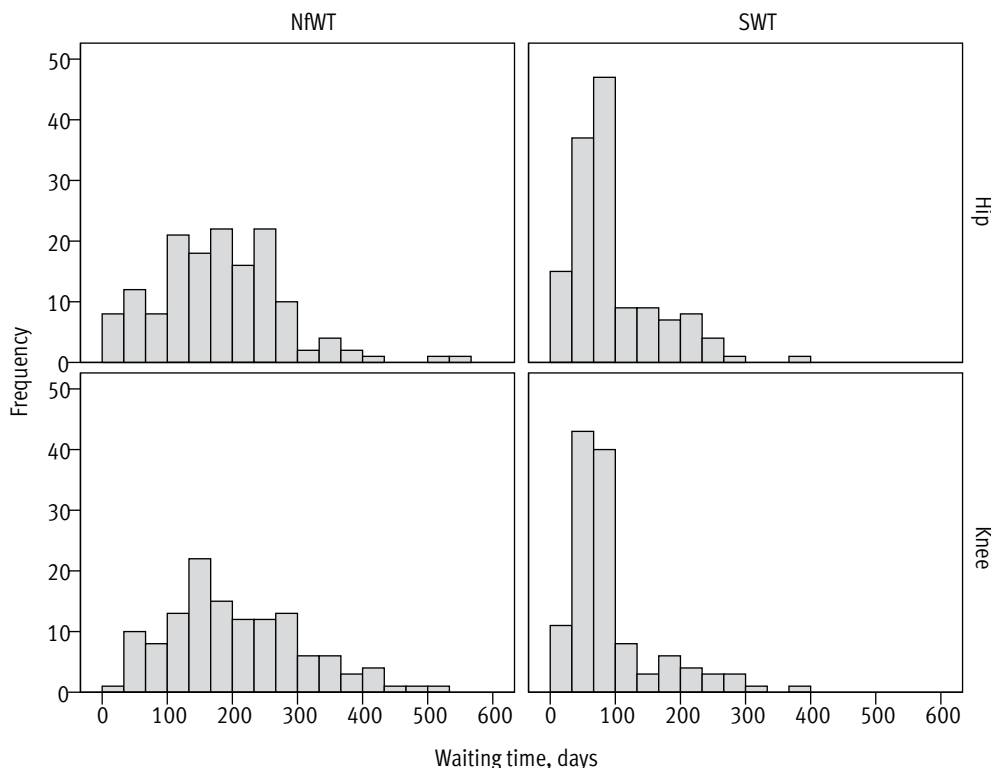
<sup>h</sup> Visual Analogy Scale (0 to 100).

### 8.2.1 Waiting times in the randomized groups (II, III)

There were statistically significant differences in mean waiting times between both randomized groups in hip and knee patients. The mean waiting time in the SWT hip patients was 74 (SD 145) days, and in NfWT it was 194 (SD 175) days. Thus difference was 120 (95% CI, 93–143) days. Similarly in knee patients the mean waiting times for SWT and NfWT groups were 95 (SD 81) and 239 (SD 135) days. Thus the difference was 144 (95% CI, 155–188) days, respectively. In both hip and knee groups there were patients in the SWT group who waited longer than three months. There

were also patients in both NfWT groups (26 patients (10%) in the hip group and 18 (7%) in the knee group) who were operated on within three months after having been placed on the waiting list. On the other hand, 71.2% in the SWT group were operated on within three months, and in the NfWT group, 79 (22.3%) waited longer than one year. (Figure 8.)

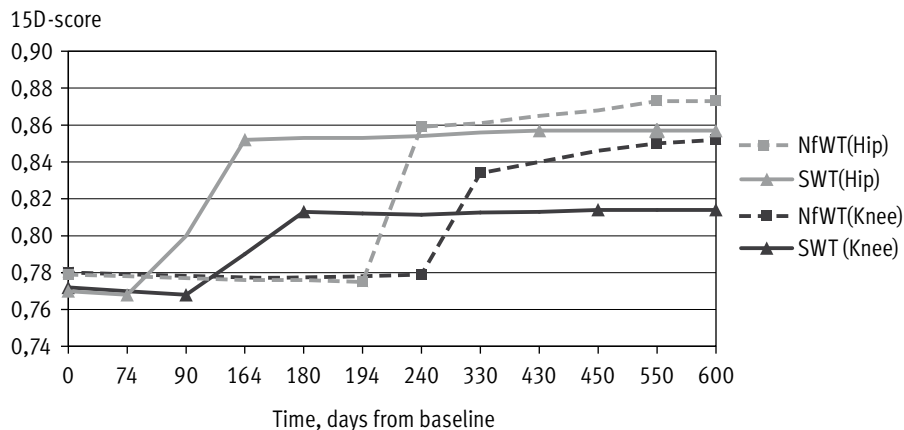
**Figure 8.** The distribution of waiting times in the randomized groups (Horizontally: 0 = NfWT, 1 = SWT, Vertically: 1 = hip, 2 = knee).



8.2.2 Effect of waiting time on health outcomes (II, III)

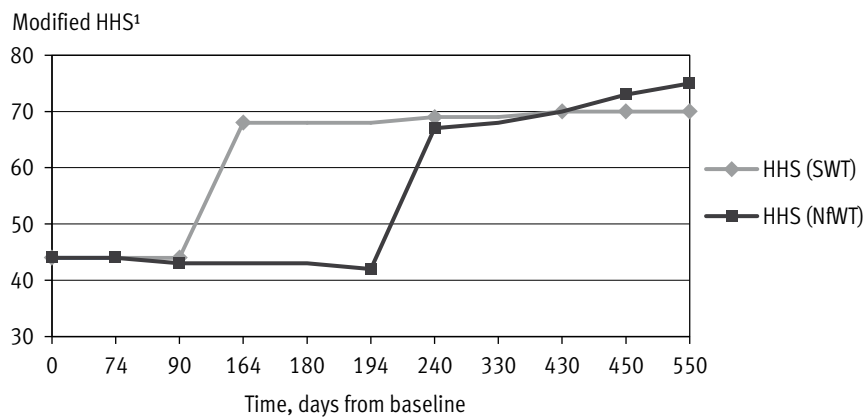
The length of waiting time alone did not affect the health outcomes in both groups. The mean 15D score at admission in the hip patients’ SWT group was 0.768 and in the NfWT group it was 0.769 (95% CI for mean difference: from -0.022 to 0.024). In knee patients, the mean 15D score at admission in the SWT group was 0.768 and in the NfWT group it was 0.779 (95% CI for mean difference: from -0.020 to 0.029). (Figure 9.)

**Figure 9.** The mean 15D scores in the randomized hip and knee patients during the observation time.



In hip patients, the mean self-reported HHS score in the SWT group was 42.99 and in the NfWT group it was 41.73 (95% CI for mean difference: from -5.233 to 1.816) (Figure 10).

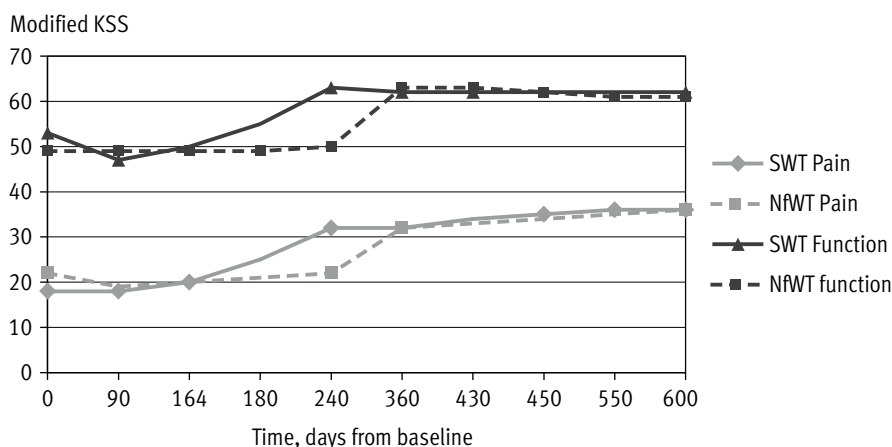
**Figure 10.** The mean Modified HHS during the whole observation period.



HHS<sup>1</sup>, score 0–91 (from worst to best).

The knee-specific pain and function scores were reported separately. At baseline, the mean pain score in the SWT patients was 18.93, and in the NfWT patients it was 22.79 (95% CI for mean difference: from  $-1.478$  to  $3.969$ ) and the mean function scores were 48.42 and 46.95 (95% CI for mean difference: from  $-6.659$  to  $3.493$ ), respectively (Figure 11).

**Figure 11.** The mean Modified KSS (pain and function) during the whole observation period.



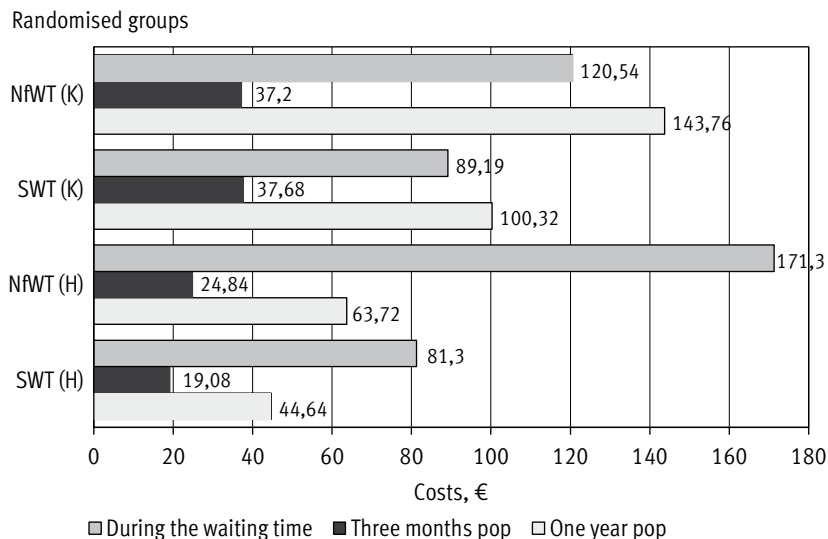
KSS modified pain and function score (pain: 0–50, function 0–100).

During the waiting time there was no significant deterioration in the mean pain or function scores. There were improvements from admission to 12 months postoperatively in the mean 15D score and in the pain and function scores in all groups.

### 8.2.3 Costs of disease-specific medication (II, III)

The mean weekly costs of disease-specific medication (DSM) between the SWT and NfWT hip patients' groups were EUR 5.56 and EUR 5.63, respectively. The total medication costs during the waiting time were significantly higher in the NfWT group. In knee patients, the mean costs of disease-specific medication during the waiting time in the SWT group were EUR 5.33 and in the NfWT group they were EUR 3.57 in a week. The difference is statistically significant ( $p = 0.029$ ). However, there was no statistically significant difference in the mean medication costs during the whole waiting time period. The mean total medication costs increased again during the year after TJR in both randomized groups, being highest in the knee patients' NfWT group (EUR 143) (Figure 12).

**Figure 12.** The mean total costs of disease-specific medication in the randomized groups from baseline to one year postoperatively (H = hip, K = knee, SWT = short waiting time, NfWT = non fixed waiting time).



### 8.3 The cost-utility (IV)

Due to dropouts, the final cost-utility analyses are based on 550 (66%) of the randomized patients who completed all four questionnaires, with a mean (SD) age of 66 (9.9; range; from 33 to 89) years, of which 345 (63%) were women and an additional 243 (hip n = 130, knee n = 113) were in the SWT group and 307 (hip n = 149, knee n = 158) were in the NfWT group (Table 6, p. 47).

There were statistically significant differences in the knee patients’ mean 15D scores after the operation; those in the NfWT group had better HRQoL. The 15D scores at each measurement point are presented in Table 7 (p. 52).

In ITT analyses, the mean (SD) total costs of TJR among hip replacement patients were EUR 9,986 (3,540) in the SWT group and EUR 10,472 (4,686) in the NfWT group, and EUR 9,809 (4,085) and EUR 9,801 (3116) among knee patients, respectively (Table 8, p. 52).

During the equally long follow-up period, the SWT hip patients experienced on average 1.341 QALYs and the NfWT patients 1.327 QALYs. Furthermore, the SWT knee patients experienced on average 1.453 QALYs and the NfWT patients 1.467 QALYs (Table 9, p. 53).

In the ITT analyses of hip patients, the SWT group gained more QALYs at lower cost than patients in the NfWT group. In knee patients, the results were opposite. However, the difference in knee patients was insignificant.

**Table 7.** Mean 15D scores in hip and knee patients in the randomized groups.

Measurement point	Hip			Knee		
	SWT n = 130	NFWT n = 149	95% CI for mean difference	SWT n = 113	NFWT n = 158	95% CI for mean difference
At baseline	0.770	0.779	-0.036 to 0.026	0.772	0.779	-0.004 to 0.030
±SD	(0.090)	(0.100)		(0.18)	(0.12)	
At admission	0.772	0.771	-0.023 to 0.021	0.778	0.786	-0.008 to 0.014
±SD	(0.090)	(0.095)		(0.105)	(0.125)	
Three months pop	0.945	0.854	-0.277 to 0.094	0.811	0.842	0.033 to 0.058*
±SD	(1.063)	(0.105)		(0.12)	(0.11)	
One year pop	0.854	0.904	-0.033 to 0.133	0.823	0.852	0.019 to 0.035*
±SD	(0.116)	(0.503)		(0.14)	(0.10)	

\*p < 0.01.

**Table 8.** Mean costs of health care utilization; differences between the randomized groups in hip and knee patients.

Measurement periods	Hip			Knee		
	SWT (n = 130) Mean costs (€)	NFWT (n = 149) Mean costs (€)	95% CI for mean difference	SWT (n = 113) Mean costs (€)	NFWT (n = 158) Mean costs (€)	95% CI for mean difference
During the waiting time	441	422	-236 to 198	627	463	-502 to 174
±SD	(748)	(1,081)		(1491)	(1,232)	
From admission to 3 months postoperatively	8,015	8,274	-280 to 798	7,827	8,031	-168 to 576
±SD	(2,288)	(2,243)		(1451)	(1,613)	
From 3 months to 12 months postoperatively	1,591	1,886	-477 to 1,068	1,561	1,309	-938 to 434
±SD	(2,405)	(4,036)		(3,121)	(2,344)	
Mean Total costs	9,986	10,472	-486 to 1,458	9,809	9,801	-909 to 894
±SD	(3,540)	(4,686)		(4,085)	(3,116)	

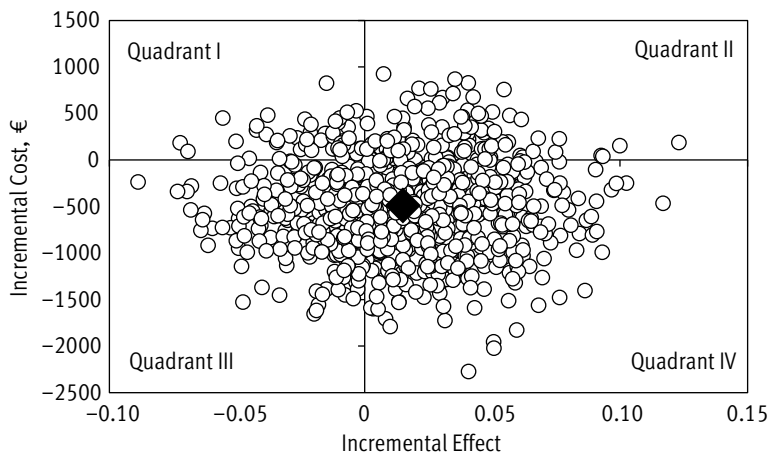
**Table 9.** The results of cost-utility analyses of total joint replacement in hip and knee patients (ITT analyses).

Strategy	Cost C (€)	Incremental costs ΔC	Utility U (QALYs)	Incremental utility ΔU	Cost-utility C/U	ICUR ΔC/ΔU
SWT (hip)	9,986		1.3414		7,444.46	
NfWT (hip)	10,474	488	1.3265	-0.0149	7,895.97	32,751
SWT (knee)	9,809		1.4528		6,751.79	
NfWT (knee)	9,802	-7	1.4669	0.0141	6,682.12	496

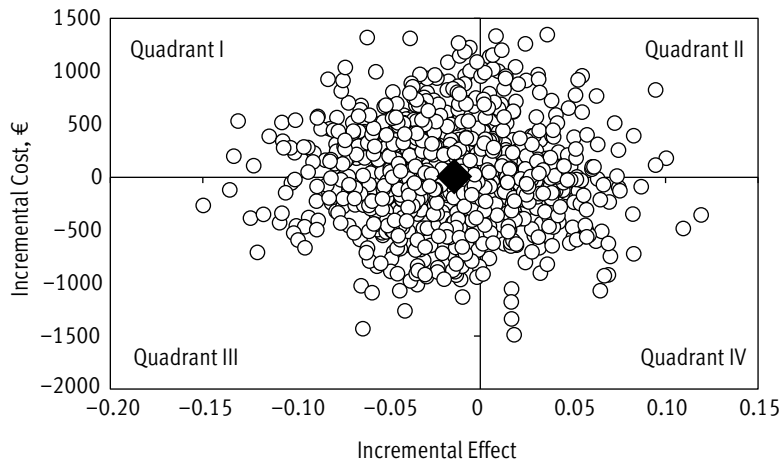
### 8.3.1 Sensitivity analysis

The point estimates in the ITT analysis above thus suggest strong dominance for SWT among hip patients. On the basis of probabilistic sensitivity analyses in hip patients, the 95% CI for mean difference in QALYs was from -0.048 to 0.076 and in costs from EUR -1,453 to EUR 464. In knee patients, the 95% CI for mean difference in QALYs was from -0.095 to 0.063, and in costs from EUR -913 to EUR 955 (Figures 13 and 14).

**Figure 13.** Cost-effectiveness plane in hip replacement (diamond stands for the base case result). In 59.5% of simulated cases, SWT was both less costly and more effective (quadrant IV); in 25.4% less costly and less effective (quadrant III); in 10.5% more costly and more effective (quadrant II); in 4.6% more costly and less effective (quadrant I).

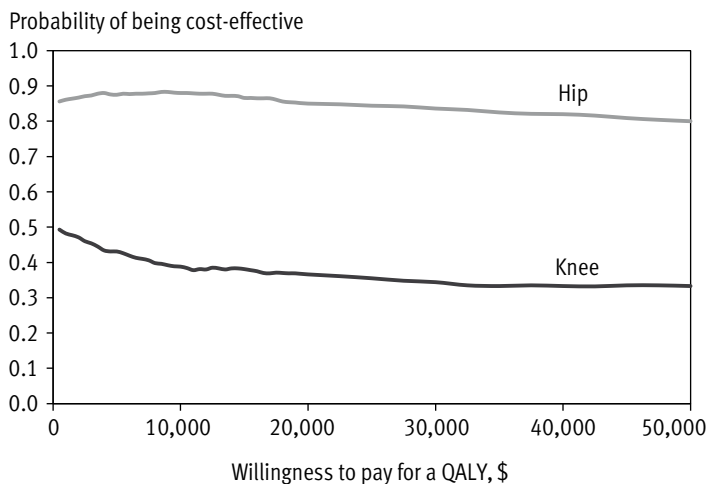


**Figure 14.** Cost-effectiveness plane in knee replacement (diamond stands for the base case result). In 17.5% of simulated cases, SWT was both less costly and more effective (quadrant IV); in 31.7% less costly and less effective (quadrant III); in 16.8% more costly and more effective (quadrant II); in 34.0% more costly and less effective (quadrant I).



If the willingness to pay for a QALY is EUR 20,000, the probability of SWT being cost-effective is about 85% in hip patients and about 40% in knee patients (Figure 15).

**Figure 15.** Cost-effectiveness acceptability curves for the short waiting time in hip and knee replacement.





Because there were patients in both randomized groups who did not follow the main protocol, we carried out a secondary per-protocol analysis. As a result, the mean total cost among hip patients was EUR 10,302 ( $\pm 3788$ ) in the SWT group and EUR 10,402 ( $\pm 4854$ ) in the NfWT group, and EUR 9,374 ( $\pm 3,259$ ) and EUR 9,904 ( $\pm 3,115$ ) among knee patients, respectively. During the equally long follow-up period, the SWT hip patients experienced, on average, 1.3536 QALYs and the NfWT patients 1.3879 QALYs. Correspondingly, the SWT knee patients experienced, on average, 1.4428 QALYs and the NfWT patients 1.5022 QALYs. Point estimates thus suggest an ICUR of EUR 2,941 for NfWT among hip patients and of EUR 8,983 among knee patients. This analysis seemed to add to the uncertainty over whether the difference in waiting time actually makes much difference in terms of cost-utility (Table 10).

**Table 10.** Per-protocol analysis of cost-utility of Total Joint Replacement in hip and knee patients.

Strategy	Costs C (€)	Incremental costs $\Delta C$	Utility (QALYs)	Incremental utility $\Delta QALYS$	Cost-utility C/QALYs	ICUR $\Delta C/\Delta QALYS$
Hip(SWT)	10,302		1,354		7,608.57	
Hip (NfWT)	10,402	-100	1,388	-0.034	7,494.24	2,941
Knee (SWT)	9,374		1,443		6,496.19	
Knee (NfWT)	9,904	-530	1,502	-0.059	6,593.87	8,983

According to the present study, there does not seem to be a significant difference in the cost-utility of short and longer waiting times for TJR, at least given the waiting time difference between our study groups.

## 9 DISCUSSION

Osteoarthritis is the most common cause of musculoskeletal disability and pain in the world. Furthermore, as a result of the prevalence of OA, the number of patients on waiting lists and the waiting times for elective total joint replacements are increasing. There are studies that have evaluated the effect of waiting time on HRQoL and other outcomes of TJR, but few of these have evaluated the effect of waiting time on these outcomes in two different waiting time groups. The main aim of the present study was to evaluate the effect of different lengths of waiting time on HRQoL and other health outcomes before and after TJR. In the study, patients were randomized into short and non-fixed waiting time groups and followed on the basis of the ITT principle. This study was conducted during the time period when waiting times for elective surgery were long and the share of patients having waited for over 6 months of all operated patients were highest in knee replacements (41%) and in hip replacements (32%) (Mikkola et al. 2008).

### 9.1 HRQoL in patients waiting for Total Joint Replacement

According to earlier studies, patients waiting for total joint replacement have poor quality of life and progressive pain and reduction in physical functioning (Derrett et al. 1999; Kelly et al. 2001; Croft et al. 2002; Ackerman et al. 2005; Hirvonen et al. 2006; McHugh et al. 2008). However, a systematic review of the impact of waiting time for TJR on pain and functional status concluded that there was strong evidence that pain (in hip and knee OA) and self-reported functioning (in hip OA) did not deteriorate during a waiting time less than 180 days (Hoogeboom et al. 2009). The impact of waiting time was mostly seen as deterioration in disease-specific measurements. According to Hoogeboom et al. (2009), there is conflicting evidence for change in self-rated functioning in patients with knee OA, and indefinite results were reported for long waiting times (> 6 months), though pain might increase in hip OA patients. In our study, the mean total scores of HRQoL instruments did not deteriorate while waiting. Furthermore, the mean modified KS score in the NfWT group was better at admission than at baseline. These findings are in line with earlier studies. However, some studies have shown opposite findings. For example Bachrach-Lindström et al. (2008), found that a long waiting time (> 6 months) is detrimental to patients' HRQoL, causing pain, reduced functional condition, and an increased need for support from relatives, thus limiting independence in daily life (Bachrach-Lindström et al. 2008).

Why did the HRQoL not deteriorate during the waiting time? One explanation might be that when patients are placed on the waiting list, the awareness of upcoming surgery might bring psychological relief, and at the same time, pain and functional ability are managed with appropriate medication during the waiting time. This was seen as an increasing use of DSM during the waiting time in all patient groups. According to Conner-Spady et al. (2007) patients' view of waiting times are not only related to quality of life, but also to prior expectations and notions of fairness and priority. Ac-

According to earlier literature reviews, no other studies have been published in which the waiting time effect was tested by randomizing the patients into two different waiting time groups at baseline.

## 9.2 Baseline co-morbidity and HRQoL

In Finland, 44% of the working population and 82% of senior citizens have at least one chronic disease. Some studies have suggested that the use of a generic HRQoL instrument in the studies of OA, where co-morbidity is common, would be useful in characterizing the global burden of this disease (Piccorillo 2000; Salaffi et al. 2005). The findings of the present study have shown that OA patients with co-morbidity have poorer HRQoL than patients without co-morbidity. However, those patients without co-morbidity seem to rate their burden of OA, when measured by VAS, even higher than OA patients with co-morbidity before surgery. Assessing the co-morbidity condition and HRQoL at baseline or even earlier in time, when patients are waiting with a referral for access to a waiting list, the severity of co-morbidity might operate as an instrument to help in prioritization in medical decision-making for health care delivery.

