The expected life-time cost-effectiveness of non-surgical treatment compared to direct total knee arthroplasty for patients with severe osteoarthritis of the knee

A cost-utility analysis from perspective of the Norwegian health care service developed by decision analytic modelling

by
Lars Asphaug

Master thesis in Health Economics, Policy and Management

Department of Health Management and Health Economics
The Faculty of Medicine
University of Oslo
May 2016

The expected life-time cost-effectiveness of non-surgical conservative treatment compared to direct total knee arthroplasty for patients with severe osteoarthritis of the knee

Lars Asphaug

http://www.duo.uio.no/
Print: Reprosentralen, Universitetet i Oslo

Abstract

Background The aim of this thesis was to investigate whether a non-surgical conservative treatment programme could be a viable alternative to directly undergoing arthroplastic surgery for patients with severe osteoarthritis of the knee, and to evaluate the cost-effectiveness of this strategy. In Norway, a requirement for becoming a candidate for arthroplastic surgery is to have undergone and exhausted the effects of conservative treatment. Recent studies have found that Norwegian sufferers of knee osteoarthritis are not given adequate conservative treatment. This study compares the costs and QALYs of giving a non-surgical conservative treatment programme that is complacent with international guidelines relative to TKA directly, over the expected remaining life-time of the patients.

Methods The thesis is a cost-utility analysis developed by a decision-analytic model. A Markov-model was used for eight different cohorts of patients.

Results The strategy of a non-surgical treatment programme followed by a TKA in the event of a failed native knee is a less effective, cost-saving alternative to undergoing a primary total knee arthroplasty directly for all cohorts. There was considerable uncertainty around the decision, and a value of information analysis revealed that research should be allocated towards reducing the uncertainty of the utility parameters of the model to enable a more robust decision.

Conclusions At low willingness to pay per QALY thresholds, the non-surgical conservative treatment programme is cost-effective, however in moderate to higher ranges of willingness to pay, the TKA-only strategy is the cost-effective alternative.

Acknowledgements

I would like to thank my supervisor Hans Olav Melberg for his always positive support, helpful feedback along the way, and for giving me space to independently sharpen my researching skills. I also need to thank Ole Aars for initially telling me about the project. For teaching me all I know about decision analytic modelling I would like to thank Eline Aas.

A huge thank you to Signe Flottorp and Stijn Van de Velde at the Norwegian Institute of Public Health for all the help with the clinical side of things, for all the interesting discussions at our meetings, and generally being very welcoming and easy to work with. Thanks also to Albert Paus at Rikshospitalet for his valuable insights and feedback.

As this thesis concludes my masters degree, I have to thank all my brilliant fellow students during these two years. For this thesis, thanks especially to Tom Stephens who helped me work out critical issues in the survival analysis.

I would also like to thank my parents for giving me a hard-working attitude, and always being supportive and believing in my academic potential through the years when I doubted my own abilities.

Lastly, most of all thank you to Anja for all that she is, no explanation needed.

Contents

List of Tables

List of Figures

Abbreviations

AIC	Akaike's Information Criterion
BMI	Body mass index
CEAC	Cost-effectiveness acceptability curve
CUA	Cost-utility analysis
DRG	Diagnosis related groups
EVPI	Expected value of perfect information
EVPPI	Expected value of perfect information for parameters
GP	General practitioner
HRQoL	Health-related quality-of-life
ICER	Incremental cost-effectiveness ratio
NAR	Norwegian Arthroplasty Register
NIPH	Norwegian Institute of Public Health
NMB	Net monetary benefit
NOK	Norwegian kroner
OA	Osteoarthritis
OARSI	Osteoarthritis Research Society International
PSA	Probabilistic sensitivity analysis
PT	Physiotherapist
QALY	Quality-adjusted life-year
RCT	Randomised controlled trial
SKAR	Swedish Arthroplasty Register
TKA	Total knee arthroplasty
TTO	Time trade-off
UKA	Unicompartmental knee arthroplasty
WOMAC	Western Ontario and McMaster Universities Osteoarthritis Index
WTP	Willingness-to-pay