## 9.3 Randomization and waiting time

This study is the first clinical trial where patients were randomly allocated to different waiting time groups when placed on the waiting list for elective surgery. Randomization was based on a flowchart of four phases: enrolment, intervention allocation, follow-up, and data analysis of a parallel randomized trial of two groups modified from the CONSORT Statement 2010 (Schulz and Grimes 2002). The patients were randomized to different waiting time groups (SWT and NfWT) at baseline and were followed according to the ITT principle. The patients awaiting TJR were prospectively followed from the time of being placed on the waiting list to admission, with waiting time recorded precisely, and further for a year postoperatively, providing evidence of the effect of waiting time on pre- and postoperative health status. Randomization was successfully completed, and the groups did not differ from each other at baseline. Further, the sample size estimate was based on the primary outcome variable, the 15D score.

## 9.4 Health outcome measures and costs of osteoarthritis

Both generic and disease-specific instruments were used. The generic 15D has been used successfully in earlier studies dealing with hip and knee replacement and thus facilitates a comparison of the pre-surgery scores in these studies (Rissanen et al. 1995, 1996 and 1997; Räsänen et al. 2005, 2006 and 2007). Earlier research has shown that in most of the important properties (reliability, content validity, sensitivity in terms of discriminatory power, and responsiveness to change), the 15D instrument compares at least equally with other similar instruments that produce a valuation-

based single index number. And recent research has since confirmed that, especially in the rehabilitation of musculoskeletal disorders, the 15D instrument was at least as responsive as the SF-6D and much more responsive than the EQ-5D. Furthermore, the scores of 15D are suitable for QALY calculation. Both HHS in hip patients and KSS in knee patients are well-known, standardized and widely used disease-specific instruments to measure both pain and functioning in OA patients. We used these instruments as modified. In an earlier study by Mahomed et al. (2001), a comparison between self-report HHS and surgeon-assessed HHS was made. They concluded that the measurements showed excellent concordance and supported the use of self-reported HHS (Mahomed et al. 2001). The disease-specific, self-reported modified KSS was used to measure only knee pain and physical function, and it has been found to be responsive to these changes in the study by Lingard et al. (2001).

The utilization and costs of health care and social services were surprisingly low during the whole observation period, which is against common assumptions, though there is some scientific evidence concerning costs related to OA during the waiting time. According to the study of Quan et al. (2002), waiting for a surgery may have an independent economic effect. If a delay for surgery equates with failed problem resolution, then it is reasonable to presume that additional health services will be required until the definitive surgical service is performed and it may increase overall health service expenditures. However, Quan et al. (2002) concluded that no evidence was found to suggest that waiting for common surgical procedures is correlated with higher health service expenditures pre- or postoperatively. They found that waiting time is not a proxy for health service use and health service costs do not decrease markedly after surgery (Quan et al. 2002). Our findings support these results, as patients with OA did not use health care services significantly more than the general population of the same age. However, Fielden et al. (2005) suggested that longer waits (> 5 months) incur higher economic costs while waiting, and March et al. (2002) found that poorer pre-surgery health status predicts higher expenditures during the first postoperative year.

Some uncertainty was embedded in the results of utilization and costs of health care and social services. There might be several reasons for this. Firstly, the questionnaire was self-reported and the questions were presented three months retrospectively. Secondly, elderly people may under- or overestimate the utilization; the data showed that the responses were partly incomplete, and some patients answered that they had utilized services but did not mention the number of visits or what kind of services were used. Thirdly, we evaluated only the services and costs related to OA. These things might lead to underestimation of health care and social services.

In the present study, the mean length of a hospital stay was 2 days for hip patients and 3 days for knee patients. According to the TJR statistics, in Finland the mean costs of surgery between the years 2002 and 2003 was EUR 7,421 and costs including hospital days were EUR 7,643: EUR 7,462 for hip patients and EUR 7,729 for knee patients (Remes et al. 2007). In this study the costs of surgery were drawn from the Finnish Hospital Discharge Register (FHDR); the mean cost for THR was EUR 7,852, and for

TKR it was EUR 7,655. The other costs related to OA were calculated by multiplying the Finnish unit costs by the amount of service use as reported by patients.

### 9.5 Cost-utility of Total Joint Replacement

Earlier non-randomized studies have found that total joint replacements of hip and knee are effective and cost-effective interventions (Bourne 1996; Rissanen et al. 1997; Segal et al. 2004; Räsänen et al. 2006; Navarro Espigares and Torres 2008; Xie et al. 2010; Higashi and Barendregt 2011). However, according to several studies, knee patients gain more QALYs at lower cost than hip patients (Segal et al. 2004; Navarro Espigares and Torres 2008). These findings are partly opposite to ours. According to the results of the ITT analyses of this randomized study, a waiting time  $\leq 3$  months might have a positive effect on the cost-utility of THR. But in knee patients, no evidence was found that patients in SWT would benefit more than those waiting over 3 months.

### 9.6 Limitations of the study

Dropout is a potential weakness of an RCT; that was also the case in this study. Firstly, 235 of the eligible patients refused to participate in the study and 162 were lost in the follow-up during the study. We tested whether those ( $n = 162$ ) who dropped out after randomization at any stage of the follow-up differed in the baseline characteristics from those who remained in the study to the very end. The only statistically significant differences were found in the mean age (dropouts slightly older) and in the proportion of living alone (among dropouts slightly higher). Secondly, 126 patients in the SWT group waited more than three months; and there were patients in the NfWT group who were operated within three months after having been placed on the waiting list. However, the primary analysis was based on the ITT principle to avoid bias associated with a non-random loss of participants. In addition, per-protocol analyses were performed as supplementary analyses. In the SWT group, those compliant with the allocated waiting time of three months (and all of the patients in the NfWT group were included in the analysis) supported the main findings and did not show a statistically or clinically important difference in HRQoL between randomized groups.

Thirdly, establishing comparable QALYs and costs between the SWT and NfWT groups did include challenges. As a solution, the mean follow-up time was set the same in both groups. However, we do not know exactly how the HRQoL (and costs) in the SWT group developed during the time from the last measurement in that group to the final measurement in the NfWT group. The HRQoL may have deteriorated slightly due to ageing or co-morbidity. But because the mean time difference between the last measurements in the groups was only 4–8 months, the change would probably be negligible, and therefore our assumption of no change may be justified. On the other hand, had the SWT group incurred further costs contrary to our assumption, its total cost would have been underestimated. As these changes would probably be marginal,

they might not have affected our conclusions. However, there is a lot of variance around the point estimates, and the differences in costs and QALYs between the per-protocol groups were not statistically significant in hip or knee patients.

Furthermore, data on medication were obtained from patients' self-reports, and there was no distinction between self-care and prescription medication. For temporary medication, we used mean dosages. Due to these factors, the differences between the randomized groups may have been underestimated.

### 9.7 Reliability and validity of the study

In this study, we focused on OA of the hip and knee joint, which is the most common cause of musculoskeletal disability and pain in the world. It is a chronic disease affecting patients' quality of life, and its incidence and prevalence will increase in the future due to the ageing of the population. This might lead to increasing volumes of TJR and further lengthening of waiting times. A randomized controlled trial (RCT) is a type of scientific experiment commonly used in testing the efficacy or effectiveness of health care services or health care technologies. For health care interventions such as new therapies or pharmaceuticals, many countries have formal requirements for provision of safety and efficacy data prior to product licensing (Schulz and Grimes 2002). RCTs are not common in surgery; however, we successfully completed the procedure to randomize patients into two different waiting time groups. All eligible patients had a chance to be recruited to the study and had the possibility to be allocated to the SWT group. We did not change the hospitals' routine procedure.

As far as HRQoL is concerned, we cannot be sure that there was no change in extraneous influences that might have occurred, such as an attitude change over time. This could lead to a difference in the responses provided. When a respondent answers a set of test items, the score obtained represents only a limited sample of behaviour (Crocker and Algina 1986). As a result, the scores may change due to some characteristic of the respondent, which may lead to errors of measurement. These kinds of errors will reduce the accuracy and consistency of the instrument and test scores (Golafshani 2003).

In this study, both generic and disease-specific instruments we used are referred to as stable and validated. If we are dealing with a repeatable measure, then the results should be similar in consecutive measurements if no change in the variable to be measured has occurred. In this study, the data were collected from four different hospitals, and the data collection was repeated four times during the study. This study was conducted during the time when the waiting times for elective surgery were long and the New Healthcare Reform was introduced in Finland. According to the reform, elective surgery must be provided within three months or at least within six months. The results of this study seem to suggest that longer waiting times, at least as long as those experienced in this study, do not affect HRQoL or other health outcomes related to TJR in hip and knee patients.

## 10 CONCLUSIONS AND SUGGESTIONS FOR FURTHER STUDIES

On the basis of the present study, the following conclusions can be drawn:

- 1) It is possible to perform randomization in a clinical trial, where patients are placed on a waiting list for a non-urgent surgical procedure. In this study, each patient had an equal possibility to be assigned by chance, rather than by choice, either to the short waiting time group or to the non-fixed waiting time group.
- 2) When patients are placed on the waiting list for TJR, they have deteriorated HRQoL, and the length of the waiting time, at least as experienced in this study, does not have a significant effect on HRQoL during the waiting time.
- 3) We do not have an accurate or credible answer to the question of an optimal length of the waiting time. According to the present study, the length of waiting time did not alone affect the cost-utility of the TJR, the severity of pain, or deterioration in functioning.
- 4) Understanding patients' views on waiting for surgery has implications for better management of waiting times for TJR. When implementing waiting time reform, health authorities need instruments for decision-making on optimizing the waiting times. Rapid access to care is not always related to better outcomes of the intervention, and consequently, the length of the waiting time should not categorically be the same for each patient.
- 5) Both generic and disease-specific outcome measures are needed to evaluate outcomes of before and after surgery in different patient groups.
- 6) Co-morbidity affects OA patients' HRQoL already when placed on the waiting list, and co-morbidity should be considered when the decision about surgery is made.

## REFERENCES

- Ackerman IN, Graves SE, Wicks IP, Bennell KL, Osborne RH. Severely compromised quality of life in women and those of lower socioeconomic status waiting for joint replacement surgery. *Arthritis Rheum* 2005; 53: 653–658.
- Adie S, Kwan A, Naylor JM, Harris IA, Mittal R. Cryotherapy following total knee replacement. *Cochrane Database of Systematic Reviews* 2012, Issue 9. Art. No.: CD007911. DOI: 10.1002/14651858.CD007911.pub2.
- Albin B, Hjelm K, Zhang WC. Healthcare systems in Sweden and China. Legal and formal organizational aspects. *Health Research Policy and Systems* 2010; 8: 20.
- American College of Rheumatology. Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). General Description. ACR. Retrieved 6 June 2012.
- Anell A. Swedish healthcare under pressure. *Health Econ* 2005; 14: S237–S254.
- Aromaa A, Koskinen S, eds. Health and functional capacity in Finland. Baseline results of the Health 2000 Health Examination Survey. Helsinki: National Public Health Institute, Publications of the National Public Health Institute B12, 2004.
- Atkinson MH. Osteoarthrosis. *Can Fam Physician*, 1984; 30: 1503–1507.
- Avouac J, Grossec L, Dougados M. Efficacy and safety of opioids for osteoarthritis. A meta-analysis of randomized controlled trials. *Osteoarthritis and Cartilage* 2007; 15: 957–965.
- Baar ME van, Dekker J, Oostendorp RA et al. Effectiveness of exercise in patients with osteoarthritis of hip or knee: nine months' follow up. *Ann Rheum Dis* 2001; 60: 1123–1130.
- Bachrach-Lindström M, Karlsson S, Petterson LG, Johansson T. Patients on the waiting list for total joint hip replacement. A 1 year follow-up study. *Scandinavian Caring Sciences* 2008; 22: 536–542.
- Barber J, Thompson S. Analysis and interpretation of cost data in randomised controlled trials. Review of published studies. *BMJ* 1998; 317 (7167): 1195–1200.
- Bourne RB. The cost-effectiveness of total joint replacement. The Canadian experience. *Knee* 1996, 3: 161–162.
- Bowden A, Fox-Rushby JA. A systematic and critical review of the process of translation and adaptation of generic health-related quality of life measures in Africa, Asia, Eastern Europe, the Middle East, South America. *Soc Sci Med* 2003; 57 (7): 1289–1306.
- Bowling A. *Measuring health. A review of quality of life measurement scales*. Buckingham: Open University Press, 2005.



Brauer CA, Rosen AB, Olchanski NV, Neuman PJ. Cost-utility analyses in orthopaedic surgery. *J Bone Joint Surg Am* 2005; 87 (6): 1253–1259.

Brazier J, Ratcliffe J, Salomon JA, Tsuchiya A. Measuring and valuing health benefits for economic evaluation. Oxford: Oxford University Press, 2007.

Brunenberg DE, Steyn MJ van, Sluimer JC, Bekebrede LL, Bulstra SK, Joore MA. Joint recovery programme versus usual care. An economic evaluation of a clinical pathway for joint replacement surgery. *Med Care* 2005; 43 (10): 1018–1026.

Conner-Spady BL, Johnston GL, Sanmartin C, John J McGurran JJ, Noseworthy TM. A bird can't fly on one wing. Patient views on waiting for hip and knee replacement surgery. *Health Expect*. 2007; 10 (2): 108–116

Crocker L, Algina J. Introduction to classical and modern test theory. Toronto: Holt, Rinehart and Winston, 1986.

Croft P, Lewis M, Wynn Jones C, Coggon D, Cooper C. Health status in patients awaiting hip replacement for osteoarthritis. *Rheumatology* .2002; 9: 1001–1007.

Current Care Quideliness. Helsinki: Finnish Medical Society Duodecim, 2007. Available at: <<http://www.kaypahoito.fi>>.

Derret S, Paul C, Morris J. Waiting for elective surgery. Effect on health-related quality of life. *Int J Qual Health Care* 1999; 11: 47–57.

Donaldson C, Baker R, Mason H et al. Methods for the economic evaluation of healthcare programmes. Oxford: Oxford University Press, 2005.

Drummond M, Schulpher M, Torrance G, O'Brien B, Stoddart G. Methods for the economic evaluation of healthcare programmes. Oxford: Oxford University Press, 2005.

Es PP van, Luijsterburg P, Dekker J et al. Cost-effectiveness of exercise therapy versus general practitioner care for osteoarthritis of the hip. Design of a randomized clinical trial. *BMC Musculoskelet Disord* 2011; 12: 232. doi: 10.1186/1471-2474-12-232.

Ethgen O, Bruyère O, Richy F, Dardenness 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): 363–374.

Fielden JM, Cumming JM, Horne JG, Devane PA, Slack A, Gallagher LM. Waiting for hip arthroplasty economic costs and health outcomes. *J Arthroplasty* 2005; 20: 8.

Finnish statistics on medicine 2006. Helsinki: National agency for medicine and Social Insurance Institution of Finland, 2007.

Finnish statistics on medicine 2010. Helsinki: National agency for medicine and Social Insurance Institution of Finland, 2011.

Flood J. The role of acetaminophen in the treatment of osteoarthritis. *Am J Manag Care*, 2010; 16 Suppl Management: 48–54.

Golafshani N. Understanding reliability and validity in qualitative research. *The Qualitative Report* 2003; 8 (4): 597–607. Available at: <<http://www.nova.edu/ssss/QR/QR8-4/golafshani.pdf>>.

Gold M, Siegel J, Russell L, Weinstein M. Cost-effectiveness in health and in medicine. New York, NY: Oxford University Press, 1996.

Gravelle H, Siciliani L. Is waiting time prioritization welfare improving? *Health Economics* 2008; 17 (2): 167–184.

Hagen TP, Kaarboe OK. The Norwegian hospital reform of 2002: The central state takes over ownership of public hospitals. *Health Policy*. 2006;76(3):320–333.

Hanning M, Winblad Spångberg U. Maximum waiting time. A threat to clinical freedom? Implementation of a policy to reduce waiting times. *Health Policy* 2000; 52: 15–32.

Harrison AJ, Appleby J. Optimising waiting. A view from the English National Health Service. *Health Econ Policy Law* 2010; 5: 397–409.

Harris W. Traumatic arthritis of hip after dislocation and acetabular fractures. Treatment by mold arthroplasty. *J Bone Joint Surg Am* 1969; 51: 4.

Hawker GA, Stewart L, French MR et al. Understanding the pain experience in hip and knee osteoarthritis. An OARSI/OMERACT initiative. *Osteoarthritis and Cartilage* 2008; 16: 415–422.

Hawthorne G, Richardson J, Day NA. A comparison of five Multi Attribute Utility Instruments. Centre for Health Program Evaluation. Working paper 140 presented to the Twenty Second Australian Conference of Health Economists 2000. Melbourne, 2000.

Hawthorne G, Richardson J, Day NA. A comparison of the assessment of quality of life (AQoL) with four other generic utility instruments. *Annals of Medicine* 2001; 33: 358–370

Heliövaara M, Mäkelä M, Sievers K et al. Tuki- ja liikuntaelinten sairaudet Suomessa. Helsinki: Kansaneläkelaitos, Julkaisuja AL:35, 1993.

Higashi H, Barendregt JJ. Cost-effectiveness of total hip and knee replacement for Australian population with osteoarthritis. Discrete-event simulation model. *PLoS* 2011; 6 (9): e25403.

Hirvonen J. Effect of waiting time on health outcomes and service utilization. A prospective randomized study on patients admitted to hospital for hip or knee replacement. Research report 170. Helsinki: Stakes, 2007.

Hirvonen J, Blom M, Tuominen U et al. Health-related quality of life in patients waiting for major joint replacement. A comparison between patients and population. *Health Qual Life Outcomes*. 2006; 4: 3.

Hirvonen J, Blom M, Tuominen U et al. Is longer waiting time associated with health and social services utilization before treatment? A randomized study. *J Health Serv Res Policy* 2007a; 12 (4): 209–214.

Hirvonen J, Blom M, Tuominen U et al. Evaluating waiting time effect on health outcomes at admission. A prospective randomized study on patients with osteoarthritis of the knee joint. *Journal of Evaluation in Clinical Practice* 2007b; 13: 728–733.

Hirvonen J, Blom M, Tuominen U et al. The effect of waiting time on health-related quality of life, pain and physical function in patients admitted to hospital for primary total hip replacement. A randomised controlled trial. *Value in Health* 2009; 12: 942–947.

Hoogeboom TJ, Eney CHM van den, Sluisz G van der et al. The impact of waiting for total joint replacement on pain and functional status. A systematic review. *Osteoarthritis and Cartilage* 2009; 17: 1420–1427.

Hujanen T, Kapiainen S, Tuominen U, Pekurinen M. Unit costs in Finnish healthcare 2006. Helsinki: Stakes, Working paper 3, 2008. Available at: <[http://www.terveydenhuollon\\_yksikkokustannukset\\_Suomessa\\_2006](http://www.terveydenhuollon_yksikkokustannukset_Suomessa_2006)>.

Insall JN, Dorr LD, Scott RD, Scott WN. Rationale of the knee society clinical rating system. *Clinical Orthopaedics and Related Research* 1989; 248: 13–14.

ICD10. International statistical Classification of Diseases and Related Health Problems: I 10:3, 1994. World Health Organization (WHO).

Iversen T. A theory of hospital waiting lists. *J Health Econ* 1993; 12: 55–71.

Jordan KM, Arden NK, Doherty M ym. EULAR Recommendations 2003. An evidence based approach to the management of knee osteoarthritis. Report of a Task Force of the Standing Committee for International Clinical Studies Including Therapeutic Trials (ESCISIT). *Ann Rheum Dis* 2003; 62: 1145–1155.

Karlsö 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.

Kelly KD, Voaklander DC, Johnston DWC, Newman SC, Suarez-Almazor ME. Change in pain and function while waiting for major joint arthroplasty. *J Arthroplasty* 2001; 16 (3): 351–359.

Kuster MS. Exercise recommendations after total joint replacement. A review of the current literature and proposal of scientifically based guidelines. *Sports Med* 2002; 32 (7): 433–445.

Lavernia CJ, Guzman JF, Gachupin-Garcia A. Cost effectiveness and quality of life in knee arthroplasty. *Clin Orthop Relat Res* 1997; 345: 134–139.

Lingard E, Katz J, Wright R, Wright E, Sledge C. Validity and responsiveness of the Knee Society Clinical Rating System in comparison with SF-36 and WOMAC. *JBJS* 2001; 83: 1856–1864.

Lyman GH. Economic analysis of randomized, controlled trials. *Curr Oncol Rep* 2001; 3: 396–403.

Mahomed NN, Arndt DC, McGrory BJ, Harris WH. The Harris hip score. Comparison of patients self-report with surgeon assessment. *J Ahrthroplasty* 2001; 16: 575–580.

March LM, Cross M, Tribe K, Lapsley H, Courtenay B, Brooks P. Cost of joint replacement surgery for osteoarthritis. The patients' perspective. *J Rheumatol* 2002; 29: 1006–1014.

McDowell P, Newell C. *Measuring Health. A guide to rating scales and questionnaires*. New York, NY: Oxford University Press, 1996.

McHugh GA, Campell M, Silman AS, Kay PR, Luker KA. Patients wait for a hip or knee replacement. Is there any prioritization for surgery? *J Eval Clin Pract* 2008; 14 (3): 361–367.

Mealli F, Li F. *Statistical methods in causal inference. Lectures*, University of Helsinki, 16–18.5.2011.

Mikkola H, Peltola M, Pekurinen M et al. Jonotetut leikkaukset Suomessa 1998–2006. Hoitotakuun vaikutusten arviointia. In: Pekurinen M, Mikkola H, Tuominen U. *Hoitotakuun talous*. Helsinki: Stakes, Reports 5, 2008.

Mikkola H, Peltola M, Tuominen U, Malmivaara A. A waiting time reform, lasting or short time effects? 2011 Submitted to Health Policy.

Ministry of Social Affairs and Health. 2004. Available at: <[http://www.stm.fi/en/social\\_and\\_health\\_services/legislation](http://www.stm.fi/en/social_and_health_services/legislation)>.

Montin L, Suominen T, Katajisto J, Lepistö J, Leino-Kilpi H. Economic outcomes from patients' perspective and health-related quality of life after total hip arthroplasty. *Scandic J Caring Sci* 2009; 23 (1): 11–20.

Mooch J, Kohlmann T. Comparing preference-based quality-of-life measures. Results from rehabilitation patients with musculoskeletal, cardiovascular, or psychosomatic disorders. *Qual Life Res* 2008; 17:485–495.

National Joint Registry 2011. Available at: <<http://www.njrcentre.org.uk>>.

National Research and Development Centre for Welfare and Health. Operative inpatients service 2006. Available at: <<http://www.stakes.fi/tilastot/tilastotiedotteet/2007>>.

Navarro Espigare JL, Hernandez Torres E. Cost-outcome analysis of joint replacement. Evidence from Spanish public hospital. *Gaceta Sanitaria* 2008; 22 (4): 336–343.

National Institute for Health and Clinical Excellence 2009. Available at: <<http://www.nice.org.uk>>.

Noseworthy TW, Sanmartin C, Conner-Spady B et al. Towards establishing evidence-based benchmarks for acceptable waiting times for joint replacement surgery. Report #2, October 2005. Available at: <[www.wcwl.ca/library/cihr\\_synthesis\\_reports/](http://www.wcwl.ca/library/cihr_synthesis_reports/)> (Version current at June 15, 2007).

Núñez M, Núñez E, Segur JM et al. Health-related quality of life and costs in patients with osteoarthritis on waiting list for total knee replacement. *Osteoarthritis Cartilage* 2007; 15: 258–265.

OECD. Health at a Glance 2011. OECD Indicators. Paris: OECD, OECD Publishing, 2011. Available at: <[http://dx.doi.org/10.1787/health\\_glance-2011-en](http://dx.doi.org/10.1787/health_glance-2011-en)>.

Oliver A. Health care reform in England. Progress and plans. *Health Policy Monitor, Surveys* 7, 2006. Available at: <<http://www.hpm.org>>.

Pharmaceutical Information Centre in Finland. *Pharmaca Fennica* CD-rom 1, 2004.

Piccirillo JF. Comorbidity as a new data element. Washington, DC: Committee on Standards, Commission on Cancer, 2000.

Quan H, Lafreniere R, Johnson D. Health service costs for patients on the waiting list. *Journal Canadien de chirurgie* 2002; 45 (1): 34–42.

Rawlins M, Culyer A. National Institute for Clinical Excellence and its value judgements. *BMJ* 2004; 329: 224–227.

Remes V, Peltola M, Häkkinen U et al. Perfect – Tekonivelkirurgia. Lonkan ja polven tekonivelkirurgian kustannukset ja vaikuttavuus. Helsinki: Stakes, Työpapereita 29, 2007.

Rissanen P, Aro S, Slätis P, Sintonen H, Paavolainen P. Health and quality of life before and after hip or knee arthroplasty. *J Arthroplasty* 1995; 10: 169–175.

Rissanen P, Aro S, Sintonen H, Slätis P, Paavolainen P. Quality of life and functional ability in hip and knee replacements. A prospective study. *Qual Life Res* 1996; 5: 1.

Rissanen P, Aro S, Sintonen H et al. Costs and cost-effectiveness in hip and knee replacements. A prospective study. *Int J Technol Assess Health Care* 1997; 13: 575–588.

Räsänen P, Sintonen H, Ryyänänen P, Blom M, Semberg-Konttinen V, Roine RP. Measuring cost-effectiveness of secondary health care. Feasibility and potential utilization of results. *Int J Technol Assess Health Care* 2005; 21 (1) 22–31.

Räsänen P, Roine E, Sintonen H, Semberg-Konttinen V, Ryyänänen OP, Roine R. Use of quality-adjusted life years for estimation of effectiveness of health care. A systematic literature review. *Int J Technol Assess Health Care* 2006; 22 (2): 235–241.

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

Schaafsma J. Are there better ways to determine wait times? *CMAJ*; 174 (11): 1551-1552.

Salaffi F, Carotti M, Grassi W. Health-related quality of life in patients with hip or knee osteoarthritis. Comparison of generic and disease-specific instruments. *Clin Rheumatol* 2005; 24: 29–37.

Scharf HP, Mansmann U, Streitberger K et al. Acupuncture and knee osteoarthritis. A three-armed randomized trial. *Ann Intern Med* 2006 4; 145 (1): 12–20.

Schulz KF, Grimes UA. Generation of allocation sequences in randomised trials. Chance, not choice. *Lancet* 2002; (Feb) 9; 359 (9305): 515–519.

Segal L, Day SE, Chapman AB, Osborne RH. Can we reduce burden from osteoarthritis? An evidence-based priority-setting model. *Bone and Joint Disorder: Prevention and Control* 2004; 180: 11–17. Available at: <[http://www.mja.com.au/public/issues/180\\_05\\_010304/seg10357\\_fm.html](http://www.mja.com.au/public/issues/180_05_010304/seg10357_fm.html)>.

Shah Pankaj B. Intention-to-treat and per-protocol analysis. *CMAJ* 2011 April 5; 183 (6): 696.

Siciliani L, Hurst J. Explaining waiting time variations for elective surgery across OECD countries. *OECD Health Working Papers* 2003; 7: 10–11.

Siciliani L, Hurst J. Explaining waiting time variations for elective surgery across OECD countries. *OECD Economic Studies* 2004; 1 (38): 95–123.

Sintonen H. The 15d measure of health-related quality of life: reliability, validity and sensitivity of its health state descriptive system. Melbourne: National Centre for Health Program Evaluation, Working paper 41, 1994a. Available at: <<http://www.buseco.monash.edu.au/centres/che/publications.php#4>>.

Sintonen H. Outcome measurement in acid-related diseases. *Pharmacoeconomics* 1994b; 5 (Suppl. 3): 17–24.

Sintonen H. The 15D measure of health-related quality of life. II feasibility and validity of its valuation system. Melbourne: National Centre for Health Program Evaluation, Working paper 42, 1995. Available at: <<http://www.buseco.monash.edu.au/centres/che/publications.php#4>>.

Sintonen H. The 15D instrument of health related quality of life. Properties and applications. *Ann Med* 2001; 33; 328–336.

Sobolev B, Kuramoto L. *Analysis of waiting time data in health services research*. New York, NY: Springer, 2008.

Stavem K. Reliability, validity and responsiveness of two multiattribute utility measures in patients with chronic obstructive pulmonary disease. *Quality of Life Research. An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation* 1999; 8: 45–54.

Swedish Association of Local Authorities and Regions (SALAR). *Swedish health Care in an international context. A comparison of care needs, costs, and outcomes*. Stockholm: SALAR, 2005. Available at: <<http://www.skl.se>>.

THL. Lonkka- ja polvipteeseit Suomessa 2010 [Finnish Arthroplasty Register]. Helsinki: THL, Tilastoraportti 23, 2011. Available at: <[http://www.stakes.fi/tilastot/tilastotiedotteet/2011/Tr23\\_11.pdf](http://www.stakes.fi/tilastot/tilastotiedotteet/2011/Tr23_11.pdf)>.

Tuominen U, Sintonen H, Hirvonen J et al. The effect of waiting time on health outcomes and use and costs of medication in hip replacement patients. A randomised clinical trial. *Osteoarthritis and Cartilage* 2009; 17: 1144–1150.

Tuominen U, Sintonen H, Hirvonen J et al. Is longer waiting time for total knee replacement associated with health outcomes and medication costs? A randomized clinical trial. *Value in Health* 2010; 13 (8): 998–1004

Willcox S, Seddon M, Dunn S, Tudor Edwards R, Pearse J, Tu JV. Measuring and reducing waiting times. A cross-national comparison of strategies. *Market Watch. Health Affairs* 2007; 26 (4): 1078–1087.

Xie F, Lo NN, Tarride JE, O'Reilly D, Lee HP. Total or partial knee replacement? Cost-utility analysis in patients with knee osteoarthritis based on a 2-year observational study. *Eur J Health Econ* 2010; 11: 27–34.

Xuan J, Kirchoefer L, Boyer J et al. Effects of comorbidity on health related quality of life. Scores: An analysis of clinical trial data. *Clin Ther* 1999; 21: 2.

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 (ESCIIT). *Ann Rheum Dis* 2005; 64: 669–681.

Zhang W, Doherty M, Peat G et al. EULAR evidence-based recommendations for the diagnosis of knee osteoarthritis. *Ann Rheum. Dis.* 2010a; 69 (3): 483–489.

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* 2010b; 18: 476–499.

## APPENDICES

### QUALITY OF LIFE QUESTIONNAIRE (15D©)

Please read through all the alternative responses to each question before placing a cross (x) against the alternative that best describes **your present health status**. Continue through all 15 questions in this manner, giving only **one** answer to each.

#### QUESTION 1. MOBILITY

- 1 ( ) I am able to walk normally (without difficulty) indoors, outdoors and on stairs.
- 2 ( ) I am able to walk without difficulty indoors, but outdoors and/or on stairs I have slight difficulties.
- 3 ( ) I am able to walk without help indoors (with or without an appliance), but outdoors and/or on stairs only with considerable difficulty or with help from others.
- 4 ( ) I am able to walk indoors only with help from others.
- 5 ( ) I am completely bed-ridden and unable to move about.

#### QUESTION 2. VISION

- 1 ( ) I see normally, i.e. I can read newspapers and TV text without difficulty (with or without glasses).
- 2 ( ) I can read papers and/or TV text with slight difficulty (with or without glasses).
- 3 ( ) I can read papers and/or TV text with considerable difficulty (with or without glasses).
- 4 ( ) I cannot read papers or TV text either with glasses or without, but I can see enough to walk about without guidance.
- 5 ( ) I cannot see enough to walk about without a guide, i.e. I am almost or completely blind.

#### QUESTION 3. HEARING

- 1 ( ) I can hear normally, i.e. normal speech (with or without a hearing aid).
- 2 ( ) I hear normal speech with a little difficulty.
- 3 ( ) I hear normal speech with considerable difficulty; in conversation I need voices to be louder than normal.
- 4 ( ) I hear even loud voices poorly; I am almost deaf.
- 5 ( ) I am completely deaf.

#### QUESTION 4. BREATHING

- 1 ( ) I am able to breathe normally, i.e. with no shortness of breath or other breathing difficulty.
- 2 ( ) I have shortness of breath during heavy work or sports, or when walking briskly on flat ground or slightly uphill.
- 3 ( ) I have shortness of breath when walking on flat ground at the same speed as others my age.
- 4 ( ) I get shortness of breath even after light activity, e.g. washing or dressing myself.
- 5 ( ) I have breathing difficulties almost all the time, even when resting.



**QUESTION 5. SLEEPING**

- 1 ( ) I am able to sleep normally, i.e. I have no problems with sleeping.
- 2 ( ) I have slight problems with sleeping, e.g. difficulty in falling asleep, or sometimes waking at night.
- 3 ( ) I have moderate problems with sleeping, e.g. disturbed sleep, or feeling I have not slept enough.
- 4 ( ) I have great problems with sleeping, e.g. having to use sleeping pills often or routinely, or usually waking at night and/or too early in the morning.
- 5 ( ) I suffer severe sleeplessness, e.g. sleep is almost impossible, even with full use of sleeping pills or staying awake most of the night.

**QUESTION 6. EATING**

- 1 ( ) I am able to eat normally, i.e. with no help from others.
- 2 ( ) I am able to eat by myself with minor difficulty (e.g. slowly, clumsily, shakily, or with special appliances).
- 3 ( ) I need some help from another person in eating.
- 4 ( ) I am unable to eat by myself at all, so I must be fed by another person.
- 5 ( ) I am unable to eat at all, so I am fed either by tube or intravenously.

**QUESTION 7. SPEECH**

- 1 ( ) I am able to speak normally, i.e. clearly, audibly and fluently.
- 2 ( ) I have slight speech difficulties, e.g. occasional fumbling for words, mumbling, or changes of pitch.
- 3 ( ) I can make myself understood, but my speech is disjointed, faltering, stuttering or stammering.
- 4 ( ) Most people have great difficulty understanding my speech.
- 5 ( ) I can only make myself understood by gestures.

**QUESTION 8. EXCRETION**

- 1 ( ) My bladder and bowel work normally and without problems.
- 2 ( ) I have slight problems with my bladder and/or bowel function, e.g. difficulties with urination, or loose or hard bowels.
- 3 ( ) I have marked problems with my bladder and/or bowel function, e.g. occasional 'accidents', or severe constipation or diarrhoea.
- 4 ( ) I have serious problems with my bladder and/or bowel function, e.g. routine 'accidents', or need of catheterization or enemas.
- 5 ( ) I have no control over my bladder and/or bowel function.

**QUESTION 9. USUAL ACTIVITIES**

- 1 ( ) I am able to perform my usual activities (e.g. employment, studying, housework, free-time activities) without difficulty.
- 2 ( ) I am able to perform my usual activities slightly less effectively or with minor difficulty.
- 3 ( ) I am able to perform my usual activities much less effectively, with considerable difficulty, or not completely.
- 4 ( ) I can only manage a small proportion of my previously usual activities.
- 5 ( ) I am unable to manage any of my previously usual activities.

**QUESTION 10. MENTAL FUNCTION**

- 1 ( ) I am able to think clearly and logically, and my memory functions well.
- 2 ( ) I have slight difficulties in thinking clearly and logically, or my memory sometimes fails me.
- 3 ( ) I have marked difficulties in thinking clearly and logically, or my memory is somewhat impaired.
- 4 ( ) I have great difficulties in thinking clearly and logically, or my memory is seriously impaired.
- 5 ( ) I am permanently confused and disoriented in place and time.

**QUESTION 11. DISCOMFORT AND SYMPTOMS**

- 1 ( ) I have no physical discomfort or symptoms, e.g. pain, ache, nausea, itching, etc.
- 2 ( ) I have mild physical discomfort or symptoms, e.g. pain, ache, nausea, itching, etc.
- 3 ( ) I have marked physical discomfort or symptoms, e.g. pain, ache, nausea, itching, etc.
- 4 ( ) I have severe physical discomfort or symptoms, e.g. pain, ache, nausea, itching, etc.
- 5 ( ) I have unbearable physical discomfort or symptoms, e.g. pain, ache, nausea, itching, etc.

**QUESTION 12. DEPRESSION**

- 1 ( ) I do not feel at all sad, melancholic or depressed.
- 2 ( ) I feel slightly sad, melancholic or depressed.
- 3 ( ) I feel moderately sad, melancholic or depressed.
- 4 ( ) I feel very sad, melancholic or depressed.
- 5 ( ) I feel extremely sad, melancholic or depressed.

**QUESTION 13. DISTRESS**

- 1 ( ) I do not feel at all anxious, stressed or nervous.
- 2 ( ) I feel slightly anxious, stressed or nervous.
- 3 ( ) I feel moderately anxious, stressed or nervous.
- 4 ( ) I feel very anxious, stressed or nervous.
- 5 ( ) I feel extremely anxious, stressed or nervous.

**QUESTION 14. VITALITY**

- 1 ( ) I feel healthy and energetic.
- 2 ( ) I feel slightly weary, tired or feeble.
- 3 ( ) I feel moderately weary, tired or feeble.
- 4 ( ) I feel very weary, tired or feeble, almost exhausted.
- 5 ( ) I feel extremely weary, tired or feeble, totally exhausted.

**QUESTION 15. SEXUAL ACTIVITY**

- 1 ( ) My state of health has no adverse effect on my sexual activity.
- 2 ( ) My state of health has a slight effect on my sexual activity.
- 3 ( ) My state of health has a considerable effect on my sexual activity.
- 4 ( ) My state of health makes sexual activity almost impossible.
- 5 ( ) My state of health makes sexual activity impossible.

## MODIFIED DISEASE-SPECIFIC MEASURES

Hip patients			Knee patients		
Item	Description	Points	Item	Description	Points
Pain	1 None	44	Pain	1 None	50
	2 Slight, occasional	40		2 Mild and occasional	45
	3 Mild pain, rarely with unusual activities	30		3 Mild, stairs only	40
	4 Moderate pain, some limitation of ordinary activities	20		4 Mild, walking and stairs	40
	5 Marked pain, limitation of activities	10		5 Moderate, occasional	20
	6 Totally disable, pain in bed	0		6 Moderate, continual	10
			7 Severe	0	
Function (Gait)			Function		
1. Limp	1 None	11	1. Walking	1 > 1.5 km / unlimited	50
	2 Slight	8		2 1–1.5 km	40
	3 Moderate	5		3 100–500 m	20
	4 Severe	0		4 Indoors only	10
2. Support	1 None	11	5 Unable	0	
	2 Cane for long walks	7	2. Stairs	1 Normal, up and down	50
	3 Cane most of the time	5		2 Normal up, down with rail	40
	4 One crutch	3		3 Up and down with rail	30
	5 Two canes	2		4 Up with rail, unable down	15
	6 Two crutches	0		5 Unable	0
	7 Not able to walk	0			
3. Distance walked	1 > 1.5 km / unlimited	11			
	2 1–1.5 km	8			
	3 100–500 m	5			
	4 Indoors only	2			
	5 Unable	0			
Function (Activities)					
1. Stairs	1 Normally	4			
	2 Normally, using a railing	2			
	3 Step by step, using a railing	1			
	4 Unable to do stairs	0			
2. Shoes and socks	1 With ease	4			
	2 With difficulty	3			
	3 Unable	0			
3. Sitting	1 In ordinary chair one hour	5			
	2 On a higher chair for one hour	3			
	3 Unable to sit	0			



## ORIGINAL PUBLICATIONS



Tuominen U, Blom M, Hirvonen J, Hietaniemi K, Lehto MUK, Paavolainen P, Seitsalo S, Sintonen H, Rissanen P. Effects of comorbidities to health-related quality of life in patients awaiting major joint replacement. *Health and Quality of Life Outcomes* 2007; 5: 16; doi:10.1186/1477-7525-5-16.





Research

Open Access

## The effect of co-morbidities on health-related quality of life in patients placed on the waiting list for total joint replacement

Ulla Tuominen\*<sup>1,2</sup>, Marja Blom<sup>3,4</sup>, Johanna Hirvonen<sup>2</sup>, Seppo Seitsalo<sup>5</sup>,  
Matti Lehto<sup>6</sup>, Pekka Paavolainen<sup>5</sup>, Kalevi Hietaniemi<sup>3</sup>, Pekka Rissanen<sup>7</sup> and  
Harri Sintonen<sup>2</sup>

Address: <sup>1</sup>National Research and Development Centre for Welfare and Health, Helsinki, Finland, <sup>2</sup>University of Helsinki, Finland, <sup>3</sup>HUCH Hospital Area, Espoo, Finland, <sup>4</sup>Academy of Finland, <sup>5</sup>Orton Orthopaedic Hospital, Helsinki, Finland, <sup>6</sup>Coxa, Hospital for Joint Replacement, Medical Research Fund of Tampere University Hospital, Tampere, Finland and <sup>7</sup>University of Tampere, Finland

Email: Ulla Tuominen\* - ulla.tuominen@stakes.fi; Marja Blom - marja.blom@hus.fi; Johanna Hirvonen - johanna.hirvonen@mikkelihamk.fi; Seppo Seitsalo - seppo.seitsalo@invalidisaatio.fi; Matti Lehto - matti.lehto@pshp.fi; Pekka Paavolainen - pekka.paavolainen@invalidisaatio.fi; Kalevi Hietaniemi - kalevi.hietaniemi@hus.fi; Pekka Rissanen - pekka.rissanen@uta.fi; Harri Sintonen - harri.sintonen@helsinki.fi

\* Corresponding author

Published: 15 March 2007

Received: 5 February 2007

Health and Quality of Life Outcomes 2007, 5:16 doi:10.1186/1477-7525-5-16

Accepted: 15 March 2007

This article is available from: <http://www.hqlo.com/content/5/1/16>

© 2007 Tuominen et al; licensee BioMed Central.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

**Background:** Co-morbidity is a powerful predictor of health care outcomes and costs, as well as an important cofounder in epidemiologic studies. The effect of co-morbidities is generally related to mortality or complications. This study evaluated the association between co-morbidity and health-related quality of life (HRQoL) in patients awaiting total joint replacement.

**Methods:** A total of 893 patients were recruited to the study between August 2002 and November 2003 in four Finnish hospitals. The effect of co-morbidity on HRQoL was measured by the generic 15D instrument and by a Visual Analog Scale (VAS). Comparative variance analysis of socio-demographic and clinical characteristics was described by using either an independent samples t-test or the Chi-square test. The differences in each of the 15D dimensions and the overall 15D single index score for patients were calculated. Two-sided p-values were calculated with the Levene Test for Equality of Variances.

**Results:** Patients with co-morbidity totaled 649; the incidence of co-morbidity was 73%. The mean number of co-morbidities among the patients was two. At baseline the 15D score in patients with and without co-morbidity was 0.778 vs 0.816, respectively. The difference of the score (0.038) was clinically and statistically significant ( $P < 0.001$ ). The patients' scores with and without co-morbidity on the different 15D dimensions related to osteoarthritis—moving, sleeping, usual activities, discomfort and symptoms, vitality and sexual activity—were low in both groups. Patients with co-morbidity scored lower on the dimensions of moving, vitality and sexual activity compared to the patients without co-morbidity. Co-morbidity was significantly associated with a reduced HRQoL. Patients without co-morbidity had poorer VAS, arthritis had strong effect to their quality of life compared to the patients with co-morbidity.

**Conclusion:** Assessing co-morbidity in patients placed on the waiting list for joint replacement may be useful method to prioritization in medical decision-making for healthcare delivery. The assessment of co-morbidities during waiting time is important as well as evaluating how the co-morbidity may affect the final outcomes of the total joint replacement.

## Background

Chronic diseases have been shown to negatively affect people's quality of life (QoL), and they are a common reason for disability and early death. According to the Finnish National Health 2000 Survey half of the Finnish population aged over 30 has at least one chronic disease, while 44 % of the working population and 82 % of senior citizens have at least one chronic disease [1]. The decline of the populations' self-reported sense of well-being is a consequence of ageing, with no observable difference between women and men. The commonest chronic diseases in Finland are: cardiovascular disease, musculoskeletal disease, diabetes, and lung diseases [1].

Co-morbid or coexisting disease refers to the occurrence of two or more diseases in the same individual. The presence of co-morbidity has a pervasive effect on QoL, research, and clinical practice through its influence on diagnosis, prognosis, treatment and, health care delivery. Each co-morbid disease may have its own effect on QoL while also having a clinical effect on patients' sense of well-being. It is also important for studies of patients with chronic disease in whom mortality is rare and the goal of medical care is to control the course of the disease and maximize the quality of life [2]. Evaluations on how different diseases affect patients' health-related quality of life (HRQoL) focus mostly on the index disease, considering the effect of co-morbidities to a lesser extent. However when the focus is on the consequences—mortality, complications or in the costs of the medical care – co-morbidity becomes an important denominator.

Co-morbidity can also play an important role in different types of studies. Randomized controlled trials and prognostic studies might be complicated by co-morbidity. It can either act as a cofounder, threatening the internal validity, or as an effect modifier, threatening the internal and external validity of the study: therefore an efficient method is needed to measure co-morbidity [3].

In Finland, total joint replacements (TJR) are surgical procedures with high volume and long waiting times. In 2003, primary hip and knee replacements were carried out for almost 15500 patients [4]. For patients with primary total joint replacement of hip the median waiting time was 155 days, and for patients with primary total joint replacement of knee, 205 days [4]. One reason for these long waiting times is that OA is not itself life threatening. However, some previous studies have reported that those awaiting hip or knee replacement have a significantly poorer quality of life and that arthritis becomes a chronic and heavy burden to the patients [5,6]. Also few studies having examined waiting time effects on health status in OA patients have not been able to show that patients having to wait longer would suffer from pain and functional

difficulties or poorer HRQoL than those with shorter waiting [7,8].

In 2002, a prospective multi-centre study was started in four Finnish hospitals. The aim of this larger study is to assess the effect, the costs and the cost effectiveness of the waiting time in patients awaiting TJR. This report is part of this ongoing study. The objective of this paper is to evaluate the effect of co-morbidity on HRQoL in OA patients when placed on the waiting list for TJR.

## Methods

### Data collection

Between August 2002 and November 2003, a total of 893 OA patients were enrolled in the study in four Finnish hospitals: the Helsinki University Central Hospital: Surgical Hospital, the Helsinki University Central Hospital: Jorvi Hospital, the Coxa Hospital for Joint Replacement and the Orton Orthopedic Hospital. Patients were recruited into the study through contact with orthopedic and practice staff.

The key inclusion criteria were a need for a primary TJR due to OA of the hip or knee joint as evaluated by the hospital surgeon; patient was aged 16 or older and placed on the waiting list in a research hospital, and the patient was willing and mentally able to participate in the study. The key exclusion criteria were patients with rheumatoid arthritis, congenital hemophilia or congenital deformities, and fractures. Patients completed a self-administered questionnaire when placed on the waiting list for TJR. The questionnaires were distributed to the patients at the hospital and were returned by post. Common guidelines for administering the questionnaires were provided at each hospital. The patients completed a socio-demographic form, reported their co-morbidities as diagnosed by a medical doctor, completed the visual analog scale (VAS), and also completed separate questionnaires for self-reported sense of well-being and HRQoL. Each patient provided informed consent. The study was approved by the Helsinki University Central Hospital Surgery Ethics Committee.

In this study, the co-morbidity data was collected from the patients' reported co-morbidities as diagnosed by a medical doctor. The patients were assigned to subgroups according to co-morbidity status. The patients' reported diseases were classified according to the ICD10 [9], giving nine diagnoses groups in total: tumors, diabetes mellitus, respiratory disease, cardiovascular disease, high cholesterol, mental health problems, musculoskeletal system diseases, endocrinological problems, and visual or hearing problems.

### Co-morbidity and Quality of Life

It is difficult to choose the most appropriate co-morbidity measurement because comparative data on how the available instruments perform in different disease settings are limited [10,11]. There are several reports about different measurements and measurement combinations for assessing co-morbidity. Instruments used in clinical research to calculate co-morbidity include, for example, the Carlson Index [13], the Cumulative Illness Rating Scale (CIRS) [14], the Index of Coexistent Diseases (ICED) [15] and the Kaplan Index [16]. Both generic and disease-specific QoL instruments are used to assess the effect of co-morbidities in clinical trials [12,17].