Chapter 1

Introduction

Replacing a body part with cement, metal and polyethylene is not a trivial event in a patient's medical history; some might even say its a last resort. A total knee arthroplasty (TKA), knee joint replacement, is one of the final treatment options for osteoarthritis of the knee provided for by the Norwegian health care service. The disease, estimated to affect 7.1% of the population [1], is a slow-progressing condition with the need for arthroplasty manifesting at a high age. With Norwegian adults increasingly joining the global trend of overweight and obesity [2], more weight will be put on old, troubled knees in the future. In other words, the demand for arthroplastic surgery is likely not shifting downwards any time soon. There are however alternatives to surgery, even for patients with advanced symptoms, and this analysis is an economic evaluation of an alternative strategy to total knee arthroplasty.

A requirement for becoming a candidate of arthroplasty is having gone through and exhausted the effects of non-surgical conservative treatment. Recent studies [3, 4] have found that Norwegian knee osteoarthritis patients are given non-surgical treatment of varying, lacking or poor quality – possible explanatory factors as to why the effects of conservative treatment becomes exhausted. The result is that annually, just under 6 000 individuals undergo primary total knee arthroplasty. This group has been increasing by about a thousand every five years since the mid-1990s [5].

New evidence from a Danish randomised controlled trial comparing non-surgical

conservative treatment head-to-head with direct total knee arthroplasty [6], suggests that not only can guideline-compliant conservative efforts improve the knee in terms of less pain and better functioning for patients originally candidates for TKA, but also that arthroplastic surgery can be delayed or possibly avoided altogether.

The Directorate of Health ordered a report from the Norwegian Institute of Public Health (NIPH) on whether non-surgical conservative treatment can delay or avoid the need of arthroplastic surgery, and whether such efforts can be considered cost-effective. The purpose of this present study is to answer that on behalf of the NIPH. The research question was defined as,
"is a non-surgical conservative treatment programme compared to directly undergoing total knee arthroplasty a cost-effective alternative over the remaining life-time of the patients?".

The analysis makes use of the above mentioned new evidence and incorporates it into a Norwegian setting, comparing to total knee arthroplasty in the framework of a cost-utility analysis developed with the perspective of the Norwegian health care service. The method used to evaluate the relative cost-effectiveness is a decision analytic model.

The study is developed to be informative for patients who are candidates for arthroplastic surgery at different age groups and gender. The outcomes of interest are defined by the nature of the framework: incremental costs and the incremental quality-adjusted life-years gained from the non-surgical conservative treatment programme compared to undergoing total knee arthroplasty directly.

The thesis is structured as ten chapters with four technical appendices supplementing the analysis. The second chapter elaborates on the problem and available recommended treatment. Following that, an explanation of the framework and method employed is given. The model itself is explained in chapter four, and the input data is given in chapter five. The sixth chapter shows the estimated results from the model, and the uncertainty around them. An analysis of the value of uncertainty-reducing information is given in chapter seven. Chapter eight showcases the changes to model results under other structural assumptions. The two last chapters is a discussion of the whole analysis and the conclusion.

Chapter 2

Background

2.1 Osteoarthritis of the knee

Osteoarthritis (OA) is a disorder affecting the synovial joints of the hands, neck, lower back, big toe, hip, and knee [7]. Synovial joints connect two bones and allow movement of the bones against one another, in the knee – the femur and the tibia. The bone-ends are covered with articular cartilage, which together with the synovial membranes make up the joint cavity [7]. Idiopathic osteoarthritis is characterised by a localised loss of the articular cartilage. Traditionally the cartilage loss was thought to be exclusively degenerative, i.e. due to biomechanical wear and tear. The contemporary perception is that the onset of cartilage loss can be biochemically caused as well; a metabolic imbalance in the cartilage cells themselves can make the cells slightly favour a catabolic metabolism [7]. With the localised loss, the focal stress of the patient's body weight on the site increases relative to the rest of the cartilage. As this tends accentuate the loss, a repair mechanism begins. At the loss-site, the bone is remodelled, and at the joint margins osteophytes (overgrowth of new bone) forms. The remodelling and overgrowth of new bone can cause the knee to become misaligned, increasing the focal stress and subsequent cartilage loss even further [8]. This usually gives a patient with knee osteoarthritis a feeling of pain in the affected knee with active motion, and functional impairment at the more severe stages [9].