In this study the effect of co-morbidity was assessed as the difference in HRQoL between the patients with and those without co-morbidity. HRQoL was measured by using the generic, multidimensional, standardized, and self-administered 15D instrument. The 15D is a Multi-Attribute-Utility-Scale (MAU) measurement instrument [18] that measures quality of life in 15 dimensions: moving, vision, hearing, breathing, sleeping, eating, speech, eliminating, vitality, mental functions, discomfort and symptoms, depression, distress, energy and sexual activity. Each dimension has a single question with 5 possible answer options. The 15D can be used as a profile measure or to give a single index score by means of population-based preference weights. The index score is between 0 (being dead) and 1 (being totally healthy). Completing the 15D questionnaire takes 5–10 minutes and it describes the respondents HRQoL at that point in time. The minimum clinically important difference (MCID) in the 15D single index score is interpreted as a difference of  $\pm 0.03$  or more, which corresponds to the minimum difference that people can generally distinguish [19]. The 15D favourably compares with other quality of life instruments—such as EuroQol (EQ-5D), Health Utility Index (HUI 1–3), Short Form-36 (SF-36), and the Nottingham Health Profile (NHP)—in most of the important properties (e.g. responsiveness, reliability and validity) [20–22].

To evaluate the patients' sense of well-being, we used VAS, which is a health-state rating scale. The measurement consists of a line on a page with clearly defined endpoints. The most preferred health state is placed at one end of the line and the least preferred at the other end [23,24]. In this study, VAS was a horizontal 100 mm long line (100 mm = 100%) illustrating the patients' health state deficiency at that moment. It was used to evaluate the effect of arthritis on health. Patients were asked to mark on the line which part of the deficiency of health is due to arthritis. The higher the number was—on scale from 0 to 100—the more powerful was the effect of arthritis. In addition, the patients' self-reported state of health was described with a

five-point scale, representing health states from excellent to worst.

### Statistical analysis

At the baseline (when placed on the waiting list) descriptive statistics were used to describe the socio-demographic and clinical characteristics of the patients. Comparative variance analyses of socio-demographic and clinical characteristics were described by using either an independent samples t-test or the Chi-square test depending on whether it was a continuous or nominal scale. The differences in each of the 15D dimensions and the overall 15D single index score for patients were calculated. Two-sided p-values were calculated with the Levene Test for Equality of Variances, with the minimum significance level set at 5% (P-value < 0.05). The mean differences between each of the dimensions were also calculated. Missing values for the 15D were predicted by means of a regression model with the patient's responses for other dimensions, and also with data from the patients, with age and gender as explanatory variables [22]. The missing values were estimated if a minimum of 80% of dimensions had been completed. Furthermore, the incidences of co-morbid diagnoses and also the mean 15D score were calculated for each diagnosis group in the model. Data analyses were performed using SPSS version 12.0.1 for Windows.

### Results

Of the 914 eligible patients recruited into this study, twenty one were excluded because they didn't return the questionnaire, leaving 893 patients in the study group. The mean age of patients waiting for TJR was 66 years (range 24–88) and 63% of the participants were female. Patients with co-morbidity totaled 649, while patients without co-morbidity totaled 244. The mean age in the patients with co-morbidity was 67 years (range 25–87), versus 64 years (range 24–88) without co-morbidity (P < 0.001). Patient's demographics are reported in Table 1. The incidence of co-morbidity as a secondary or tertiary illness was 73%. The mean number of co-morbidities among the patients was two, and 363 (56% of co-morbid patients) had three or more diagnosis. The BMI was high in both groups (>25 which is the limit of overweight and >30 is a limit of obese) but in the co-morbid group, the BMI was higher than in the patients without co-morbidity (P < 0.001). The patients' health state based on the five-point scale was worse in the co-morbidity group. However in the patients with co-morbidity, the effect of the OA in the health-state deficiency as measured by VAS was 62% versus 76% (P < 0.001) in the patients without co-morbidity (Table 1).

The most common secondary diagnosis was cardiovascular disease (n = 419, 63%), followed by high cholesterol (n = 225, 33%), diabetes mellitus (n = 225, 33%) and

**Table 1: Demographic and clinical characteristics of TJR patients with and without co-morbidity when placed on the waiting list. Independent Sample T-test or Chi-square test**

Characteristics	Patients with co-morbidity	Patients without co-morbidity	P-value
	(n = 649–662 <sup>1</sup> )	(n = 242–244 <sup>1</sup> )	
Age, years (mean ± SD)	67 ± 10	64 ± 11	0.000***
Females [n, (%)]	422(63)	132 (54)	0.008**
Marital status [n, (%)]			
Married	411 (62)	160 (65)	0.155
Housing [n,%]			
Living alone	227 (34)	62 (25)	0.008**
Basic education [n,%]			
Low level	544 (82)	190(78)	0.130
Professional examination [n,%]			
Low level	486 (75)	175 (73)	0.279
Employment status [n, (%)]			
Retired	546 (82)	172(71)	0.000***
Health status [n, (%)]			
Fair or poor	497 (75)	133 (55)	0.000***
VAS <sup>2</sup> (mean ± SD)	62 ± 26	76 ± 24	0.000***
BMI <sup>3</sup> (mean ± SD)	28.8 ± 4	27.2 ± 4	0.000***

<sup>1</sup>Number of observation varies due to missing values; <sup>2</sup>VAS, visual analog scale (100 the worst, 0 the best value); <sup>3</sup>BMI, body mass index (wt/ht<sup>2</sup>); \*\* P < 0.01, \*\*\*P < 0.001

endocrinological problems (n = 225, 33%). The worst 15D score of 0.637 was for a patient with mental problems (Table 2).

At baseline the 15D score in patients with and without co-morbidity was 0.778 vs 0.816, respectively. The difference of the score (0.038) was clinically and statistically significant (P < 0.001). The patients' scores with and without co-morbidity on the different 15D dimensions related to OA—moving, sleeping, usual activities, discomfort and symptoms, vitality and sexual activity—were low in both groups. Patients with co-morbidity scored lower on the dimensions of moving, vitality and sexual activity compared to the patients without co-morbidity, while patients without co-morbidity scored lower in the dimensions of seeing, hearing, breathing and elimination. The deteriora-

tion of HRQoL was significantly associated with co-morbidity (Table 3).

## Discussion

The aim of this study was to assess the effect of co-morbidities on HRQoL at baseline in patients awaiting major joint replacement in four Finnish hospitals. The main finding of this study was that the HRQoL of all TJR patients was poor but significantly worse in the patients with co-morbidity. A secondary finding was that the VAS health-rating instrument appears more appropriate as a disease-specific instrument, as in this study, patients did not necessarily assign the primary disease (arthritis in most cases) as being the most prominent in affecting their well-being. Furthermore, the five-point self-rated health scale showed that the health state of patients with co-morbidity was worse than the patients without co-morbidity,

**Table 2: The co-morbidities among osteoarthritis patients and the mean 15D score in each diagnose groups, when placed on the waiting list for TJR**

Co-morbidity	Patients, n (%)	15D (± SD)
Cardiovascular disease	419 (63%)	0.775 (± 0.09)
High cholesterol	225 (34%)	0.769 (± 0.09)
Diabetes mellitus	225 (34%)	0.769 (± 0.09)
Endocrinological problems	225 (34%)	0.769 (± 0.09)
Other muskuloskeletal diseases	117 (18%)	0.758 (± 0.09)
Respiratory diseases	91 (14%)	0.758 (± 0.09)
Visual or hearing problems	54 (8%)	0.780 (± 0.07)
Tumors	25 (4%)	0.735 (± 0.08)
Mental problems	14 (2%)	0.637 (± 0.09)

**Table 3: 15D dimensions and score at baseline in total joint replacement patients with and without co-morbidity. Levene's Test for Equality of Variances**

Dimensions	Patients co-morbidity (n = 662)		Patients no co-morbidity (n = 242)		Mean difference	P-value
	Mean	SE	Mean	SE		
Moving	0.590	0.137	0.626	0.137	0.036	0.000***
Seeing	0.905	0.181	0.934	0.152	0.030	0.015*
Hearing	0.915	0.154	0.945	0.129	0.031	0.003**
Breathing	0.812	0.235	0.937	0.138	0.125	0.000***
Sleeping	0.723	0.217	0.738	0.210	0.014	0.375
Eating	0.991	0.057	0.997	0.032	0.006	0.051
Communication	0.986	0.070	0.993	0.046	0.007	0.102
Elimination	0.827	0.216	0.910	0.162	0.083	0.000***
Usual activities	0.643	0.222	0.679	0.239	0.036	0.045*
Mental	0.889	0.176	0.906	0.162	0.017	0.180
Discomfort and symptoms	0.506	0.236	0.540	0.240	0.033	0.065
Depression	0.837	0.171	0.854	0.167	0.016	0.196
Distress	0.848	0.177	0.865	0.162	0.017	0.168
Vitality	0.737	0.179	0.789	0.153	0.052	0.000***
Sexual activity	0.732	0.281	0.783	0.243	0.051	0.008**
15D-index	0.778	0.092	0.816	0.281	0.038	0.000***

\*P < 0.05; \*\*P < 0.01; \*\*\*P < 0.001

which is in line with our results of a decline from the baseline of HRQoL as measured by 15D.

Few studies have assessed the effect of co-morbidities on the quality of life (QoL) by either a generic or a disease-specific instrument. In 1999 Xuan et al. [2] reported how different measuring methods differ when evaluating the different effects of co-morbidities related to QoL. They found that co-morbidity extensively affects generic QoL, whereas the effect is considerably smaller for disease-specific measures. Salaffi et al. [17] studied the relationship between OA, co-morbidity and HRQoL in older adults compared with matched healthy controls. They found that 55% of patients reporting at least one chronic coexisting disease and OA of the lower extremities suffer a significant impact on multiple dimensions of HRQoL compared with healthy controls. The most significant impacts were seen in physical functions, physical roles, and pain. Both of these findings are in line with our results.

Cardiovascular disease is the most common disease in Finland [1] and in our study, it was also the most common coexisting disease. According to Shan et al. [14] if OA is related in some way to a co-morbidity i.e. cardiovascular disease, it might negatively affect the outcomes of the joint replacement. Furthermore, several studies have testified that arthritis is considered to be a risk factor for other co-morbidity conditions such as hypertension, heart disease, diabetes and chronic lung disease [17,25].

Limitations of this study include the severity of the co-morbidity not being known, and the fact that arthritis was necessarily classified as the primary disease. However, we can assume that the patients' coexisting diseases were not life threatening, because the severity of the co-morbidity in the patients is usually an exclusion criterion for the surgical operation. The other limitation related to co-morbidity data was that the data was collected from the patients, not from the patient's medical records. However the patients were asked to name only those co-morbidities diagnosed by a medical doctor.

Co-morbidity has been commonly measured as an index in studies where medical records have been used to investigate mortality, complications or the costs of the medical care. A significantly strong association between co-morbidity and mortality, complications and increasing in hospital costs [2,12,14,25-28] has been shown, but to our knowledge there are no previous studies on the effect of co-morbidity on the HRQoL in the patients waiting for TJR.

Our findings show statistically significant differences in the 15D-index between the groups with and without co-morbidity. This suggests that a generic measurement instrument is sensitive enough to identify the effects of co-morbidities. This study provides evidence of co-morbidity being a factor that significantly affects HRQoL, and which can be assessed when the patients are placed on the waiting list. Moreover, some studies have suggested that the use of a generic HRQoL measurement in the studies of OA

where co-morbidity is common would be useful in characterizing the global burden of this disease [16,17].

## Conclusion

Severity of OA was the only inclusion criteria of patients when placed on the waiting list, and the surgery was performed according to the hospital's routine procedure, this study had no effect to this. In these analyses, we found that the HRQoL of all TJR patients was poor but significantly worse in the patients with co-morbidity. Further, VAS health-rating instrument appears more appropriate as a disease-specific instrument, as being the most prominent in affecting patients' well-being, and also the five-point self-rated health scale showed that the health state of patients with co-morbidity was worse than the patients without co-morbidity. Assessing the co-morbidity condition at baseline might operate as an instrument to help in prioritization in medical decision-making for healthcare delivery. The assessment of co-morbidities during waiting time is important as well as evaluating how the co-morbidity may affect the final outcomes of the whole procedure.

## Competing interests

The author(s) declare that they have no competing interests.

## Authors' contributions

UT was the correspondence author of the manuscript and responsible for the integrity of the work as a whole. She contributed as a principal researcher and writer including drafting the article and the analysis and interpretation of the data. MB was the leader of the research project. She made contributions and design, acquisition and interpretation of the data and participated in the writing process by commenting the manuscript. JH made contributions to design, acquisition, and interpretation of data. HS and PR contributed as specialists in the field, and were involved in the design of the study and hypothesis formation. PP, SS, ML and KH contributed as specialists in the field of orthopaedic surgery. They made contributions to design and acquisition of data.

## Acknowledgements

This study was financially supported by the Academy of Finland (no.51871), Helsinki University Central Hospital, Jorvi Hospital; Coxa Hospital for Joint Replacement, Medical Research Fund of Tampere University Hospital; Helsinki University Central Hospital, Surgical Hospital and Orton Orthopaedic Hospital.

## References

- National Health Institution: **Health in 2000. Finnish National Survey 2003.**
- Xuan J, Kirchoerfer LJ, Boyer JG, Norwood GJ: **Effects of comorbidity on health related quality of life. Scores: An analysis of clinical trial data.** *Clin Ther* 1999, 21:383-403.
- de Groot V, Beckerman H, Lankhorst GJ, Bouter LM: **How to measure comorbidity, a critical review of available methods.** *J Clin Epidemiol* 2003, 56:221-229.
- National Research and Development Centre for Welfare and Health: **Operative inpatients service 2003.** [[http://www.stakes.info/2/9/hoitojaksor\\_2003.asp](http://www.stakes.info/2/9/hoitojaksor_2003.asp)].
- Hirvonen J, Blom M, Tuominen U, Seitsalo S, Lehto M, Paavolainen P, Hietaniemi K, Rissanen P, Sintonen H: **Health-related quality of life in patients waiting for major joint replacement. A comparison between patients and populations.** *Health Qual Life Outcomes* 2006, 4:3.
- Croft P, Lewis M, Wynn-Jones C, Coggon D, Cooper C: **Health status in patients awaiting hip replacement for osteoarthritis.** *Rheumatology* 2002, 9:1001-1007.
- Kelly KD, Voaklander DC, Johnston DW, Newman SC, Suarez-Almazor ME: **Change in pain and function while waiting for major joint arthroplasty.** *J Arthroplasty* 2001, 16:351-359.
- Derret S, Paul C, Morris J: **Waiting for elective surgery: effect on health-related quality of life.** *Int J Qual Health Care* 1999, 11:47-57.
- International statistical Classification of Diseases and Related Health Problems: **ICD10.** *World Health Organization (WHO)* 1994, 10:3.
- Imamura K, Black N: **Does comorbidity affect the outcome of surgery? Total hip replacement in UK and Japan.** *Int J Qual Health Care* 1998, 10:113-123.
- Gabriel S, Crowson C, O'Fallon A: **A Comparison of two comorbidity instruments in arthritis.** *J Clin Epidemiol* 1999, 52:1137-1142.
- Greenfield S, Apolone G, McNeil BJ, Cleary PD: **The importance of co-existent disease in the occurrence of postoperative complications and one-year recovery in patients undergoing total hip replacement. Comorbidity and outcomes after hip replacement.** *Med Care* 1993, 31:141-154.
- Charlson ME, Pompei P, Ales KL, MacKenzie CR: **A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation.** *J Chronic Dis* 1987, 40:373-383.
- Shah AN, Vail TP, Taylor D, Pietrobbon R: **Comorbid illness affects hospital costs related to hip arthroplasty. Quantification of health status and implications for fair reimbursement and surgeon comparisons.** *J Arthroplasty* 2004, 19:700-705.
- Linn BS, Linn MW, Gurel L: **Cumulative illness rating scale.** *J Am Geriatr Soc* 1968, 16:622-626.
- Piccirillo JF: **Comorbidity as a New Data Element.** In *Presentation to the Committee on Standards Commission on Cancer* Edited by: Louis St, MO. Washington University School of Medicine; 2000.
- Salaffi F, Carotti M, Stancati A, Grassi W: **Health-related quality of life in older adults with symptomatic hip and knee osteoarthritis: comparison with matched healthy controls.** *Aging Clin Exp Res* 2005, 17:255-263.
- Brazier J, Derveril M: *The use of health-related quality of life measures in economic evaluation. Health Services Research Methods. Guide to best practice* (BMJ) books, London; 1998.
- Sintonen H: **The 15D measure of health-related quality of life: properties and applications.** *Ann Med* 2001, 33:328-336.
- Bowling A: *Measuring Health. A review of quality of life measurement scores* 3rd edition. Open University Press, Berkshire; 2005.
- Sintonen H: *The 15d measure of health-related quality of life: reliability, validity and sensitivity of its health state descriptive system. Working paper 41. NCHPE* edition. 1994 [[http://www.buseco.monash.edu.au/centres/che/publications.php#4\\_09/06/05](http://www.buseco.monash.edu.au/centres/che/publications.php#4_09/06/05)]. Australia. Melbourne
- Sintonen H: [<http://www.15d-instrument.net/15d/>].
- McDowell P, Newell C: *Measuring Health. A guide to rating scales and questionnaires* 2nd edition. Oxford University Press, New York; 1996.
- Rissanen P, Aro S, Sintonen H, Slati S, Paavolainen P: **Quality of life and functional ability in hip and knee replacements: A prospective study.** *Qual Life Res* 1996, 1:56-64.
- Mont MA, Mears SC, Jones LC, Rajadhyaksha AD, Krackow AM, Bawa M, Hungerford DS: **Is coding of diagnosis, comorbidities and complications in Total Knee Arthroplasty Accurate?** *J Arthroplasty* 2002, 17:767-772.
- Bleichrodt H, Crainich D, Eekhoudt L: **The effect of comorbidities on treatment decisions.** *J Health Econ* 2003, 22:805-820.
- Roche JJ, Wenn RT, Sahota O, Moran CG: **Effect of comorbidities and postoperative complications on mortality after hip fracture in elderly people: prospective observational cohort study.** *BMJ* 2005, 331:1374.
- Paavolainen P, Pukkala E, Pulkkinen P, Visuri T: **Causes of death after total hip Arthroplasty.** *J Arthroplasty* 2002, 17:274-281.

**Tuominen U, Sintonen H, Hirvonen J, Hietaniemi K, Lehto MUK, Paavolainen P, Seitsalo S, Blom M. The effect of waiting time on health outcomes and use and costs of medication in hip replacement patients. A randomised clinical trial. *Osteoarthritis and Cartilage* 2009; 17: 1144–1150.**





## The effect of waiting time on health and quality of life outcomes and costs of medication in hip replacement patients: a randomized clinical trial

U. Tuominen<sup>†,‡,\*</sup>, H. Sintonen<sup>‡,§</sup>, J. Hirvonen<sup>†</sup>, S. Seitsalo<sup>||</sup>, P. Paavolainen<sup>||</sup>, M. Lehto<sup>¶</sup>,  
K. Hietaniemi<sup>#</sup> and M. Blom<sup>‡,††</sup>

<sup>†</sup> National Research and Development Centre for Welfare and Health, Helsinki, Finland

<sup>‡</sup> Department of Public Health, University of Helsinki, Finland

<sup>§</sup> FinOHTA, Helsinki, Finland

<sup>||</sup> Orton Orthopaedic Hospital, Helsinki, Finland

<sup>¶</sup> Hospital District of Pirkanmaa, Tampere, Finland

<sup>#</sup> Hospital District of Helsinki and Uusimaa, Finland

<sup>††</sup> Division of Social Pharmacy, Faculty of Pharmacy University of Helsinki, Finland

### Summary

**Objective:** To evaluate the effect of waiting time on health and quality of life outcomes and costs of medication in total hip replacement (THR) patients in a randomized clinical trial.

**Methods:** 395 THR patients were recruited into the study. When placed on the waiting list, patients were randomized into a short ( $\leq 3$  months) or a non-fixed waiting time (NFWT) ( $> 3$  months) group. In the final analyses 309 patients (179 women) with a mean age of 65 years were included. Health-related quality of life (HRQoL) (generic 15D), and pain and function (modified Harris Hip Score (HHS)) were calculated when placed on the waiting list, at hospital admission, and at 3 and 12 months postoperatively. The costs of disease-specific medication were calculated at the same measurement points. All analyses were performed using the intention-to-treat (ITT) principal.

**Results:** Of the recruited patients, 309 (78%) completed the follow-up (short group 140 and non-fixed group 169 patients). The mean waiting time was 74 days in the short and 194 days in the NFWT groups. In the ITT analyses there were no statistically significant differences between the groups in the weekly use and costs of medication, HRQoL or HHS at baseline, at admission, or 3 or 12 months after surgery. The only difference was in total medication costs during the waiting time period, at EUR 83 and 171, respectively.

**Conclusions:** The length of the waiting time did not generate different effects on the studied health and quality of life outcomes of the randomized groups. However, those in short waiting time group reached earlier better HRQoL.

© 2009 Osteoarthritis Research Society International. Published by Elsevier Ltd. All rights reserved.

**Key words:** Waiting time, Randomization, HRQoL, 15D, Harris Hip Score, Medication.

### Introduction

According to the Health 2000 Health Examination Survey in Finland, 12% of the population aged over 65 had hip arthritis<sup>1</sup>, while 8460 total hip replacements (THRs) were performed in 2004 with the median waiting time to surgery being 153 days. Between 1987 and 2004, the THR rate rose on average 5% annually<sup>2</sup>. In 2005, the Finnish Social Insurance Institution paid EUR 85 million in drug reimbursements to 990,637 patients with musculoskeletal diseases, with the mean cost per patient being euro (EUR) 86 per year<sup>3</sup>.

Earlier studies have established that arthritis causes pain, reduces the range of motion, and creates difficulties in participating in daily activities, which in turn affect quality of life. Because the waiting times for surgery are long, the disease becomes an increasingly chronic burden to patients<sup>4–7</sup>. Moreover, the duration of conservative treatment and the use of medication (analgesics and anti-inflammatory drugs) increase with longer waiting time. Prior studies have demonstrated that poorer health status preoperatively is predictive of higher out-of-pocket costs for patients during the first year postoperatively<sup>8</sup>, and waiting more than 6 months is associated with higher mean total costs while longer waiting times results in deterioration in physical function while waiting<sup>7</sup>.

The aim of this prospective randomized controlled trial was to identify the effects of waiting time on health and health-related quality of life (HRQoL) outcomes and the use and costs of disease-specific medication (DSM) among two patients group; a short waiting time group (SWT) ( $\leq 3$  months) and a non-fixed waiting time (NFWT) group ( $> 3$  months). Measurements took place when first placed on the waiting list, at admission, and 3 and 12 months after the THR.

\*Address correspondence and reprint requests to: Ulla Tuominen, National Research and Development Centre for Welfare and Health, Lintulahdenkuja 4, FI-00531 Helsinki, Finland. Tel: 358-503399646; Fax: 358-939672485; E-mail: ulla.tuominen@thl.fi, harri.sintonen@helsinki.fi, johanna.hirvonen@thl.fi, seppo.seitsalo@invalidisaatio.fi, pekka.paavolainen@invalidisaatio.fi, matti.lehto@pshp.fi, kalevi.hietaniemi@hus.fi, marja.blom@helsinki.fi

Received 19 August 2008; revision accepted 18 March 2009.

To our knowledge no prior studies have looked at the effect of waiting time on HRQoL and the use and costs of DSM among hip replacement patients in a randomized study design. The question of whether the length of waiting time for THR affects the use and costs of medication and health and quality of life outcomes is a contested issue.

## Methods

### DATA COLLECTION

Between August 2002 and November 2003, 513 THR patients in three Finnish hospitals were invited to participate in the study: two of the hospitals (the Surgical Hospital and Jorvi Hospital) are part of the Helsinki University Central Hospital, while the third is the Coxa Hospital for Joint Replacement. Patients were recruited into the study through contact with orthopedic and practice staff.

The key inclusion criteria were the need for a primary THR due to osteoarthritis of the hip joint as evaluated by the hospital surgeon, the patient was aged 16 or older and placed on the waiting list in a research hospital, and the patient was willing and mentally able to participate in the study. The key exclusion criteria were patients with rheumatoid arthritis, fractures, and congenital haemophilia or congenital deformities.

### RANDOMIZATION

When placed on the hospital waiting list, patients were randomly assigned to one of two groups: (1) a SWT with a maximum wait of 3

months, or (2) a NFWT with surgery performed according to the hospital's routine procedure, with the waiting period measured from the date the patient was added to the waiting list to the date of admission for surgery. The number of patients placed on the waiting list varied from 1 month to another, being specific to each hospital. Therefore, no advance estimate could be made of the number of patients to be placed on the list. The patients randomized into the SWT group could only be operated on in one of four operating periods during the year, and only half of the hospital's 1-month surgical capacity could be allocated to the SWT group, so the number of SWTs was restricted and determined specifically for each hospital.

Patients were recruited into the study in over three (for one of the hospitals) or four recruitment periods (for the other two hospitals), each period lasting 3 months in order to avoid the waiting time for the SWT group exceeding 3 months. Patients in the SWT group were operated within 2 weeks following the end of each recruitment period. The size of the NFWT group was not restricted so as to ensure that all eligible patients placed on the waiting list had an opportunity to be recruited to the study. Randomization took place during the whole recruitment period (3 months), allowing everyone the possibility of a SWT. The two patient groups therefore differed in size.

Computer-generated randomization sequences were produced by the National Research and Development Centre for Welfare and Health and supplied to the hospitals using consecutively numbered and sealed opaque envelopes. The patient's named nurse assigned participants to their groups after the decision for surgery had been made. The randomization envelopes contained information on whether the patient belonged to the SWT or NFWT group. Surgeons were blind to patient allocation. For ethical reasons double-blinding was not possible.

Patients completed a self-administered questionnaire when placed on the waiting list, at admission, and at 3 and 12 months postoperatively. The

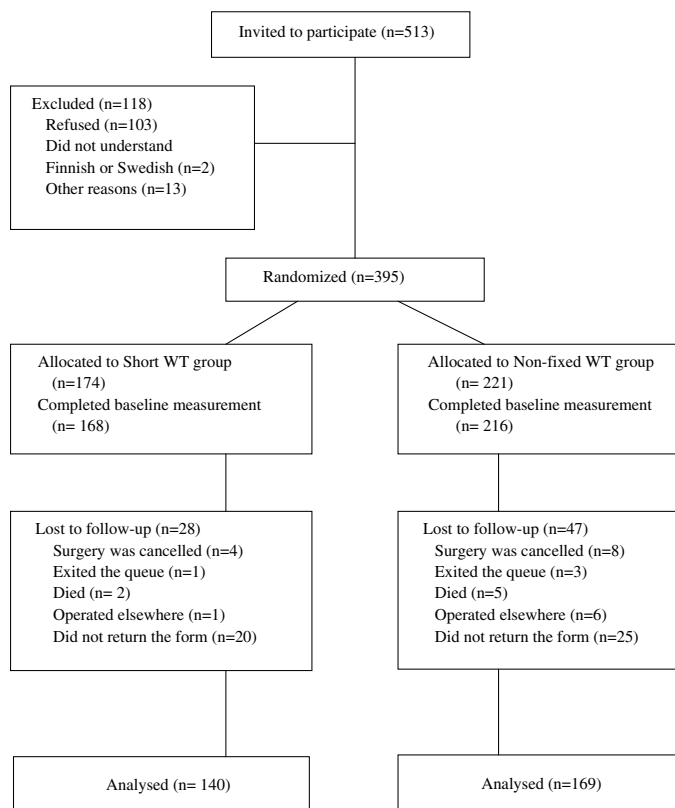


Fig. 1. Flow of patients through the trial.

Table I  
Baseline characteristics of randomized groups and those lost to follow-up

Characteristics at baseline	SWT ( <i>n</i> = 145)	NFWT ( <i>n</i> = 175)	Lost to follow-up ( <i>n</i> = 64)	<i>P</i> -value*
Age, years (mean ± SD)	66 ± 9.8	64 ± 10.1	68 ± 10.1	0.053
Females [ <i>n</i> , (%)]	88 (49.2)	91 (52.3)	38 (62.3)	0.539
Housing [ <i>n</i> , (%)]				
Living alone	50 (33.6)	45 (26.5)	20 (33.9)	0.313
Basic education [ <i>n</i> , (%)]				
Lower level	118 (79.2)	133 (78.7)	46 (78.0)	0.494
Comorbidity, yes [ <i>n</i> , (%)]	103 (69.1)	116 (68.2)	40 (67.8)	0.504
Medication to arthritis, yes [ <i>n</i> , (%)]	136 (91.3)	149 (87.6)	56 (94.9)	0.136
Medication costs €/week ± SD	5.4 ± 5.8	6.3 ± 7.0	5.5 ± 7.2	0.997
BMI <sup>‡</sup> ± SD	27.9 ± 3.9	27.7 ± 3.8	27.7 ± 4.6	0.865
HHS <sup>‡</sup> ± SD	43.9 ± 13.5	44.09 ± 14.4	44.3 ± 12.9	0.853
Pain score ± SD	16.8 ± 7.6	17.4 ± 6.8	16.2 ± 7.0	0.388
Function score ± SD	27.1 ± 8.6	27.2 ± 9.0	27.7 ± 8.1	0.612
15D score** (mean ± SD)	0.767 ± 0.09	0.764 ± 0.12	0.753 ± 0.01	0.295

\*Between patients who completed all questionnaires and those lost to follow-up.

‡Body mass index (kg/m<sup>2</sup>).

‡The scale 0–91, worst to best.

\*\*The scale 0–1, worst to best.

questionnaires were either distributed to the patients at the hospital or in some cases mailed to patients, as happened with one hospital for the third and fourth questionnaires. All questionnaires were returned by post. Common guidelines for administering the questionnaires were provided in each hospital. The patients completed a sociodemographic form, reported their medication and comorbidities as diagnosed by a medical doctor, and completed the disease-specific modified Harris Hip Score (HHS) and a separate questionnaire for HRQoL. Each patient provided informed consent. The study was approved by the Helsinki University Central Hospital Surgery Ethics Committee.

#### MEASUREMENT INSTRUMENTS

HRQoL was measured by the generic 15D instrument. The 15D is composed of 15 dimensions: moving, vision, hearing, breathing, sleeping, eating, speech, eliminating, vitality, usual activities, mental function, discomfort and symptoms, depression, distress, and sexual activity. Each dimension has five ordinal levels to choose from. The 15D can be used as a profile measure or to give a single index score by means of population-based preference weights. The index score (15D score) ranges from 0 (dead) to 1 (completely

healthy)<sup>9</sup>. Completing the 15D questionnaire takes 5–10 min and it describes the HRQoL of the respondent at present. A difference of >|0.03| in the 15D score is clinically important in the sense that on average people can feel the difference<sup>10</sup>. In most of the important properties (reliability, content validity, sensitivity in terms of discriminatory power and responsiveness to change) the 15D compares at least equally with other similar instruments that produce a valuation-based single index number<sup>10–14</sup>.

By using the mean 15D scores from each measurement point and assuming a linear change in the scores between the measurement points we also estimated the possible gain in quality-adjusted life years (QALY gain) for both groups within the observation period.

The disease-specific modified HHS was used to measure hip pain and function. The self-report HHS consists of two sections: pain (0–44 points) and functional activities of daily living and gait (0–47). The total score ranges from 0 to 91, with higher scores representing better health states<sup>15</sup>. A further 9 points of the total HHS would normally be assigned to describing deformity and range of motion, but were excluded because these items could not be measured in a patient self-report questionnaire. The performance of self-reported HHS is comparable to that of surgeon-assessed HHS and has shown to be less burdensome to patients than physician-administered HHS<sup>16</sup>. The HHS and 15D score serve as measures of disease severity at baseline (preoperatively).

The use of DSM (analgesics and anti-inflammatory drugs) during the week preceding every measurement point was measured based on self-report. The unit costs of medication per tablet were obtained from CD-Pharmacy<sup>17</sup>. The costs of medication during the waiting period were calculated as a product of the weeks spent on the waiting list and the medication costs per week. The calculations for medication costs were made at each of the four measurement points.

#### STATISTICAL ANALYSIS

The sample size estimate was based on the primary outcome variable 15D. A subgroup of 177 patients would provide the 80% power (two-tailed  $\alpha$  error 5%) to detect clinically important differences ( $\Delta 0.03$ ) in the 15D score between the randomized groups.

Primary analyses were conducted with an intention-to-treat (ITT) principle<sup>18</sup>, so that patients were followed in the groups to which they were randomly allocated. Also a supplementary per-protocol analysis was carried out at admission by excluding the patients from SWT group who were admitted beyond the short waiting time (waiting time > 3 months). When comparing the mean scores of 15D, total HHS, pain, function and costs of DSM at the various points of follow-up between SWT and NFWT, general linear model for repeated measures was used. Univariate (tests of within-subject effects) *P*-values were calculated, while a *P*-value < 0.05 was considered statistically significant. Also the mean differences, standard deviations and confidence intervals (CIs) of 95% of the variables 15D, total HHS, pain, function and costs of DMS were presented for the estimated effect.

Descriptive characteristics at baseline in the randomized groups and the patients who were lost to follow-up were compared using either the *F*-test or

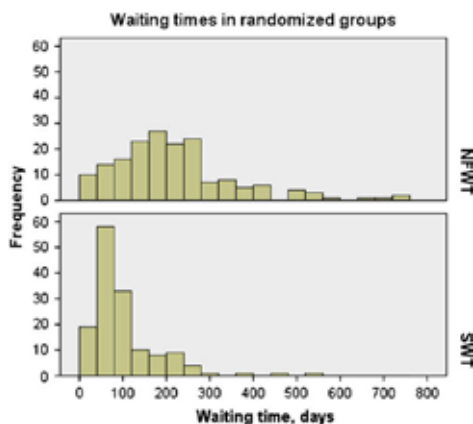


Fig. 2. The distribution of patients by waiting time in the randomized groups.

Table II  
Outcomes at admission, ITT-analysis

Outcome measure	SWT mean ( $\pm$ SD) ( <i>n</i> = 139–140)	NFWT mean ( $\pm$ SD) ( <i>n</i> = 167–169)	Mean difference	95% CI	<i>P</i> -value
15D score*	0.768 (0.107)	0.769 (0.107)	0.001	–0.022 to 0.024	0.931
Self-report HHS†	42.99 (15.32)	41.73 (14.08)	–1.262	–5.233 to 1.816	0.456
HHS pain score	17.63 (7.64)	17.14 (8.55)	–0.493	–2.332 to 1.346	0.602
HHS function score	25.86 (9.08)	24.79 (8.42)	–1.067	–3.051 to 0.915	0.29
Costs of medication‡	5.56 (6.80)	5.63 (6.22)	0.793	–1.399 to 1.557	0.916
Costs of medication**	81.3 (129.54)	171.3 (264.36)	88.49	42.49 to 134.90	0.000***
Waiting time, days	74 (145)	194 (175)	117.80	93.03 to 142.57	0.000***

\*\*\**P* < 0.001.

\*The scale 0–1, worst to best.

†The scale 0–91, worst to best.

‡Weekly costs of medication in EUR (EUR is 1.6 US\$).

\*\*Medication costs during the waiting time.

the chi-squared test, depending on whether the variable was on a continuous or nominal scale.

Missing values on the 15D were replaced if a minimum of 80% of dimensions had been completed, using a regression model with the patient's responses for other dimensions, age and gender as explanatory variables<sup>11</sup>. Data analyses were performed using SPSS versions 14 and 16 for Windows.

## Results

Of the 513 eligible patients invited to participate in the study, 118 refused to participate and were excluded. Their mean age was 70 years (SD  $\pm$  11) and 64% were women. The most frequently quoted reasons for refusal were being too tired or unwillingness to complete questionnaires. Thus, 395 patients after providing a signed informed consent were randomly allocated to either the SWT (*n* = 174) or NFWT (*n* = 221) group. Of these patients, 86 were lost to follow-up during the waiting time and were not included in the final analyses. Of these patients 11 did not return the questionnaire at baseline, although they had signed informed consent and were randomized, seven were operated on elsewhere, operations for 12 patients were canceled, and seven died while waiting and 45 did not return the questionnaire at admission. All analyses are based on 309 (78%) patients (179 women) with a mean (SD) age of 65 ( $\pm$ 9.9; range; from 33 to 87) years, of which 140 were in the SWT and 169 in the NFWT group (Fig. 1).

### BASELINE CHARACTERISTICS

The baseline characteristics of the groups were similar and are reported in Table I. The mean ( $\pm$ SD) 15D score in the SWT group was 0.767 ( $\pm$ 0.09) and 0.764 ( $\pm$ 0.12)

in the NFWT group; the difference was not statistically significant or clinically important (*P* = 0.295). The mean ( $\pm$ SD) total HHS was 43.9 ( $\pm$ 13.6) and 44.1 ( $\pm$ 14.1) in the SWT and NFWT group, respectively (*P* = 0.853). The percentage of patients receiving DSM was 91% (*n* = 136) in the SWT and 89% (*n* = 149) in the NFWT group. The mean weekly medication costs were EUR 5.4 and EUR 6.3 in the SWT and NFWT groups, respectively. The difference was not statistically significant (Table I).

A comparison between patients who completed the questionnaire and those who were lost to follow-up showed no statistically significant differences between the groups (Table I).

### OUTCOMES AND WAITING TIME

At admission there was a statistically significant difference in mean waiting time between the groups: 74 ( $\pm$ 145) days in the SWT and 194 ( $\pm$ 175) days in the NFWT group (95% CI: 93.03–142.57, *P* < 0.001) (Fig. 2). There were no statistically significant differences between the groups in the mean 15D score, total HHS, pain and function or in the weekly cost of medication. However, due to a shorter waiting time the total cost of DSM during the waiting period was lower in the SWT group (EUR 81.3) than in the NFWT group (EUR 171.3) (95% CI: 42.49–134.90, *P* < 0.001) (Table II).

A per-protocol analysis was performed as a supplementary analysis. In the SWT group (*n* = 92), those compliant with allocated waiting time 59 ( $\pm$ 21) days and all patients (*n* = 170) in the NFWT group were included in the analysis. Similar results were obtained in a per-protocol analysis as in ITT analysis (Table III).

Table III  
Outcomes at admission, per-protocol analysis

Outcome measure	SWT mean ( $\pm$ SD) ( <i>n</i> = 92)*	NFWT mean ( $\pm$ SD) ( <i>n</i> = 170)	Mean difference	95% CI	<i>P</i> -value
15D-score	0.764 (0.11)	0.770 (0.11)	0.004	–0.023 to 0.032	0.733
Self-report HHS	42.95 (16.27)	41.73 (14.0)	–1.251	–5.198 to 2.766	0.548
HHS pain score	17.42 (8.18)	17.14 (8.56)	–0.283	–2.432 to 1.866	0.795
HHS function score	25.90 (9.47)	24.80 (8.42)	–1.106	–3.442 to 1.230	0.351
Costs of medication‡	5.22 (6.74)	5.63 (6.22)	0.41	–1.277 to 2.098	0.632
Waiting time days	59.28 (21.40)	218.50 (140.35)	159.211	137.79 to 180.635	0.000***

\*\*\**P* < 0.001.

\*49 SWT patients admitted beyond SWT were excluded.

‡Weekly costs of medication in EUR (EUR is 1.6 US\$).

Table IV  
Descriptive statistics of the outcomes in SWT and NFWT groups; mean and standard deviation

Outcome	At baseline		At admission		3 months postoperatively		1-year postoperatively	
	SWT	NFWT	SWT	NFWT	SWT	NFWT	SWT	NFWT
15D								
Mean	0.770	0.779	0.772	0.775	0.852	0.859	0.857	0.873
±SD	0.090	0.100	0.090	0.090	0.090	0.100	0.110	0.110
Total HHS								
Mean	43.75	43.05	43.55	41.82	67.02	67.12	72.18	74.57
±SD	13.78	14.58	15.27	14.48	16.54	17.15	16.91	16.9
Function score								
Mean	26.68	26.19	26.11	24.93	34.93	35.05	39.05	39.24
±SD	8.73	8.95	8.97	8.56	8.41	9.04	7.67	8.5
Pain score								
Mean	16.82	17.43	17.63	17.13	32.4	32.75	33.57	35.1
±SD	7.6	6.8	7.6	8.55	10	9.5	9.78	10.06
Costs of medication (€)								
Mean	5.06	6.41	5.31	5.57	1.59	1.96	0.8	0.98
±SD	5.78	7.23	6.73	6.4	4.12	4.47	2.45	2.71

#### OUTCOMES AFTER THR

The use and costs of medication had decreased in both groups at 3 months and 1 year after the THR: 43% ( $n=55$ ) in the SWT and 48% ( $n=71$ ) in the NFWT group reported using medication for arthritis. In the SWT group the weekly costs at 3 months were EUR 1.58 and in the NFWT group EUR 1.96 and after 1 year EUR 0.80 and EUR 0.98, respectively. The differences were statistically significant between different measurement points ( $F=63.08$ ;  $P<0.001$ ), but not between the randomized groups ( $F=1.08$ ;  $P=0.372$ ).

At 3 months the mean 15D score was 0.852 in the SWT and 0.859 in the NFWT group, and after 1 year 0.856 and 0.873, while the mean total HHS was after 3 months 67.02 and 67.12, respectively and after 1 year 72.18 and 74.57, respectively. The differences between the groups are not statistically significant (Table IV). There were statistically significant differences between different measurement points, but not between the randomized groups (Table V).

There was an improvement from 3 months to 12 months postoperatively in the mean 15D score and total HHS, pain and function and a decrease in the cost of medication in both groups. SWT resulted in a gain of 0.028 QALYs

(area between the curves in Fig. 3) during the observation period.

#### Conclusions

Scientific evidence on the relationship between waiting time and THR outcomes is inconsistent and the absence of randomized trials has prevented an assessment of whether longer waiting is somehow related to health and quality of life outcomes. To our knowledge, the present study is the first to assess the use and costs of DSM in THR patients randomly allocated to short and NFWT groups. The study also analyzed whether the length of the waiting time was related to HRQoL, hip pain and function as measured by the 15D and HHS, respectively.

The main finding was that the time spent on the waiting list was not related to the weekly use and costs of DSM, with no differences found between the randomized groups. The weekly costs were almost identical in both groups at each of the four measurement points. In an earlier study among these patients it was found that they used very little health and social services during the waiting time and only a minority of patients received visits to their home from nurse, chiroprapist or physician with no statistically

Table V  
Effect of the waiting time in the randomized groups. General linear model, repeated measures analysis, tests of within-subjects effect

Measurement instrument	Source	F	P-value
15D	Time	207.324	0.000*
	Time × randomized group	0.694	0.531‡
HHS	Time	470.763	0.000*
	Time × randomized group	1.429	0.239‡
Function score	Time	308.701	0.000*
	Time × randomized group	0.703	0.511‡
Costs of medication	Time	86.104	0.000*
	Time × randomized group	1.04	0.365‡

\*Time effect – differences between the four measurement points.

‡Group effect – differences between randomized groups.

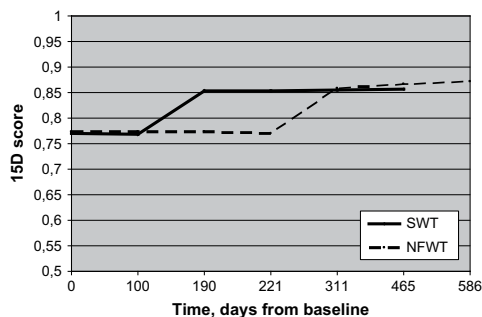


Fig. 3. The QALY gain (area between the curves) due to short waiting time during the observation period.

significant difference between the randomized groups<sup>19</sup>. Similarly, the mean HRQoL scores and pain and function scores were almost identical in both groups at each of the four measurement points, and reached their highest values 12 months postoperatively. However, the SWT resulted in a small QALY gain of 0.028, and reached better quality of life 3 months earlier than those in the NFWT group.

In this study, patients' HRQoL at baseline was deteriorated and they had pain and difficulties in functioning and daily activities. However, the length of waiting time alone did not affect HRQoL or pain and function scores, which is in line with the findings of some earlier studies<sup>4,5,7,20,21</sup>. However THR has been found to be effective<sup>12,13,22,23</sup>, with this study also showing significant improvement in HRQoL and in pain and function 3 and 12 months postoperatively in both randomized groups.

To our knowledge no studies have so far been published on the effect of waiting time on the use and costs of DSM. Fielden *et al.*<sup>6</sup> estimated the costs and outcomes of the waiting time for THR in a prospective study. They found that waiting more than 6 months was associated with higher total costs and deterioration in physical function while waiting. These cost results are not comparable to ours, as they estimated all costs during the waiting period<sup>6</sup>.

The strengths of this study were that the patients awaiting THR were prospectively followed from the time of being placed on the waiting list to admission—with waiting times recorded precisely—and for a further year of follow-up postoperatively, providing evidence of the effect of WT on pre- and postoperative health status. Further the patients were randomly assigned to either a SWT or NFWT group. The findings were based on the simultaneous use of patient-reported generic and disease-specific instruments as outcome measures. Also the results based on ITT and per-protocol analyses indicated that there were no statistical significant differences in health outcomes between the randomized groups. This was further tested with repeated measures analyses. The within-subject test indicated that there was a significant time effect, i.e., the outcomes did change over time, but the changes were identical, which is consistent with the finding that the interaction was not significant.

Some limitations pertain to this study. First, patients who refused to participate the study were older than those in the study groups. Second, medication information was obtained only from patient self-reports, while there was no distinction between self-care and prescription medication. For temporary medication we used mean dosages. Third, a total of 49 patients in the SWT group waited more than 3 months. The reasons were hospitals' limited capacity to carry out THR within the 3 months waiting time period or the patient's unwillingness to have THR within 3 months. Due to these factors, the differences between the randomized groups may have been underestimated. However the primary analysis was based on the ITT principle to address the question of clinical effectiveness and to avoid the bias associated with a non-random loss of participants. In addition the supplementary per-protocol analysis at admission where the SWT patients who were admitted beyond SWT were excluded from the analysis — supported the main findings and did not show a statistically or clinically important difference in HRQoL between randomized groups.

According to this study the length of the waiting time, at least as realized in practice in this study, did not result in different health outcomes in the randomized groups in three

cross-sectional follow-up measurements from baseline to 1-year postoperatively. The length of the waiting time reflected in the total medication costs of the waiting period and in a small QALY gain in the SWT, as it reached the same HRQoL level as the NFWT group on average 3 months earlier.

### Conflict of interest

The authors declare that they have no competing interests.

### Acknowledgments

This study was financially supported by the Academy of Finland (no. 51871), HUS HUCH Jorvi Hospital, Coxa Hospital for Joint Replacement, Medical Research Fund of Tampere University Hospital, HUS HUCH Surgical Hospital, Orton Orthopaedic Hospital and Yrjö Jahnsson Foundation.

### References

1. Aromaa A, Koskinen S, Eds. Health and Functional Capacity in Finland. Baseline Results of the Health 2000 Health Examination Survey. Helsinki, 2004.
2. National Research and Development Centre for Welfare and Health. Operative Inpatient Service, [http://www.stakes.info/0\\_1\\_7.asp](http://www.stakes.info/0_1_7.asp); 2006.
3. National Health Insurance Institution. Available in: <http://www.kela.fi>; 2006.
4. Kill S, Wright I, Jones RS. Change in Harris hip score in patients on the waiting list for total hip replacement. *Ann R Coll Surg Engl* 2001;82:2.
5. Croft P, Lewis M, Wynn Jones C, Coggon D, Cooper C. Health status in patients awaiting hip replacement for osteoarthritis. *Rheumatology* 2002;41:9.
6. Hajat S, Fitzpatrick R, Morris R, *et al.* Does waiting for total hip replacement surgery matter? Prospective cohort study. *J Health Serv Res Policy* 2002;7.
7. Fielden JM, Cumming JM, Horne JG, Devane PA, Slack A, Gallagher LM. Waiting for hip arthroplasty economic costs and health outcomes. *J Arthroplasty* 2005;20:8.
8. March L, Cross M, Tribe K, Lapsley H, Courtenay B, Brooks P. Costs of joint replacement surgery for osteoarthritis: the patients' perspective. *J Rheumatol* 2002;29:1006.
9. Sintonen H. The 15D measure of health-related quality of life. Properties and applications. *Ann Med* 2001;33:328–36.
10. Sintonen H. Outcome measurement in acid-related diseases. *Pharmacoeconomics* 1994;5(Suppl 3):17–24.
11. Sintonen H. The 15D Measure of Health-related Quality of Life. II Feasibility and Validity of Its Valuation System. Melbourne: National Centre for Health Program Evaluation 1995. Working paper 42, Available from: <http://www.buseco.monash.edu.au/centres/che/publication.php#4>.
12. Rissanen P, Aro S, Sintonen H, Asikainen K, Sliätis P, Paavola P. Costs and cost-effectiveness in hip and knee replacements. A prospective study. *Int J Technol Assess Health Care* 1997;13:575–88.
13. Brazier J, Derveril MI. The use of health-related quality of life measures in economic evaluation. In: *Health Services Research Methods. Guide to Best Practice*. London: BMJ Books; 1998.
14. Stavem K. Reliability, validity and responsiveness of two utility measures in patients with chronic obstructive pulmonary disease. *Qual Life Res* 1999;8(1–2):45–54.
15. Harris W. Traumatic arthritis of hip after dislocation and acetabular fractures: treatment by mold arthroplasty. *J Bone Joint Surg Am* 1969; 51:4.
16. Hawthorne G, Richardson J, Day NA. A comparison of the assessment of quality of life (AQoL) with four other generic utility instruments. *Ann Med* 2001;33:358–70.
17. Pharmaceutical Information Centre in Finland: *Pharmaca Fennica CD-rom* 1, 2004.
18. Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ* 1999;319: 670–4.
19. Hirvonen J, Blom M, Tuominen U, Seitsalo S, Lehto M, Paavola P, *et al.* Is longer waiting time associated with health and social services utilization before treatment? A randomized study. *J Health Serv Res Policy* 2007 Oct;12(4):209–14.