Osteoarthritis of the knee was in 2008 estimated to be prevalent in approxi-

Table 2.1: Estimated prevalence of knee osteoarthritis in Norway

Age	**Females**	**Males**
74-76	28.2%	14.9%
64-66	16.1%	14.0%
54-56	10.5%	8.3%
44-46	4.9%	3.6%
34-36	1.8%	1.7%
24-26	0.6%	0.0%
All ages	7.9%	6.2%

mately 8% of women, and 6% of men in a representative sample of the Norwegian population [1]. The estimate was based on the self-reporting of 3 266 inhabitants in Ullensaker municipality. The disease was more prevalent in older individuals, and more so for women than men [1]. Generalising this to the Norwegian 2016 population of 5.2 million means that as many as 364,000 Norwegians might be burdened by the disease. As visible in table 2.1, adapted from Grotle et al. [1], the disease is more prevalent in females than males for all ages.

A 2012 study that followed 561 middle aged British women over 15 years estimated the annual incidence rate (new cases per year) of radiographically confirmed osteoarthritis of the knee at 2.3%. They also estimated that the annual rate of progression, calculated from having at least a grade 2 radiographic knee at baseline (a definite confirmation of the presence of OA), and at least one grade higher 15 years later, was 2.8% [10]. A similar study from Sweden estimated the incidence rate to 4% annually, and rate of progression of 8% per year [11]. This gives some indication that knee osteoarthritis is a disease that develops and progresses slowly, meaning the more severe stages – the focus of this analysis – are unlikely to manifest before older adulthood.

The most recent systematic review and meta-analysis (2015) on the risk factors for development of knee OA in older adults (at least 50 years old) found that the most prominent risk factors were being overweight (defined as a body mass index (BMI) between 25 and 30 kg/m^2), being obese (BMI in excess of 30 kg/m^2), having a previous knee injury, and being of female gender [12]. The size of the increased risks of the patient characteristics were expressed in odds-ratios, and are listed in

Table 2.2: Most prominent risk factors for development of knee osteoarthritis

Characteristic	**Odds-ratio**	**95% confidence interval**
Female gender	1.68	(1.37, 2.07)
Overweight (BMI $\geq$ 25 kg/m^2)	1.98	(1.57, 2.20)
Obesity (BMI $\geq$ 30 kg/m^2)	2.66	(2.15, 3.28)
Previous knee injury	2.83	(1.91, 4.19)

table 2.2.

The diagnosis is partially set by assessing radiographic findings such as reduced joint space due to the loss of cartilage, a presence of subchondral sclerosis (the bone thickening repair function), and osteophytes [13]. To rule out competing diagnoses, the patient should fulfil at least one of three clinical criteria: age $>$ 50 years; morning stiffness $<$ 30 minutes; or creptius (popping sound in the joint) with active motion [14]. Patients with moderate and severe knee osteoarthritis have a very reduced, or totally closed joint space, and a high presence of bone thickening and overgrowth, contributing to considerable pain and reduced physical functioning [13].

2.2 Treatment

There exists no cure for osteoarthritis once it is present, even though considerable research has been sunk into developing disease modifying drugs targeting the biochemical imbalance suspected of at least partially causing the onset [15].

The treatment of knee osteoarthritis once present has three main goals; alleviate pain, adjust misalignment, and prevent progression. The appropriate type of treatment depends on the level of severity of the symptoms. With increasing severity, the effects of non-invasive treatments can become exhausted and the patient might eventually need arthroplastic surgery (joint replacement) [16].

In Norway, the Directorate of Health requires patients with both moderate and severe symptoms to have undergone and exhausted the effect of non-surgical conservative treatment before becoming candidates for arthroplastic surgery [17]. As we will learn, the exhaustion of non-surgical treatment might also have something

to do with the quality of the supply of this treatment, in addition to the actual effectiveness of it under ideal circumstances.

There are surgical interventions available that also conserves the patient's native knee such as arthroscopy and osteotomy [16], however, the former has in recent years been found to not be more effective than placebo [18] or non-surgical treatment [19]. Osteotomies are used to bend out the knee to correct a misalignment by cutting a wedge in the tibia [16].