20. Kelly KD, Voaklander DC, Johnston DWC, Newman SC, Suarez-Almazor ME. Change in pain and function while waiting for major joint arthroplasty. *J Arthroplasty* 2001;16(3):351–9.
  21. Hirvonen J, Blom M, Tuominen U, Sintonen H. The use of Patients-Reported Outcomes (PROs) in patients who underwent major joint replacement and those who left the queue. *Patient Reported Outcomes. Newsletter* 2006;36:31–2.
  22. Räsänen P, Paavolainen P, Sintonen H, *et al.* Effectiveness of hip or knee replacement surgery in terms of quality-adjusted life years and costs. *Acta Orthopaedica* 2007;78(1):108–15.
  23. Wilke RJ, Burke LB, Erickson P. Measuring the treatment impact: a review of patients-reported outcomes and other efficacy endpoints in approved product labels. *Control Clin Trials* 2004;25: 535–52.
-





**Tuominen U, Sintonen H, Hirvonen J, Hietaniemi K, Lehto MUK, Paavolainen P, Seitsalo S, Blom M. Is longer waiting time for total knee replacement associated with health outcomes and medication costs? A randomized clinical trial. *Value in Health* 2010; 13 (8): 998–1004.**



# Is Longer Waiting Time for Total Knee Replacement Associated with Health Outcomes and Medication Costs? Randomized Clinical Trial

Ulla Tuominen, PhD,<sup>1,2</sup> Harri Sintonen, PhD,<sup>2</sup> Johanna Hirvonen, PhD,<sup>3</sup> Seppo Seitsalo, MD, PhD,<sup>4</sup>  
Pekka Paavolainen, MD, PhD,<sup>4</sup> Matti Lehto, MD, PhD,<sup>5</sup> Kalevi Hietaniemi, MD,<sup>6</sup> Marja Blom, PhD<sup>7</sup>

<sup>1</sup>National the Social Insurance Institution of Finland, Helsinki, Finland; <sup>2</sup>Department of Public Health, University of Helsinki, Helsinki, Finland; <sup>3</sup>Mikkeli University of Applied Sciences, Mikkeli, Finland; <sup>4</sup>Orton Orthopaedic Hospital, Helsinki, Finland; <sup>5</sup>Hospital District of Pirkanmaa, Tampere, Finland; <sup>6</sup>Hospital District of Helsinki and Uusimaa, Helsinki, Finland; <sup>7</sup>Division of Social Pharmacy, Faculty of Pharmacy University of Helsinki, Helsinki, Finland

## ABSTRACT

**Background:** The aim of this prospective randomized study was to evaluate the effect of waiting time (WT) on health-related quality of life (HRQoL), knee pain and physical function, and the use and costs of medication of patients awaiting total knee replacement.

**Methods:** When placed on the waiting list, 438 patients were randomized into a short waiting time (SWT  $\leq 3$  months) or a nonfixed waiting time (NFWT  $> 3$  months) group. HRQoL was measured by the 15D, and pain and physical function by modified Knee Society Clinical Rating System at baseline, admission, and 3 and 12 months postoperatively. The costs of medication due to osteoarthritis were calculated at the same measurement points. All analyses were performed using the intention-to-treat principle.

**Results:** The mean WT was 94 and 239 days in the SWT and NFWT groups, respectively. Apart from higher weekly cost of medication in the SWT group at admission and better HRQoL in the NFWT group 1 year postoperatively, there were no statistically significant differences between the groups in other outcomes during the follow-up.

**Conclusion:** Those in the SWT group had higher weekly costs of medication at admission, and reached better HRQoL 3 months earlier than those in the NFWT group, but the latter had better HRQoL after operation. Otherwise, the length of WT was not associated with different health and HRQoL outcomes in the groups.

**Keywords:** cost, health-related quality of life, osteoarthritis, randomized clinical trial, waiting lists.

## Introduction

Osteoarthritis (OA) is the most common cause of musculoskeletal disability and pain in the world. In Finland, 6% of men and 8% of women over the age of 30 years suffer from clinically diagnosed knee OA. The incidence of radiological and clinical knee arthritis increased with age both in men and women [1]. During 2004, a total of 5905 total knee replacements (TKRs) were performed with the median waiting time (WT) to surgery being 209 days. According to statistics, the number of TKRs increased to 9033 and median WT decreased to 149 in 2006 [2]. In 2005, the Finnish Social Insurance Institution paid EUR 85 million in drug reimbursements to 990,637 patients with musculoskeletal diseases, with the mean cost being EUR 86 per year [3].

Patients with OA experience increasing pain and progressive loss of physical function, walking, and stair climbing [4–8]. Earlier studies have established that arthritis causes difficulties in participating in daily activities, which in turn affect quality of life [9–12]. Physiotherapy and pharmacological treatment (paracetamol and nonsteroidal anti-inflammatory drugs) are used to reduce the symptoms, namely joint pain, stiffness, and swelling. But, if OA pain is not otherwise manageable, patients' ability to manage everyday tasks is essentially compromised because of OA, or if there is significant restriction of motion, or a joint malposition, a TKR should be performed [1].

In many Western countries, WTs for TKR are long [13]. In discussion on health policy in Finland, long WTs are regarded as a major problem for several reasons. They are claimed to violate the principle of equal access to treatment, which is a key performance indicator of health care in Finland. They are also claimed to result in extra suffering, extra cost to the patients and the society, and poorer treatment outcomes. Nevertheless, it is unclear to what extent these claims are true as there is little solid evidence to substantiate them.

To our knowledge, no prior studies have looked at the effect of WT on health-related quality of life (HRQoL) and the costs of disease-specific medication (DSM) among knee replacement patients in a randomized study design. The aim of this prospective randomized study was to identify the effects of WT on health and quality-of-life outcomes, and the use and costs of disease-specific medication among two different patient groups: a short WT group ( $\leq 3$  months) and a nonfixed WT group ( $> 3$  months). The question of whether the length of WT for TKR affects the costs of medication and health, and quality-of-life outcomes is a contested issue.

## Methods

### Data Collection

Between August 2002 and November 2003, 555 TKR patients in three Finnish hospitals were invited to participate in the study: two of the hospitals (the Surgical Hospital and Jorvi Hospital) are part of the Helsinki University Central Hospital, while the third is the Coxa Hospital for Joint Replacement. Patients were

Address correspondence to: Ulla Tuominen, The Social Insurance Institution, Research Department, Nordenskiöldinkatu 12 Helsinki, FI-00250 Finland. E-mail: ulla.tuominen@thl.fi  
10.1111/j.1524-4733.2010.00779.x

recruited into the study through contact with orthopedic and practice staff.

The key inclusion criteria were the need for a primary TKR due to OA of the knee joint as evaluated by the hospital surgeon, the patient was adult (aged 16 or older) and placed on the waiting list in a research hospital, and the patient was willing and mentally able to participate in the study. The key exclusion criteria were patients with rheumatoid arthritis, fractures, and congenital hemophilia or congenital deformities.

### Randomization

After being placed on the hospital waiting list, the patients were randomly assigned to one of two groups: 1) a short waiting time (hereafter SWT) with a maximum 3 months wait; or 2) a non-fixed waiting time (hereafter NFWT) with surgery performed according to the hospital's routine procedure, with the waiting period measured from the date the patient was added to the waiting list to the date of admission for surgery. The number of patients placed on the waiting list varied from 1 month to another, being specific to each hospital. Therefore, no advance estimate could be made of the number of patients to be placed on the list.

The patients were recruited into the study in three (in one hospital) or four recruitment periods (in two hospitals), each period lasting 3 months to avoid the WT for the SWT group exceeding 3 months. Patients randomized into the SWT group were operated within 2 weeks after the end of each recruitment period, and only half of the hospital's 1 month surgical capacity could be allocated to the SWT group, so the number of SWTs was restricted and determined specifically for each hospital. The size of the NFWT group was not restricted to ensure that all eligible patients placed on the waiting list had an opportunity to be recruited into the study. The two patient groups therefore differed in size.

Computer-generated randomization sequences were produced by the National Research and Development Centre for Welfare and Health, and supplied to the hospitals using consecutively numbered and sealed opaque envelopes. The patient's named nurse assigned participants to their groups after the decision for surgery had been made. The randomization envelopes contained information on whether the patient belonged to the SWT or NFWT group. Surgeons were blinded to patient allocation. For ethical reasons, double-blinding was not possible.

The patients completed a self-administered questionnaire when placed on the waiting list, at admission, and at 3 and 12 months postoperatively. The questionnaires were either distributed to the patients at the hospital or in some cases mailed to patients, as happened with one hospital for the third and fourth questionnaires. All questionnaires were returned by prepaid post. Common guidelines for administering the questionnaires were provided in each hospital. The patients completed a sociodemographic form, reported their medication and comorbidities as diagnosed by a medical doctor, and completed the disease-specific modified Knee Society Clinical Rating System (KS) and a separate questionnaire for HRQoL. Each patient provided informed consent. The study was approved by the Helsinki University Central Hospital Surgery Ethics Committee.

### Measurement Instruments

HRQoL was measured by the generic 15D instrument. The 15D is composed of 15 dimensions: moving, vision, hearing, breathing, sleeping, eating, speech, eliminating, vitality, usual activities, mental function, discomfort and symptoms, depression, distress, and sexual activity. Each dimension has five ordinal levels to

choose from. The 15D can be used as a profile measure or to give a single index score by means of population-based preference weights. The index score (15D score) ranges from 0 (dead) to 1 (completely healthy) [14]. Completing the 15D questionnaire takes 5 to 10 min, and it describes the HRQoL of the respondent at present. A difference of  $>0.031$  in the 15D score is clinically important in the sense that on average people can feel the difference [15]. The 15D was chosen for three main reasons: 1) it has been used successfully in earlier studies dealing with knee replacement and facilitates thus a comparison to the presurgery scores in these studies; 2) earlier research has shown that in most of the important properties (reliability, content validity, sensitivity in terms of discriminatory power, and responsiveness to change), the 15D compares at least equally with other similar instruments that produce a valuation-based single index number [16,17]; and 3) recent research has since confirmed that especially in the rehabilitation of musculoskeletal disorders, the 15D was at least as responsive as the SF-6D, if not better. These two instruments were anyway the most responsive ones [18].

By using the mean 15D scores from each measurement point and assuming a linear change in the scores between the measurement points, we also estimated the possible gain in quality-adjusted life-years (QALY gain) for both groups within the observation period.

The disease-specific, self-reported modified KS [19] was used to measure knee pain and physical function, because it has been found to be responsive to change [20]. Pain score (0–50 points) and function score (0–100 points) are presented separately; clinical dimensions (range of motion, stability, flexion contracture, extension lag, and alignment) could not be measured in a patient's self-administered questionnaire, and thus the total score could not be derived.

The use of DSM during the week preceding every measurement point was measured based on self-report. The unit costs of medication per tablet were obtained from CD-Pharmacy [21]. The costs of medication during the waiting period were calculated as a product of the weeks spent on the waiting list and the medication costs per week. The calculations for medication costs were made at each of the four measurement points.

### Statistical Analysis

The sample size estimate was based on the primary outcome variable 15D. A total subgroup of 177 patients would provide the 80% power (two-tailed  $\alpha$  error 5%) to detect clinically important differences  $>0.031$  in the 15D score between the randomized groups.

Primary analyses were conducted with an intention-to-treat (ITT) principle [22], so that patients were followed in the groups to which they were randomly allocated. When comparing the mean scores of 15D, knee pain, function, and costs of DSM at the various points of follow-up between SWT and NFWT, a general linear model (GLM) for repeated measures as tests of between-subject effects was used.  $P < 0.05$  was considered statistically significant. Also, the mean differences, SDs, and confidence intervals (CIs) of 95% of the variables 15D, knee pain and function, WT, and costs of DSM were presented for the estimated effect.

Baseline characteristics of the randomized groups and the patients who were lost to follow-up were compared using either the independent sample *t* test or chi-square test, depending on whether the variable was on a continuous or nominal scale.

Missing values on the 15D were replaced if a minimum of 80% of dimensions had been completed, using a regression model with the patient's responses for other dimensions, age, and

sex as explanatory variables [14]. Data analyses were performed using SPSS for Windows v.14 and v.16 (Chicago, IL).

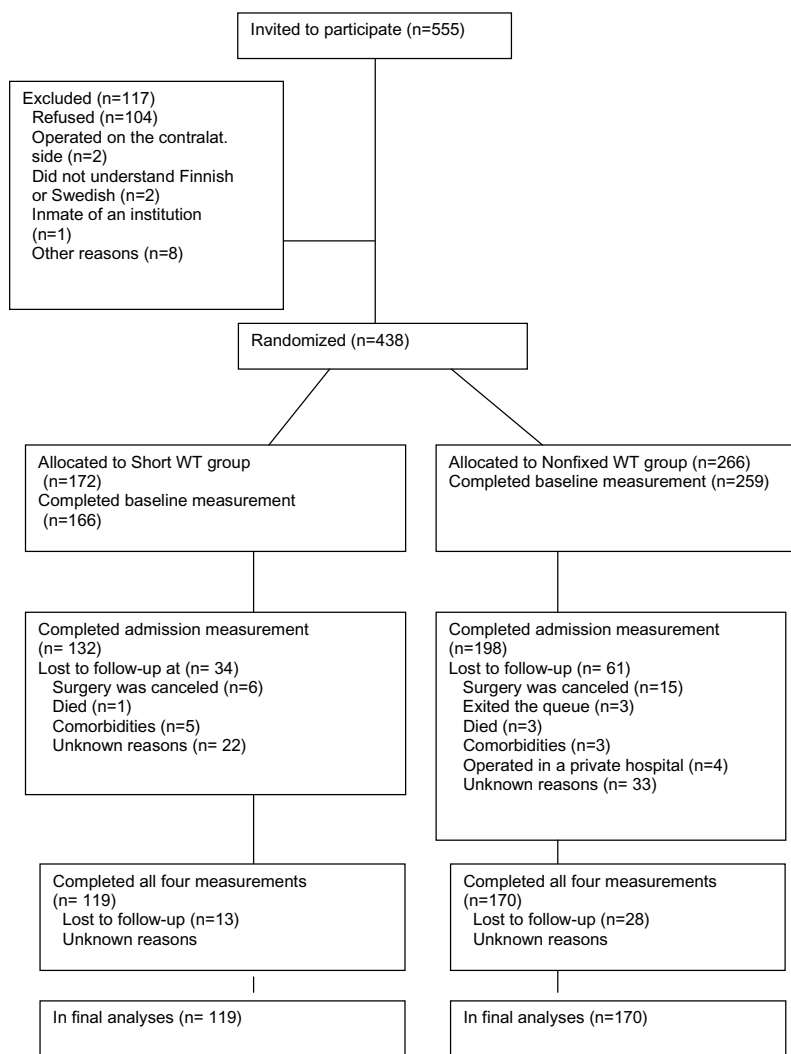
## Results

Of the 555 eligible patients invited to participate in the study, 117 patients (90 women) with a mean age of 71 refused to participate and were excluded. Thus, 438 patients after providing a signed informed consent were randomly allocated to either the SWT ( $n = 172$ ) or NFWT ( $n = 266$ ) group. Of these patients, 13 did not return the questionnaire at baseline, although they had signed informed consent. During the WT, 95 patients were lost to follow-up, 3 exited the queue, 8 had severe comorbidities,

4 were operated on in a private hospital, 29 had canceled operations, 4 died while waiting, and 45 did not return the questionnaire at admission. Primary ITT analyses are based on 330 (77%) patients (237 women) with a mean ( $\pm$ SD) age of 68 ( $\pm$ 9.9) years, of which 132 were in the SWT and 198 in the NFWT group, and the final GLM repeated analyses are based on 289 patients, who completed all four questionnaires. Of them, 119 were in the SWT and 170 in the NFWT (Fig. 1).

## Baseline Characteristics

The baseline characteristics of the groups were similar (Table 1). The mean ( $\pm$ SD) 15D score in the SWT group was 0.772



**Figure 1** Patients flow through the trial.

**Table 1** Characteristics at baseline in patients waiting for total knee replacement and those lost to follow-up

Characteristics	SWT n = 132	NFWT n = 198	Lost to follow-up n = 95	P value <sup>†</sup>	P value <sup>‡</sup>
Age, mean $\pm$ SD	67 $\pm$ 9.5	68 $\pm$ 8.9	69 $\pm$ 9.3	0.208	0.619
Sex (female), (n [%])	98 (70)	139 (69)	64 (66)	0.308	0.294
Living alone (n [%])	50 (32)	65 (37)	33 (38)	0.180	0.900
Professional education (n [%])	31 (23)	38 (19)	11 (13)	0.218	0.168
Employment status					
Retired (n [%])	107 (81)	169 (85)	71 (83)	0.050	0.834
Comorbidity (n [%])	102 (77)	154 (78)	70 (81)	0.509	0.566
Medication to arthritis (n [%])	115 (87)	175 (88)	71 (82)	0.428	0.208
Cost of medication €/week	5.72 $\pm$ 7.2	5.91 $\pm$ 6.1	4.89 $\pm$ 6.2	0.803	0.456
Body mass index (kg/m <sup>2</sup> )	29.4 $\pm$ 7.0	29.4 $\pm$ 4.4	28.55 $\pm$ 7.1	0.121	0.43
Pain score, mean $\pm$ SD <sup>§</sup>	19.9 $\pm$ 11.8	20.92 $\pm$ 12.4	19.96 $\pm$ 11.0	0.355	0.030*
Function score, mean $\pm$ SD <sup>  </sup>	48.42 $\pm$ 22.1	46.95 $\pm$ 23.3	52.38 $\pm$ 22.6	0.567	0.171
15D score, mean $\pm$ SD <sup>¶</sup>	0.772 $\pm$ 0.108	0.779 $\pm$ 0.119	0.779 $\pm$ 0.11	0.539	0.835

\* $P < 0.05$ .<sup>†</sup>Differences between the randomized groups.<sup>‡</sup>Differences between the patients who completed the questionnaire and those lost to follow-up.<sup>§</sup>The scale 0–50, worst to best.<sup>||</sup>The scale 0–100, worst to best.<sup>¶</sup>The scale 0–1, worst to best.

( $\pm 0.108$ ) and 0.779 ( $\pm 0.119$ ) in the NFWT group; the difference was not statistically significant or clinically important ( $P = 0.539$ ). The mean ( $\pm$ SD) pain score was 19.9 ( $\pm 11.8$ ) and 20.9 ( $\pm 12.4$ ) in the SWT and NFWT groups, respectively ( $P = 0.355$ ). The percentage of patients receiving DSM was 87% ( $n = 118$ ) in the SWT and 88% ( $n = 181$ ) in the NFWT group. The mean weekly medication costs were €5.7 and €5.9 in the SWT and NFWT groups, respectively. The difference was not statistically significant (Table 1).

A comparison between patients who completed the questionnaire and those who were lost to follow-up showed a statistically significant difference in pain score ( $P = 0.030$ ) (Table 1).

### Outcomes and WT

At admission, there was a statistically significant difference in mean WT between the groups: 95 ( $\pm 81$ ) days in the SWT and 239 ( $\pm 135$ ) days in the NFWT group ( $F$  value 100.6; 95% CI 116, 99–172, 22;  $P < 0.001$ ). There was a statistically significant difference between the groups in the weekly costs of DMS, at €5.33 and €3.57, respectively ( $F$  value 4.85; 95% CI 3.54–5.10;  $P = 0.029$ ). There were no differences between the randomized groups in the mean 15D, pain, and function score, and in the costs of DSM during the WT (Table 2).

### Outcomes after TKR

The use and costs of medication had decreased in both groups at 3 months and 1 year after the TKR. In the SWT group, the

weekly costs at 3 months were €3.14 and in the NFWT group €3.10, and after 1 year they were €1.74 and €2.96, respectively (Table 3). In repeated measures analyses, the differences were statistically significant between measurement points ( $F = 13.17$ ;  $P < 0.001$ ), but not between the randomized groups ( $F = 1.17$ ;  $P = 0.317$ ) (Table 4).

At 3 months, the mean 15D score was 0.813 in the SWT and 0.837 in the NFWT group, and at 1 year they were 0.813 and 0.852, respectively. The difference between the randomized groups was statistically significant ( $P = 0.012$ ) after 1 year post-operatively. The mean pain scores were at 3 months 32.7 and 34.1, respectively, and at 1 year they were 36.3 and 36.9, respectively. The mean function scores were at 3 months 62.78 and 63.8 in the SWT and NFWT respectively, and at 1 year 73.5 and 74.6, respectively. The differences between the randomized groups were not statistically significant (Table 3). There were statistically significant differences between the measurement points (Table 4).

There were improvements from admission to 12 months post-operatively in the mean 15D score, and pain and function scores, and a decrease in the cost of medication in both groups. Assuming that the final mean HRQoL score in the SWT group would carry forward until the final measurement point in the NFWT group, the latter group would gain 0.033 QALYs more than the SWT group during the whole observation period of almost 700 days (the sum of two areas between the curves in Fig. 2), even if the SWT group would gain 0.012 QALYs more in the short run by reaching a better level of HRQoL earlier (the first area

**Table 2** Outcomes at admission in total knee replacement patients, intention-to-treat analyses

Outcome	SWT Mean, ( $\pm$ SD) (n = 133–136)*	NFWT Mean, ( $\pm$ SD) (n = 194–203)*	95% Confidence interval for mean	F value	P value
15D score	0.768 (0.15)	0.779 (0.10)	–0.020 to 0.029	0.031	0.744
Pain score	18.93 (11.8)	22.79 (12.40)	–1.478 to 3.969	0.308	0.369
Function score	48.42 (22.1)	46.95 (23.31)	–6.659 to 3.493	0.737	0.540
Cost of medication €/week	5.33 (7.55)	3.57 (5.38)	3.54 to 5.10	4.85	0.029 <sup>†</sup>
Cost of medication €/waiting time (WT)	89.19 (199.19)	120.54 (194.82)	80.75 to 131.09	1.50	0.222
WT (days)	94.60 (81.3)	239.2 (135.1)	154.5 to 188.09	100.60	0.000 <sup>‡</sup>

\*Number of observations varies because of missing values.

<sup>†</sup> $P < 0.05$ , <sup>‡</sup> $P < 0.001$ .

**Table 3** Outcomes after total knee replacement, intention-to-treat analyses

Outcome	Three months postoperatively		95% Confidence interval (CI) for mean	P value	One year postoperatively		95% CI for mean	P value
	SWT	NFWT			SWT	NFWT		
15D score, mean	0.813	0.837	-0.002 to 0.05	0.06	0.813	0.852	0.01 to 0.07	0.012*
±SD	0.12	0.11			0.14	0.1		
Pain score, mean	32.7	34.07	-1.55 to 4.85	0.311	36.27	36.95	-3.11 to 4.47	0.724
±SD	13.03	13.49			13.15	12.83		
Function score, mean	62.78	63.86	-7.12 to 4.96	0.725	73.5	74.63	-4.71 to 6.98	0.703
±SD	25.58	25.22			23.32	22.28		
Costs of medication, mean (€/week)	3.14	3.1	-1.48 to 1.20	0.838	1.74	2.96	-0.36 to 2.25	0.142
±SD	5.88	7.29			3.96	4.07		

\*P &lt; 0.05.

**Table 4** Waiting time effect between the randomized groups

Outcome	Source	F value	P value
15D	Time	38.746	0.000 <sup>†</sup>
	Time* 15D	1.66	0.177*
Pain score	Time	159.686	0.000 <sup>†</sup>
	Time* pain score	0.645	0.563*
Function score	Time	118.47	0.000 <sup>†</sup>
	Time* function score	1.102	0.346*
Costs of medication	Time	13.172	0.000 <sup>†</sup>
	Time* costs of medication	1.17	0.317*

\*Difference between randomized groups.

<sup>†</sup>Difference between the measurement points.

General linear model, repeated measures, test of within-subjects effect.

between the curves in Fig. 2), although those in the NFWT had better quality of life postoperatively.

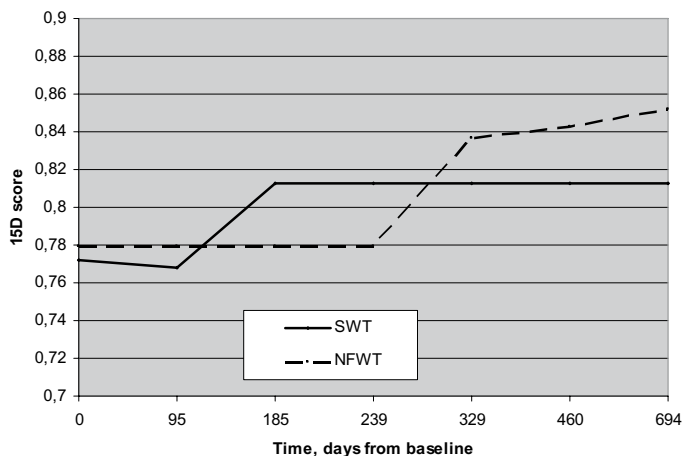
### Interpretation

Scientific evidence on the relationship between WT and TKR outcomes is inconsistent, while the absence of randomized trials has prevented an assessment of whether longer waiting is somehow related to health and quality-of-life outcomes. To our knowledge, the present study is the first to assess the use and costs of DSM in TKR patients randomly allocated to SWT and

NFWT groups. The study also analyzed whether the length of the WT was related to HRQoL, knee pain, and function as measured by the 15D and modified KS, respectively.

The main finding was that overall, there was no difference in HRQoL between the WT groups in different measurement points. Nevertheless, those in the NFWT group reached a higher level of HRQoL postoperatively and consequently gained 0.033 QALYs more than the SWT group during the whole observation period of almost 700 days. Those in the SWT group had a worse pain score at baseline, which reflected in an increased use of DSM during the waiting period. The weekly costs were almost identical in both groups at each of the three measurement points. The same applies to the mean pain and function scores, which reached their highest values 12 months postoperatively.

In this study, the patients' HRQoL at baseline was deteriorated. In fact, it seems to be worse than in two earlier studies, where the mean scores before operation measured by the 15D were 0.83 [23] and 0.81 [24] compared to our 0.77 to 0.78. The patients had pain and difficulties in functioning. Nevertheless, in an earlier study among these patients, it was found that they used very little health and social services during the WT with no statistically significant difference between the randomized groups [25]. The length of WT alone did not affect HRQoL or pain and function scores, which is in line with the findings of some earlier studies [7,8,19]. TKR has been found to be effective [24,26,27], and this study also showed a significant improvement in HRQoL,

**Figure 2** The quality-adjusted life-year gain (areas between the curves).

and in pain and function 3 and 12 months postoperatively in both randomized groups.

To our knowledge, no studies have so far been published on the effect of WT on the use and costs of DSM after the TKR. Earlier prospective studies on the costs and outcomes of the WT for TKR have found that waiting more than 6 months was associated with higher total costs and deterioration in physical function while waiting [28,29]. These cost results are not comparable to ours, as they estimated all costs during the waiting period. Nunez et al. (2007) findings on HRQoL and weekly medication costs in patients with OA on a waiting list for TKR were in line with ours.

### Strengths and Limitations

The strengths of this study were that the patients awaiting THR were prospectively followed from the time of being placed on the waiting list to admission—with WTs recorded precisely—and further for a year of follow-up postoperatively, providing evidence of the effect of WT on pre- and postoperative health status. Further, the patients were randomly assigned to either the SWT or NFWT group, the randomization was successfully completed, and the groups did not differ from each other at baseline. The findings were based on the simultaneous use of patient-reported generic and disease-specific outcome instruments. The results, based on ITT analyses, indicated that there were no statistically significant differences in health outcomes between the randomized groups during the WT. This was also tested with repeated measures analyses. The within-subject test indicated that there was a significant time effect (i.e., the outcomes did change over time), but an insignificant interaction effect (i.e., there was no difference between the groups).

Some limitations pertain to this study. First, the patients who refused to participate in the study were older, and second, a total of 29 patients in the SWT group waited more than 3 months. The reasons were hospitals' limited capacity to carry out TKR within the 3 months WT period or the patient's unwillingness to have TKR within 3 months, or they were too tired to complete the questionnaires. Because of these factors, the differences between the randomized groups may have been underestimated. Nevertheless, the primary analysis was based on the ITT principle to address the question of clinical effectiveness and to avoid the bias associated with a nonrandom loss of participants.

Third, although SF-36 has been widely used in OA area, we chose to use the 15D. SF-36 is basically a profile instrument and does not thus belong to the category of instruments producing a valuation-based single index number. Nevertheless, in 1998, the first algorithm was published to generate a valuation-based single index score, SF-6D score, from the SF-36. Hawthorne et al. [16] showed that the 15D performed at least as well as the SF-6D. Since a new algorithm has been devised for the SF-6D in 2002 [30]. Recent research with that algorithm indicates that especially in the rehabilitation of musculoskeletal disorder, the 15D was at least as responsive as the SF-6D [18].

Further, we looked at medication cost only, and medication information was obtained only from patient self-reports, while there was no distinction between self-care and prescription medication. For temporary medication, we used mean dosages. A study is ongoing, where the effect of WTs on the total cost of hip and knee replacements is being investigated.

### Conclusion

In Finland, the so-called treatment guarantee was introduced in 2005 with a maximum WT of less than 6 months [31]. Since

then, there has been discussion in health-care policy about the optimal and effective WT in elective surgery. When considering the optimal timing, information is needed on the effect of WT on key parameters. This study showed that the length of the WT, at least as realized in practice in this study, did not result in different health outcomes in the randomized groups in three cross-sectional follow-up measurements from baseline to 1 year postoperatively. Those in the SWT group reached better HRQoL 3 months earlier than those in the NFWT group, but the latter had better HRQoL after the operation and gained more QALYs overall during the study period. To be useful for future policy work, further research is needed to determine the optimal timing.

Source of financial support: This study was financially supported by the Academy of Finland (no. 51871), HUS HUCH Jorvi Hospital, Coxa Hospital for Joint Replacement, Medical Research of Tampere University Hospital, HUS HUCH Surgical Hospital, Orton Orthopaedic Hospital, and Yrjö Jahansson Foundation.

### References

- Kontio N. Current care summary 2007. Available from: <http://www.kaypahoito.fi> [Accessed August 20, 2007].
- National Research and Development Centre for Welfare and Health. 2006. Operative inpatient service. Available from: <http://www.stakes.info/0,1,7.asp> [Accessed November 29, 2006].
- Martikainen J. Finnis Statistics on Medicin 2005. National health insurance institution, statistics. 2005. Available from: <http://www.kela.fi/statistics>.
- Hoogeboom TJ, van den Ende CHM, van der Sluis G, et al. The impact of waiting for total joint replacement on pain and functional status: a systematic review. *Osteoarthritis Cartilage* 2009;17:1420–7.
- Hirvonen J, Blom M, Tuominen U, et al. Health-related quality of life in patients waiting for major joint replacement. A comparison between patients and population controls. *Health Qual Life Outcomes* 2006;4:3.
- Kapstad H, Rustoen T, Hanestad BR, et al. Changes in pain, stiffness and physical function in patients with osteoarthritis waiting for hip or knee joint replacement surgery. *Osteoarthritis Cartilage* 2007;15:837–43.
- Hawker GA, Stewart L, French MR, et al. Understanding the pain experience in hip and knee osteoarthritis—an OARSI/OMERACT initiative. *Osteoarthritis Cartilage* 2008;16:415–22.
- Boutron I, Rannou F, Jardinard-Lopez M, et al. Disability and quality of life of patients with knee or hip osteoarthritis in the primary care setting and factors associated with general practitioners' indication for prosthetic replacement within 1 year. *Osteoarthritis Cartilage* 2008;16:1024–31.
- Kelly KD, Voaklander DC, Johnston DW, et al. Change in pain and function while waiting for major joint arthroplasty. *J Arthroplasty* 2001;16:351–9.
- Hirvonen J, Blom M, Tuominen U, et al. Evaluating waiting time effect on health outcomes at admission: a prospective randomized study on patients with osteoarthritis of the knee joint. *J Eval Clin Pract* 2007;13:728–33.
- Salaffi F, Carotti M, Grassi W. Health-related quality of life in patients with hip or knee osteoarthritis: comparison of generic and disease-specific instruments. *Clin Rheumatol* 2005;24:29–37.
- Salaffi F, Carotti M, Stancati A, Grassi W. Health-related quality of life in older adults with symptomatic hip and knee osteoarthritis: a comparison with matched healthy controls. *Aging Clin Exp Res* 2005;17:255–63.
- Siciliani L, Hurst J. Explaining waiting time variations for elective surgery across OECD countries. *OECD Health Working Papers* 2003;7:10–1.
- Sintonen H. Outcome measurement in acid-related diseases. 1994;5:17–24.
- Sintonen H. The 15D instrument of health-related quality of life: properties and applications. *Ann Med* 2001;33:328–36.



- 16 Hawthorne G, Richardson J, Day NA. A comparison of the assessment of quality of life (AQoL) with four other generic utility instruments. *Ann Med* 2001;33:358–70.
- 17 Stavem K. Reliability, validity and responsiveness of two multi-attribute utility measures in patients with chronic obstructive pulmonary disease. *Qual Life Res* 1999;8:45–54.
- 18 Moock J, Kohlmann T. Comparing preference-based quality-of-life measures: results from rehabilitation patients with musculoskeletal, cardiovascular, or psychosomatic disorders. *Qual Life Res* 2008;17:485–95.
- 19 Insall JN, Dorr LD, Scott RD, et al. Rationale of the knee society clinical rating system. *Clin Orthop Relat Res* 1989;248:13–4.
- 20 Lingard E, Katz J, Wright R, et al. Validity and responsiveness of the Knee Society Clinical Rating System in comparison with SF-36 and WOMAC. *J Bone Joint Surg Am* 2001;83:1856–64.
- 21 Pharmaceutical Information Centre in Finland. *Pharmaca fennica*. 2004;CD-rom 1 Finland.
- 22 Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ* 1999;319:670–74.
- 23 Rissanen P, Aro S, Slätis P, et al. Health and quality of life before and after hip or knee arthroplasty. *J Arthroplasty* 1995;10:169–75.
- 24 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;78:108–15.
- 25 Hirvonen J, Blom M, Tuominen U, et al. Is longer waiting time associated with health and social services utilization before treatment? A randomized study. *J Health Serv Res Policy* 2007b; 12:209–14.
- 26 Lavernia CJ, Guzman JF, Gachupin-Garcia A. Cost effectiveness and quality of life in knee arthroplasty. *Clin Orthop Relat Res* 1997;345:134–9.
- 27 Rissanen P, Aro S, Sintonen H, et al. Costs and cost-effectiveness in hip and knee replacements. A prospective study. *Int J Technol Assess Health Care* 1997;13:575–88.
- 28 March L, Cross M, Tribe K, et al. Cost of joint replacement surgery for osteoarthritis: the patients' perspective. *J Rheumatol* 2002;29:1006–14.
- 29 Nunez M, Nunez E, Segur J, et al. Health-related quality of life and costs in patients with osteoarthritis on waiting list for total knee replacement. *Osteoarthritis Cartilage* 2007;15:258–65.
- 30 Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J. Health Econ* 2002;21:271–92.
- 31 Ministry of Social Affairs and Health. 2004. Available from: [http://www.stm.fi/en/social\\_and\\_health\\_services/legislation](http://www.stm.fi/en/social_and_health_services/legislation) [Accessed December 31, 2004].



**Tuominen U, Sintonen H, Hirvonen J, Hietaniemi K, Lehto MUK, Seitsalo S, Blom M. The cost-utility of waiting time in total joint replacements. RCT. 2012 (accepted for publication September 2012). International Journal of Technology Assessment in Health Care 2013; 29 (1): 27–34.**



# COST-UTILITY OF WAITING TIME IN TOTAL JOINT REPLACEMENTS: A RANDOMIZED CLINICAL TRIAL

Ulla Tuominen

*The Social Insurance Institution, Research Department*

Harri Sintonen, Pasi Aronen

*Hjelt Institute, Department of Public Health*

Johanna Hirvonen, Seppo Seitsalo, Matti Lehto, Pekka Paavolainen, Kalevi Hietaniemi, Maria Blom

*University of Helsinki, Faculty of Pharmacy*

**Keywords:** Cost-utility, Total Joint Replacement, HRQoL, 15D, waiting time

In many Western countries, long waiting times for elective surgery are a concern. Major joint replacement is an example of a type of surgery with a high volume of demand and relatively long waiting periods for patients. As populations get older, the prevalence of slowly progressive diseases, such as osteoarthritis (OA) in hip and knee joints, is increasing. Over three-quarters of a million total hip and knee replacement surgeries are done in the United States annually (1). Furthermore, according to March et al. (1997), the costs of OA have been estimated to account for up to 1–2.5 percent of the Gross National Product (GNP) in several developed countries (2). In Finland, a total of 11,104 total joint replacements (TJRs) were performed in 2004 (hip 6,600 and knee 5,905), with the median waiting time of 181 days for the surgery (hip 153 and knee 209 days). Until 2007, the number of TJRs was 17,334 (hip 7,698 and knee 9,636), with a median waiting time of 120 and 142 days, respectively (3;4). The mean waiting time for elective surgical procedures is approximately 3 months in several countries and the maximum waiting times can stretch into years.

An important question is what effect do longer waiting times, brought about by lower rates of surgery, have on patient welfare. Health status is likely to deteriorate (on average) with waiting and welfare will be lower if there is postponement of the benefit from surgery (time preference). However, the OECD Waiting Times study found surprisingly little evidence, from a review of the medical literature, of significant deterioration of health or worsening of surgical outcomes as a result of waiting for elective surgery in those countries where waiting times are up to 3 or 6 months, depending on the condition (5). Surgeons seem to be good at triage, that is, at re-prioritizing patients whose con-

ditions become unstable or deteriorate while they wait. Longer waiting may be more problematic. A study of patients on the waiting list for total hip replacement at one hospital in the United Kingdom, using a health status score specific to hip pathology, found evidence of significant deterioration and that the deterioration was greater the longer the wait. The median wait, here, was approximately 1 year (6). Similarly, a study of patients waiting for varicose vein surgery in the United Kingdom found “considerable deterioration” in their condition while waiting for surgery. In this case, the median wait was 20 months (7).

It is commonly thought that, due to waiting, there is a deterioration in the condition for which treatment is required, a loss of health-related quality of life (HRQoL) in the form of significant pain or disability, as well as an increase in the costs of surgery and use of other treatments and healthcare services pre- and postoperatively. However, as Siciliani et al. (8) suggest, eliminating waiting times altogether would not in fact constitute an optimum waiting time from the perspective of the hospital. It can be cost-effective to maintain short queues of elective patients, because the adverse health consequences of short delays are small and because there are savings in hospital capacity from allowing queues to form (8;9).

According to earlier studies, patients waiting for TJR have a poor quality of life and they have difficulties functioning in their daily activities (10–15). However, little is known about the cost-utility of total joint replacement in relation to waiting time.

The aim of our prospective, randomized, controlled trial was to compare CU of short and longer waiting times for TJR. Many observational studies have documented findings before and after operations, but such studies do not control for the natural course of the disease (4;16–18). According to literature reviews by Hoogbeem et al. (1) and ourselves, no prior studies have estimated the effect of waiting time on the cost-utility of total joint replacement in hip and knee patients using a randomised study design.

This study was financially supported by the Academy of Finland (no. 51871), HUS HUCH Jorvi Hospital, Coxa Hospital for Joint Replacement, the Medical Research Fund of Tampere University Hospital, HUS HUCH Surgical Hospital, Orton Orthopaedic Hospital and the Yrjö Jahnsson Foundation.

## MATERIALS AND METHODS

### Data Collection and Study Design

Between August 2002 and November 2003, we recruited a total of 833 patients from three Finnish hospitals to take part in this study. The three hospitals were the Surgical Hospital and Jorvi Hospital, which are both part of Helsinki University Central Hospital, and Coxa Hospital for Joint Replacement in Tampere. Patients were recruited to take part in the study through contact with orthopedic and practice staff during four (at Coxa three) recruitment periods (Supplementary Table 1, which can be viewed online at [www.journals.cambridge.org/thc2013067](http://www.journals.cambridge.org/thc2013067)). Patients' recruitment is shown in the flow chart (Figure 1) following the requirements of the CONSORT statement.

The key inclusion criteria were the need for a primary TJR due to OA of the hip or knee joint, as evaluated by the hospital surgeon, that the patient was an adult (age >16) and placed on the waiting list in a research hospital, and that the patient was willing and mentally able to participate in the study. The key exclusion criteria were patients with rheumatoid arthritis, fractures, and congenital hemophilia or congenital deformities.

### Randomization

Once the patients had been placed on the hospital waiting list, the study nurse randomly assigned them to one of two groups: (i) a short waiting time (SWT) group, with a maximum waiting period of 3 months; or (ii) a nonfixed waiting time (NfWT) group, with surgery performed according to the hospital's routine procedure and with the waiting period measured from the date the patient was added to the waiting list to the date of admission for surgery. The number of patients placed on the waiting list varied from one month to another, being specific to each hospital. Therefore, we could not estimate in advance the number of patients to be placed on the list. The patients randomized into the SWT group could only be operated on in one of four operating periods (within 2 weeks after each recruitment period) during the year. The arrangement was needed because operating rooms for the surgery of SWT patients had to be booked in advance before we could recruit the patients. For ethical reasons, all patients waiting for total joint replacement had to have an equal chance of being recruited to participate in the study in either the SWT or the NfWT group. As only half of the hospitals' 1-month surgical capacity could be allocated to the SWT group, the number of SWT patients was restricted and determined specifically for each hospital. Therefore, we needed to allocate the patients in unequal numbers to either the SWT or the NfWT group.

The researchers generated the random allocation sequence using a computer with a random number generator programmed with Visual Basic. In each hospital, after being placed on the waiting list, we informed the patient about the study and the patient provided his or her informed signed consent. The study nurse assigned participants to their groups after the decision for

surgery had been made and informed the patient of the decision. A separate randomization procedure was performed within each hospital. Surgeons were blind to patient allocation. For ethical reasons, double-blinding was not possible.

The patients used a self-administered questionnaire to report their socio-demographic data, comorbidities as diagnosed by a medical doctor, HRQoL, disease-specific medication (DSM), ability to function and the degree of pain, and the use of health and social services.

The study was approved by the Helsinki University Central Hospital Surgery Ethics Committee (registration np. 134/E6/02).

### Measurement of HRQoL

We measured HRQoL using the generic 15D instrument. The 15D is composed of fifteen dimensions: moving, vision, hearing, breathing, sleeping, eating, speech, eliminating, vitality, usual activities, mental function, discomfort and symptoms, depression, distress, and sexual activity. Each dimension has five ordinal levels to choose from. The 15D can be used as a profile measure or to give a single index score by means of population-based preference weights. The index score (15D score) ranges from 0 (dead) to 1 (completely healthy) (19). The 15D questionnaire takes 5–10 minutes to complete and it describes the HRQoL of the respondent at present. A difference of  $>|0.03|$  in the 15D score is clinically important in the sense that, on average, people can feel the difference (20). We chose to use the 15D for three main reasons: (i) it has been used successfully in earlier studies dealing with hip and knee replacement and thus facilitates a comparison of the presurgery scores in these studies; (ii) earlier research has shown that in most of the important properties (reliability, content validity, sensitivity in terms of discriminatory power and responsiveness to change), the 15D instrument compares at least equally with other similar instruments that produce a valuation-based single index number (21;22); and (iii) recent research has since confirmed that, especially in the rehabilitation of musculoskeletal disorders, the 15D instrument was at least as responsive as the SF-6D and much more responsive than the EQ-5D (23).

By using the mean 15D scores from each measurement point, and by assuming a linear change in the scores between the measurement points, we also estimated the possible gain in quality-adjusted life-years (QALY gain) for both groups within the observation period. To obtain an equally long observation period for both randomized groups, we assumed that the final HRQoL scores in the SWT group would carry forward until the mean final measurement point in the NfWT group and that members of the SWT group would incur no further costs during that time.

### Cost Data

The data on the usage of healthcare and social services were based on patients' self-reports from the waiting time to 1 year

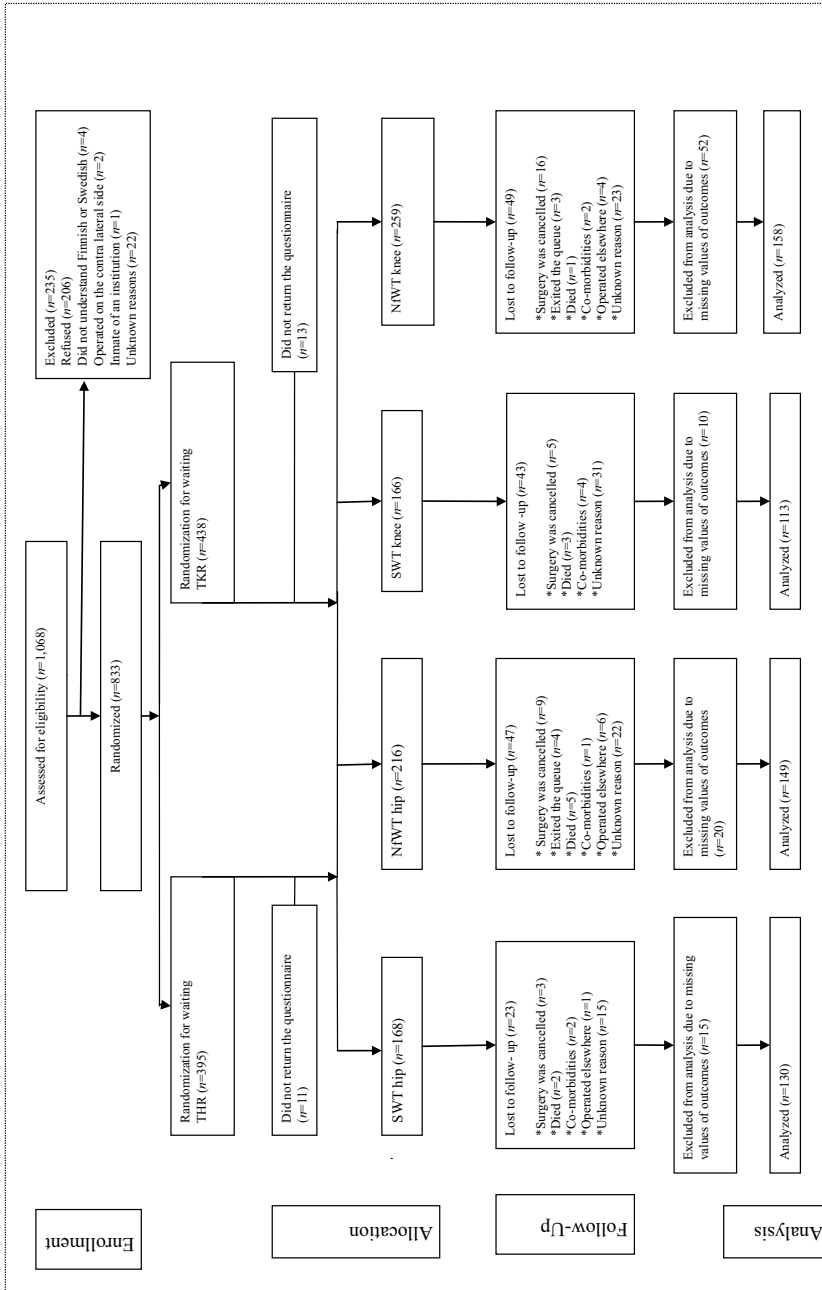


Figure 1. Flow of patients through the trial.

postoperatively, which we measured in 3-month periods. We obtained the costs for the surgery from the Finnish Hospital Discharge Register. We valued the use of healthcare and social services at Finnish unit costs for the year 2006 (24). The total direct costs include the following items: outpatient visits (doctor, nurse and chiropodist), the costs of the surgery including radiology, laboratory services, hospital days, and rehabilitation services. We multiplied the use and costs of regular social services due to OA, including meals-on-wheels, home help, laundry services, bathing services, and transportation, during the waiting time by the number of months spent on the waiting list. We carried out all analyses from a Finnish societal perspective, excluding production losses and value-added taxes.