For knees where only the inside of the joint is affected by the disease, there is an option to only replace that compartment. This procedure is the unicompartmental knee arthroplasty (UKA) [16]. In 2014, 552 patients in Norway underwent primary (first time) UKA, which roughly equates to 10% of the primary knee arthroplasties performed that year [5].

The total knee arthroplasty (TKA) is the procedure of replacing the entire joint with a prosthesis. Of the 5 500 primary TKAs performed in Norwegian hospitals in 2014, 80% were for the treatment of osteoarthritis. The most commonly inserted type of prosthesis in Norway during 1994 - 2005 was of the brand Profix for both the femur and tibia components, and in most cases a patella component was not used [5]. The Norwegian Arthroplasty Register, based at Haukeland University Hospital, has since 1994 published annual statistics on Norwegian knee arthroplasties, and thus the longest series of data is for the prostheses inserted this year. The durability of the prostheses, measured as the share of primary prostheses still intact years later, has improved since 1994. Despite this, the present analysis will in the decision model use the observed survival of the prostheses inserted in 1994. The reason is that the analysis has the life-time perspective, and it is preferable to keep the extrapolation beyond observations to a minimum. Another implication of the life-time perspective for this model is that given a failed primary prosthesis, the patient will need a new prosthesis, which is referred to as a revision TKA.

Non-invasive, or non-surgical treatment is often referred to as conservative because it is aimed at reducing symptoms while conserving the patient's native knee. Usually, non-surgical treatment is subdivided into pharmacological and non-pharmacological interventions. The Osteoarthritis Research Society International (OARSI) is comprised of an international expert panel, and their 2014 guidelines for evidence-based treatments of knee osteoarthritis [20] reviewed the appropri-

ateness of the most common treatments, based on a meta-analysis and subsequent expert consensus. The guidelines highlights the appropriateness of the non-pharmacological interventions water-based and land-based exercise (specifically strength training), self-management education, weight management, and biomechanical interventions (e.g. knee braces, and shoe insoles). Recommendations for pharmacological interventions have a lot more caveats to their appropriateness. The reason for this being that patients might have co-morbidities which makes a general recommendation difficult. Non-steriodal anti-inflammatory drugs (NSAIDs) are however, recommended for individuals without co-morbidities [20]. NSAIDs work by reducing the inflammatory response in the joint, but will usually not remove the pain and are associated with considerable risk of adverse gastrointestinal effects [16]. For this reason, OARSI recommends that NSAIDs should be co-prescribed with a proton-pump inhibitor [20]. Another class of pharmaceuticals, analgesics, are believed to be effective pain-relievers and is considered an appropriate pharmacological option for alleviating pain for knee osteoarthritis patients. Both the Norwegian electronic medical manual and the OARSI guidelines warn of the increased risks of adverse effects with long term used of both NSAIDs and analgesics.

In Norway, the primary health care service has the responsibility for managing osteoarthritis treatment, and the general practitioners (GPs) serve as gatekeepers for referral to physiotherapy and treatment in the specialist health care service [21]. A study on the self-reported quality of care of the conservative treatment provided for Norwegian knee OA patients found that while most patients had been referred to physical exercise and given pharmacological treatment, few had been referred to weight reduction help [21]. This is problematic when considering the fact that a large Norwegian 2012 population study [22] found that 72% of knee OA patients were overweight. Another Norwegian study found indications that patients were not given the recommended treatment in regards to the disease and treatment education, and self-management [4]. If patients are not given the recommended non-surgical treatment, one could argue that there is at the very least a potential that some patients can avoid surgical treatment, or postpone the surgery until after the effect of non-surgical treatment has been found to be exhausted – if they are given the recommended treatment.

The United Kingdom, the Netherlands, Australia, Sweden and Denmark have in recent years launched research projects where non-surgical interventions for hip and knee OA are provided in structured programmes [3]. An integrated approach is believed to have several potential advantages over more fragmented and unstructured provision, for example through higher quality of care, better patient satisfaction, and higher adherence to difficult life-style changes [3]. A Norwegian study where such an integrated conservative treatment model is provided is at present time being conducted in six municipalities [3]. The programme consists of an educational consultation with the general practitioner (GP) where the benefits of information, exercise and weight reduction is emphasised. The GP refers eligible patients to a physiotherapist (PT) who organises an 8 - 12 week group-based exercise programme, a healthy-eating programme and a three hour educational session about self-management of knee OA [3].