We used the total costs thus calculated and the QALYs gained during the observation period to compare the cost-utility of SWT and NfWT separately for hip and knee patients. As even the longest observation period was shorter than 2 years, no discounting was applied.

#### Statistical Analysis

The sample size estimate was based on the primary outcome variable: the 15D score. A subgroup of 177 patients would provide an 80 percent power (two-tailed  $\alpha$  error 5 percent) of detecting clinically important differences  $\Delta 0.03$  in the mean 15D score between the randomized groups. We conducted primary analyses using the intention-to-treat (ITT) principle (24), so that we could follow the patients in the groups to which they had been randomly allocated. As a secondary analysis we looked at patients in the different randomization groups with actually shorter and longer waiting times (per protocol analysis).

We compared the characteristics of the randomized groups and those who were lost to follow-up at baseline using either the independent samples *t*-test or the chi-squared test, depending on whether the variable was on a continuous or a nominal scale. In addition, we calculated the mean values for use and the costs of health and social services. To assess the degree of uncertainty in the results, we performed a probabilistic sensitivity analysis (bootstrapping with 1,000 replicates). The results are given in the form of mean incremental costs and effects with their 95 percent confidence intervals, an incremental cost-effectiveness ratio (ICER), a cost-effectiveness plane and a cost-effectiveness acceptability curve (CEAC).

We replaced the missing values on the 15D dimensions, if a minimum of 80 percent of the dimensions had been completed, using a regression model with the patient's responses for other dimensions, age and gender as explanatory variables (18). Data analyses were performed using SPSS versions 14 and 16 for Windows.

## RESULTS

### Patient Characteristics

Of the eligible patients invited to participate in the study, 235 (160 women) patients with a mean age of 70 years refused to participate and were excluded. The most frequently quoted reason for refusal was an unwillingness to complete the questionnaires. Thus, 833 patients, after providing informed consent, were randomly allocated to either the SWT ( $n = 346$ ) or NfWT ( $n = 487$ ) group (Figure 1). Of the 833 randomized patients, 24 did not return the questionnaire at baseline, although they had signed informed consent forms and had been randomized. Of the remaining 809 patients, 162 were lost to follow-up during the waiting time for various reasons and were not included in the final analyses (Figure 1). Due to missing values, the final cost-utility analyses are based on 550 (66 percent) of the randomized patients who completed the questionnaires, with a mean ( $\pm$ SD) age of 66 ( $\pm 9.9$ ; range; from 33 to 89) years, of which 345 (63 percent) were women and a further 243 (hip  $n = 130$ , knee  $n = 113$ ) were in the SWT group and 307 (hip  $n = 149$ , knee  $n = 158$ ) were in the NfWT group (Figure 1).

The baseline characteristics of the randomized groups were similar (Table 1). We have reported the details about the characteristics of these two patient groups in our earlier studies (24;25). The mean ( $\pm$ SD) 15D score in the SWT group for hip patients was 0.770 ( $\pm 0.09$ ) and 0.779 ( $\pm 0.10$ ) in the NfWT group; the difference was neither statistically significant nor clinically important (95 percent confidence interval [CI] for a mean difference from  $-0.036$  to  $0.026$ ). The mean ( $\pm$ SD) 15D scores at baseline for knee patients were 0.772 ( $\pm 0.18$ ) and 0.779 ( $\pm 0.12$ ), respectively (95 percent CI for a mean difference from  $-0.004$  to  $0.030$ ) (Table 1). The percentage of patients receiving disease specific medication (DSM) was more than 87 percent in all patients groups.

Approximately 20 percent ( $n = 162$ ) of the patients dropped out after randomization at any stage of the follow-up. The only statistically significant differences in the baseline characteristics between the dropouts and those, who remained in the study to the end of follow-up, were found in the mean age (dropouts slightly older) and in the proportion of living alone (among dropouts slightly higher) (Table 1).

### Cost-Utility

The mean waiting time for hip patients was 74 (SD  $\pm 145$ ;  $n = 145$ ) days in the SWT group and 194 (SD  $\pm 175$ ;  $n = 169$ ) days in the NfWT group, and for knee patients 94 (SD  $\pm 81$ ;  $n = 123$ ) days and 239 (SD  $\pm 135$ ;  $n = 210$ ) days, respectively.

The 15D score improved after the operation in all four groups (Table 2). The mean ( $\pm$ SD) total costs for healthcare and social services are reported in Table 3.

The mean total costs of TJR among hip replacement patients were EUR 9986 ( $\pm 3,540$ ) in the SWT group and EUR 10 472 ( $\pm 4,686$ ) in the NfWT group, and EUR 9809 ( $\pm 4,085$ ) and



**Table 1.** Baseline Characteristics in Patients Waiting for Total Joint Replacement and Those Lost to Follow-up

Characteristics at baseline	SWT n = 268	NFWT n = 379	Lost n = 162	p-value‡	p-value†
Age, years (mean ± SD)	66 (9.5)	66 (9.7)	68 (9.9)	0.687	0.023*
Females [n, (%)]	185 (65.1)	231 (60.5)	89 (62.2)	0.152	0.538
Housing [n, (%)]					
Living alone	95 (33.9)	104 (28.1)	59 (41.3)	0.066	0.010*
Basic education [n, (%)]					
Lower level	227 (81.1)	309 (83.3)	119 (83.8)	0.264	0.390
Comorbidity, yes [n, (%)]	207 (73.9)	270 (72.9)	108 (75.5)	0.406	0.330
Medication to arthritis, yes [n, (%)]	248 (88.6)	329 (88.7)	124 (86.7)	0.531	0.302
Medication to comorbidities yes [n, (%)]	215 (76.8)	273 (73.6)	112 (78.3)	0.200	0.232
Hip	145 (51.1)	169 (47.1)	70 (41.3)	0.126	0.520
BMI <sup>a</sup> (mean ± SD)	28.77 (5.68)	28.67 (5.21)	27.66 (7.07)	0.809	0.099
15D-score <sup>b</sup> (mean ± SD)	0.765 (0.105)	0.774 (0.102)	0.761 (0.120)	0.222	0.360

\*p &lt; 0.05.

‡Difference between the randomized groups.

†Difference between the patients, who remained in the study to the end of follow-up and those lost to follow-up.

<sup>a</sup>Body mass index (kg/m<sup>2</sup>).<sup>b</sup>15D-score (scale 0 = worst, 1 = best).**Table 2.** Mean 15D Scores in the Randomized Groups of Hip and Knee Patients

Measurement point	Hip SWT	NFWT	95% CI for mean difference	Knee SWT	NFWT	95% CI for mean difference
At baseline (±SD)	0.770 (0.09)	0.779 (0.10)	−0.036 to 0.026	0.772 (0.18)	0.779 (0.12)	−0.004 to 0.030
At admission (±SD)	0.772 (0.09)	0.771 (0.10)	−0.023 to 0.021	0.778 (0.11)	0.786 (0.13)	−0.008 to 0.014
Three months pop (±SD)	0.945 (1.06)	0.854 (0.11)	−0.277 to 0.094	0.811 (0.12)	0.842 (0.11)	0.033 to 0.058*
One year pop (±SD)	0.854 (0.12)	0.904 (0.50)	−0.033 to 0.133	0.823 (0.14)	0.852 (0.10)	0.019 to 0.035*

\*p &lt; 0.01.

**Table 3.** Mean Use and Costs of Healthcare and Social Services during the Waiting Time between the Randomized Groups, in Hip and Knee Patients

Items of resource use	Mean number in hip patients		Mean number in knee patients		Mean costs in hip patients (€)		Mean costs in knee patients (€)	
	SWT n = 130	NFWT n = 149	SWT n = 113	NFWT n = 158	SWT	NFWT	SWT	NFWT
Outpatients' visits	1.9	1.4	1.5	1.2	232	152	188	129
Hospital days	0.2	0.4	0.9	0.4	118	266	377	247
Healthcare service (at home)	0.3	0.1	0.1	0.1	8	4	4	1
Rehabilitation	1.0	0.2	0.5	0.3	61	16	21	14
Homecare service	0.6	0.6	1.3	1.2	21	21	38	33

<sup>a</sup>Sum of different types of hospital outpatients' units (University hospital, central hospital, district hospital, health care centre, private hospital, occupational health care unit).<sup>b</sup>Regular homecare services due to osteoarthritis.

EUR 9801 ( $\pm 3,116$ ) among knee replacement patients, respectively. During the equally long follow-up period, the SWT hip patients experienced, on average, 1.341 QALYs and the NfWT patients 1.327 QALYs. Correspondingly, the SWT knee patients experienced, on average, 1.453 QALYs and the NfWT patients 1.467 QALYs (Supplementary Table 2, which can be viewed online at [www.journals.cambridge.org/thc2013068](http://www.journals.cambridge.org/thc2013068)). Point estimates thus suggest a strong dominance for SWT among hip patients but for NfWT among knee patients (Supplementary Figures 1, 2, which can be viewed online at [www.journals.cambridge.org/thc2013069](http://www.journals.cambridge.org/thc2013069) and [www.journals.cambridge.org/thc2013070](http://www.journals.cambridge.org/thc2013070)). On the basis of probabilistic sensitivity analysis in hip patients, the 95 percent CI for the mean difference in QALYs was from  $-0.048$  to  $0.076$  and in costs from  $-1453\text{€}$  to  $464\text{€}$ . In knee patients, the 95 percent CI for the mean difference in QALYs was from  $-0.095$  to  $0.063$  and in costs from  $-913\text{€}$  to  $955\text{€}$ . If the willingness to pay for a QALY is EUR 20,000, the probability of SWT being cost-effective for hip patients is approximately 85 percent and approximately 40 percent for knee patients (Supplementary Figure 3, which can be viewed online at [www.journals.cambridge.org/thc2013071](http://www.journals.cambridge.org/thc2013071)).

In the secondary per protocol analysis the mean total costs among hip patients were EUR 10,302 ( $\pm 3788$ ) in the SWT group and EUR 10,402 ( $\pm 4854$ ) in the NfWT group, and EUR 9,374 ( $\pm 3259$ ) and EUR 9904 ( $\pm 3115$ ) among knee patients, respectively. During the equally long follow-up period, the SWT hip patients experienced, on average, 1.3536 QALYs and the NfWT patients 1.3879 QALYs. Correspondingly, the SWT knee patients experienced, on average, 1.4428 QALYs and the NfWT patients 1.5022 QALYs. Point estimates thus suggest an ICER of EUR 3000 for NfWT among hip patients and of EUR 9058 among knee patients. However, there is a lot of variance around the point estimates and the differences in costs and QALYs between the per protocol groups were not statistically significant neither in hip nor knee patients.

## DISCUSSION

Scientific evidence on the relationship between waiting time and outcomes for TJR is inconsistent. The absence of randomized trials has prevented an assessment of whether longer waiting is somehow related to HRQoL outcomes and costs. The present study compared the cost-utility of short and longer waiting times for TJR. To our knowledge, this study is the first one in which patients were randomly allocated to short and nonfixed waiting time groups when placed on the waiting list and followed according to the ITT principle.

The main finding of this study was that hip patients in the SWT group gained, on average, more QALYs at lower costs than patients in the NfWT group, suggesting a strong dominance for the SWT group. In knee patients the situation was the opposite. However, there is a high degree of uncertainty surrounding these results based on point estimates, and probabilistic sensitivity

analyses indicated that if the willingness to pay for a QALY is EUR 20 000, then the probability of a SWT being cost-effective in hip patients is approximately 85 percent and only approximately 40 percent in knee patients.

It is to be noted although that our findings may not be fully transferable to other countries. Even by using the same HRQoL instrument and valuation algorithm, the HRQoL results may not be similar due to different indications of treatment. Transferability of costs is shadowed by differences across countries e.g. in treatment practices and unit costs.

## Strengths and Limitations

There are some limitations to this study. First, a total of seventy-four patients in the SWT group waited for more than 3 months. The main reasons for this were the hospitals' limited capacity to carry out TJR within the 3-month waiting time period or the patients' unwillingness to be operated on within 3 months. Due to these factors, the differences between the randomized groups may have been underestimated and there might also be some bias in the use of health and social services. However, the primary analysis was based on the ITT principle to address the question of clinical effectiveness and to avoid the bias associated with a nonrandom loss of participants.

The per protocol analysis gave rise to further uncertainty over whether there is any real difference between the waiting time groups in cost-utility in either procedure. The point estimates suggested an ICER of EUR 3000 for NfWT among hip patients and of EUR 9058 among knee patients, but the differences in costs and QALYs between the per protocol groups were not statistically significant neither in hip nor knee patients.

Second, defining and measuring the waiting time for surgery is not a simple matter. What is the real starting point for the waiting period? According to Siciliani and Hurst (2003), one observable starting point is the time when a patient is first referred by a general practitioner to a hospital to be assessed for surgery. In the present study, the waiting time began when the practitioner first made the decision for surgery, even though patients may have already been waiting for an unknown amount of time before this decision. This might affect patients' baseline quality of life, which was poor (5).

Third, also establishing comparable QALYs and costs between the SWT and NfWT group is not without weaknesses, as the final measurements of HRQoL and costs in both groups did not take place equidistantly in time from the baseline. With our solution, the mean follow-up time is the same in both groups. However, we do not know exactly, how the HRQoL and costs in the SWT group developed during the time from the last measurement in that group to the final measurement in the NfWT group. The HRQoL may have deteriorated slightly due to ageing, but as the mean time difference between the last measurements in the groups was only 4–8 months, the change would probably be negligible; therefore, our assumption of no change may be justified. On the other hand, had the SWT group incurred further

costs contrary to our assumption, its total cost would have been underestimated. As these changes would probably been marginal, they may have not affected our conclusions.

Another possible weakness is that approximately one-third of patients dropped out during the follow-up. However, apart from being slightly older and living slightly more frequently alone, the dropouts did not deviate in a statistically significant manner in their baseline characteristics from those, who remained in the study to the end of follow-up. Thus overall, the dropout may not bias our results significantly.

Finally, the costs of medication were not included in the final analyses; the costs have been reported in our earlier studies and the findings were that the cost trends were highest during the waiting time and lowest after the operation (15;26;27).

The strengths of this study are that the patients awaiting TJR were prospectively followed from the time of first being placed on the waiting list to admission—with waiting times recorded precisely—and further for a year postoperatively, providing evidence of the effect of waiting time on pre- and postoperative health status. Furthermore, the patients were randomly assigned to the SWT and NfWT groups; the randomization was successfully completed and the groups did not differ from each other at baseline.

## CONCLUSION

According to the present study, there does not seem to be a significant difference in the cost-utility of short and longer waiting times for TJR, at least given the waiting time difference between our study groups.

## SUPPLEMENTARY MATERIAL

Supplementary Table	1:	<a href="http://www.journals.cambridge.org/thc2013067">www.journals.cambridge.org/thc2013067</a>
Supplementary Table	2:	<a href="http://www.journals.cambridge.org/thc2013068">www.journals.cambridge.org/thc2013068</a>
Supplementary Figure	1:	<a href="http://www.journals.cambridge.org/thc2013069">www.journals.cambridge.org/thc2013069</a>
Supplementary Figure	2:	<a href="http://www.journals.cambridge.org/thc2013070">www.journals.cambridge.org/thc2013070</a>
Supplementary Figure	3:	<a href="http://www.journals.cambridge.org/thc2013071">www.journals.cambridge.org/thc2013071</a>

## CONTACT INFORMATION

**Ulla Tuominen, MD, M.Sc. PhD Candidate**, ([ulla.tuominen@kela.fi](mailto:ulla.tuominen@kela.fi)), Social Insurance Institution of Finland, Helsinki, Finland

**Harri Sintonen, MD, PhD, Professor**, Hjelt Institute/Department of Public Health, University of Helsinki, Finland

**Pasi Aronen, M.Sc.**, Hospital District of Helsinki and Uusimaa, Helsinki, Finland

**Johanna Hirvonen, MD, PhD**, Mikkeli University of Applied Sciences, Mikkeli, Finland

**Seppo Seitsalo, MD, PhD, Professor**, Orton Orthopaedic Hospital, Helsinki, Finland

**Matti Lehto, MD, PhD, Professor**, University of Tampere, Finland

**Kalevi Hietaniemi, MD**, Hospital District of Helsinki and Uusimaa, Finland

**Maria Blom, MD, PhD, Professor**, Division of Social Pharmacy, Faculty of Pharmacy University of Helsinki, Finland

## CONFLICTS OF INTEREST

Harri Sintonen is the developer of the 15D and receives royalties from the electronic version of the 15D. He is a member of scientific advisory boards of MSD and Eli Lilly and has received consultancy fees or honoraria from several medical companies. The other authors report no potential conflicts of interest.

## REFERENCES

1. Hoozeboom TJ, van den Ende CHM, van der Sluis G, et al. The impact of waiting for total joint replacement on pain and functional status: a systematic review. *Osteoarthritis Cartilage*. 2009;17:1420-7.
2. March LM, Bachmeier CJM. Economic of osteoarthritis: a global perspective. *Baillieres Clin Rheumatol*. 1997;11:817-34.
3. National Research and Development Centre for Welfare and Health. Operative inpatient service 2006. [cited 2007 November 16]. Available from: <http://www.stakes.info/0,1,7.asp>.
4. Lavernia CJ, Guzman JF, Gachupin-Garcia. A cost effectiveness and quality of life in knee arthroplasty. *Clin Orthop*. 1997;345:134-9.
5. Siciliani L, Hurst J. Explaining waiting time variations for elective surgery across OECD countries. *OECD Health Working Papers*. 2003;7:10-1.
6. Kili S, Wright I, Jones RS. Change in Harris hip score in patients on the waiting list for total hip replacement. *Ann R Coll Surg Engl*. 2003;85:269-71.
7. Sarin S, Shields DA, Farrah J, Scurr JH, Coleridge-Smith PD. Does venous function deteriorate in patients waiting for varicose vein surgery? *J R Soc Med*. 1993;86.
8. Siciliani L, Hurst J. Explaining waiting time variations for elective surgery across OECD countries. *OECD Economic Studies No. 38*. 2004/1:95-123
9. Gravelle H, Siciliani L. Is waiting-time prioritisation welfare improving? *Health Econ*. 2008;17:167-84.
10. Hirvonen J, Blom M, Tuominen U, et al. Health-related quality of life in patients waiting for major joint replacement. A comparison between patients and population controls. *Health Qual Life Outcomes*. 2006;4:3.
11. Hirvonen J, Blom M, Tuominen U, et al. Evaluating waiting time effect on health outcomes at admission: a prospective randomized study on patients with osteoarthritis of the knee joint. *J Eval Clin Practice*. 2007;13:728-33.
12. Kapstad H, Rustoen T, Hanestad BR, et al. Changes in pain, stiffness and physical function in patients with osteoarthritis waiting for hip or knee joint replacement surgery. *Osteoarthritis Cartilage*. 2007;15:837-43
13. Hawker GA, Stewart L, French MR, et al. Understanding the pain experience in hip and knee osteoarthritis—an OARSI/OMERACT initiative. *Osteoarthritis Cartilage*. 2008;16:415-22.
14. Boutron I, Rannou F, Jardinard-Lopez M, et al. Disability and quality of life of patients with knee or hip osteoarthritis in the primary care setting and factors associated with general practitioners' indication for prosthetic replacement within 1 year. *Osteoarthritis Cartilage*. 2008;16:1024-31.
15. Tuominen U, Sintonen H, Hirvonen J, et al. The effect of waiting time on health outcomes and the use and costs of medication in total hip replacement patients: a randomised clinical trial. *Osteoarthritis Cartilage*. 2009;17:1144-50.

16. Laupacis A, Bourne R, Robareck C, et al. Cost of elective total hip arthroplasty during the first year. Cemented versus noncemented. *J Arthroplasty*. 1994;9:481-7.
17. 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;78:108-15.
18. Rissanen P, Aro S, Sintonen H, et al. Costs and cost-effectiveness in hip and knee replacements. A prospective study. *Int J Technol Assess Health Care*. 1997;13:575-88.
19. Sintonen H. Outcome measurement in acid-related diseases. *Pharmaco Economics*. 1994;5(Suppl.3):17-24.
20. Sintonen H. The 15D instrument of health-related quality of life: properties and applications. *Ann Med*. 2001;33:328-36.
21. Hawthorne G, Richardson J, Day NA. A comparison of the assessment of quality of life (AQoL) with four other generic utility instruments. *Ann Med*. 2001;33:358-70.
22. Stavem K. Reliability, validity and responsiveness of two multiattribute utility measures in patients with chronic obstructive pulmonary disease. *Qual Life Res*. 1999;8:45-54.
23. Moock J, Kohlmann T. Comparing preference-based quality-of-life measures: results from rehabilitation patients with musculoskeletal, cardiovascular, or psychosomatic disorders. *Qual Life Res*. 2008;17:485-95.
24. Hujanen T, Kapiainen S, Tuominen U. Unit costs in Finnish Health care. 2006. [cited 2008 April 9]. Available from: [http://www.terveydenhuollon\\_yksikkokustannukset\\_Suomessa\\_2006](http://www.terveydenhuollon_yksikkokustannukset_Suomessa_2006).
25. Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomized controlled trial. *BMI*. 1999;319:670-4.
26. Tuominen U, Sintonen H, Hirvonen J, et al. Is longer waiting time for total knee replacement associated with health outcomes and medication costs? Randomized clinical trial. *Value Health*. 2010;13:998-1004.

## RECENT PUBLICATIONS IN THE STUDIES IN SOCIAL SECURITY AND HEALTH SERIES

- 125 Pitkälä K, Savikko N, Pöysti M, Laakkonen M-L, Kautiainen H, Strandberg T, Tilvis R.** Muistisairaiden liikunnallisen kuntoutuksen vaikuttavuus. Satunnaistettu vertailututkimus. 2013. ISBN 978-951-669-909-0 (nid.), 978-951-669-910-6 (pdf)
- 124 Hakovirta M, Rantalaiho M.** Taloudellinen eriarvoisuus lasten arjessa. 2012. ISBN 978-951-669904-5 (nid.), ISBN 978-951-669-905-2 (pdf).
- 123 Kempainen T.** Well-being in socio-political context. European welfare regimes in comparison. 2012. ISBN 978-951-669-900-7 (print), ISBN 978-951-669-901-4 (pdf).
- 122 Tuovinen A-K.** Euroopan unionin toisessa jäsenvaltiossa syntyneiden sairaanhoitokustannusten korvaaminen. 2012. ISBN 978-951-669-897-0 (nid.), ISBN 978-951-669-898-7 (pdf).
- 121 Jussila H.** Päätöksenteon tukena vai hyllyssä pölyttymässä? Sosiaalipoliittisen tutkimustiedon käyttö eduskuntatyössä. 2012. ISBN 978-951-669-895-6 (nid.), ISBN 978-951-669-896-3 (pdf).
- 120 Ikonen A.** Primary care visits in the Finnish occupational health services and their connections to prevention and work-related factors. 2012. ISBN 978-951-669-882-6 (print), ISBN 978-951-669-883-3 (pdf).
- 119 Martikainen J.** Uusien lääkkeiden markkinoille tulo ja lääkekustannuksiin vaikuttaminen. 2012. ISBN 978-951-669-879-6 (nid.), 978-951-669-880-2 (pdf).
- 118 Varjonen S.** Äidin hoiva, jaettu vanhemmuus – ja vapaus valita. Perhevapaiden uudistamisen argumentointi 1970-luvulta 2000-luvulle. 2011. ISBN 978-951-669-868-0 (nid.), 978-951-669-869-7 (pdf).
- 117 Ylikännö M.** Sopivasti työtä ja vapaa-aikaa? Tutkimuksia ajankäytöstä eri elämäntilanteissa. 2011. ISBN 978-951-669-859-8 (nid.), 978-951-669-860-4 (pdf).
- 116 Mattila Y.** Suuria käännekohtia vai tasaista kehitystä? Tutkimus Suomen terveydenhuollon suuntaviivoista. 2011. ISBN 978-951-669-853-6 (nid.), ISBN 978-951-669-854-3 (pdf).
- 115 Sinokki M.** Social factors at work and the health of employees. 2011. ISBN 978-951-669-851-2 (nid.), ISBN 978-951-669-852-9 (pdf).