In this thesis the conservative interventions are modelled as a package based on a similar integrated programme from a 2015 Danish randomised controlled trial (RCT) of total knee arthroplasty versus a conservative treatment programme [6]. The conservative treatment was given in a two-step process: during the initial 12 weeks of the study the individuals randomised to conservative treatment participated in organised interventions, and for the remaining 40 weeks the individuals were encouraged to treat themselves. During this time, the individuals had intermittent telephone contact with the professionals they were acquainted with from the initial 12 weeks to keep them motivated and promote adherence.

The interventions were complacent with the OARSI guidelines and consisted of both non-pharmacological and pharmacological interventions. The first component was a neuromuscular training programme aimed at restoring a neutral functional alignment of the affected leg. The rationale being that the untreated misalignment will both be painful and progress the cartilage loss. The exercise was administered in a group directed by a physiotherapist and lasted for 12 weeks.

The second component was two 1-hour educational sessions to inform the participant of the disease, available treatment options, and self-help strategies.

If a participant's BMI exceeded 25 kg/m^2, they were given dietary advice for weight loss by a dietician in four one-hour sessions spread out over the initial 12 weeks. In the RCT, all of the individuals randomised to the conservative arm were

heavy, the mean BMI was infact 32, which is classified as obese. This is higher than the share of overweight Norwegian sufferers of knee osteoarthritis, hence an adjustment was made in the model for this.

The final non-pharmacological component to the programme were individually fitted shoe insoles by a physiotherapist to adjust for the misalignment of the knee to the foot.

58% of the individuals were at baseline given prescriptions for analgesics (acetaminophen), NSAIDs (ibuprofen), and a proton-pump inhibitor (pantoprazole) to be able to participate. At 12-month follow-up, 41% still regularly used the pharmaceuticals [6][1]. This is the best estimate that could be found for this specific patient group on such pharmacological treatment, and would therefore have to be modelled as in the trial.

The main outcome used from this study is the survival of the patients' native knee. Since the follow-up was only 12 months, the unpublished 24-month survival was obtained by contacting the main author of the study. See Appendix A for correspondence.

2.3 Analysis objectives

This master thesis has the objective of exploring the possibility of using a non-surgical programme as conservative treatment for patients who are candidates for arthroplastic surgery because of their knee osteoarthritis, and whether a such strategy could be considered a cost-effective alternative to directly undergoing primary total knee arthroplasty in a life-time perspective.

The thesis is developed as a cost-utility analysis by way of a decision analytic model, and the primary outcomes are quality-adjusted life-years and costs to the health care sector.

The research question is as follows; is a non-surgical conservative treatment programme compared to directly undergoing total knee arthroplasty a cost-effective alternative over the remaining life-time of the patients?

[1] As outlined in the RCT's technical appendix.

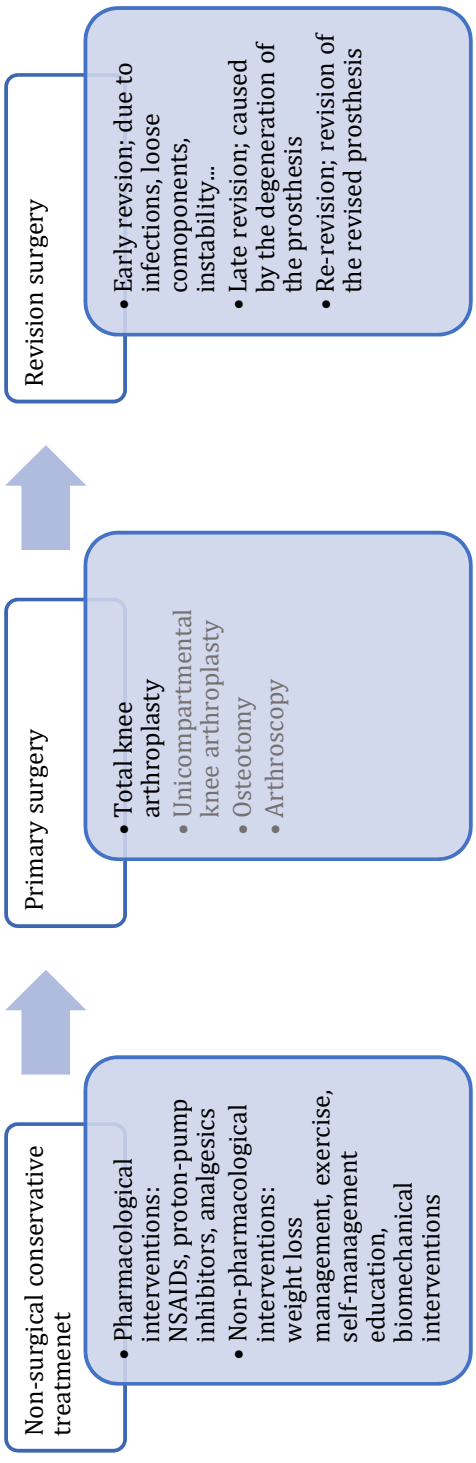


Figure 2.1: Treatment alternatives for severe knee osteoarthritis

To become a candidate for primary surgery the patient is required to undergo and exhaust the effects of non-surgical conservative treatment.

Chapter 3

Analytic framework

3.1 Economic evaluation

An economic evaluation is an analysis that compares the costs and consequences of alternative courses of action [23]. The purpose is to inform the decision maker to choose the alternative that maximizes the output under the constraint of available resources. In health care, all resources that enter into the production of health services have an opportunity cost, which is the value created at the next best alternative use of that resource. To be sure resources are allocated to the function where they maximize the production of health services needed, economic evaluations can provide appropriate measures and serve as a basis for decision making.

The three main types of economic evaluations all measure costs similarly, but are differentiable in how they quantify health benefits. The cost-effectiveness analysis expresses health benefits in natural units, such as life-years gained, or condition-specific outcomes, e.g. knee operations avoided. The cost-benefit analysis quantifies health benefits in monetary terms. The cost-utility analysis (CUA) is an extension of the cost-effectiveness analysis because health benefits are given as the product of gains in length of life and health-related quality of life (HRQoL). By using a CUA we can compare the resource allocation between all possible health problems. This is possible through the calculation of quality-adjusted life-years (QALYs) gained.

3.2 Outcomes of the cost-utility analysis

Calculating the QALY of living in a health state means weighting the time spent in that state by a utility-weight, indicating the HRQoL of living in the state. Initially there must always be someone, usually a sample of patients living in the target state, (for example severe knee osteoarthritis) describing the health-related quality-of-life. The description is usually delivered in a standardised system that is either condition specific like Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), or generic like the Euroqol group's EQ-5D [24]. The EQ-5D has five dimensions of health: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension has three levels: no problems, some problems, and extreme problems [24]. The advantages of using a generic system is that it is fast, cheap and enables the comparison of very different health states. It is of course less sensitive to very specific, albeit important changes, that condition-specific instruments can identify.

Quantifying the preferences for the HRQoL of one health state over another, requires either doing a valuation study, or using statistical modelling to adapt the results from an existing valuation study. One of the most commonly used methods for valuing described health states is the time trade-off (TTO). This method is designed to elicit the preferences of the participant for the described health state. The TTO presents individuals with a choice of living for a period t in the target heath state, or perfect health for a period x, where $x < t$. x is then incrementally shortened until the respondent is indifferent between the states, leading to the health-state utility value of $u = x/t$ [24].

Finally, to obtain the number of QALYs one can sum the products of the utility at a given time point with the cumulative survival at that time point. This gives us the area under the QALY-curve, the total number of QALYs.

Which costs are included in the analysis depends upon the perspective taken. Economic evaluations consider partial equilibria, i.e. we vary some factors while holding all others constant. Thus, even the widest perspective, the societal, excludes most parts of the economy in a general equilibrium. Usually, four categories of costs are considered in the societal perspective [23]. The first is costs to the health sector in production of the health service. These can be thought of as the

variable input factors and their prices (e.g. the wages of physicians and nurses) and the fixed input factors (e.g. the capital costs of equipment and facilities). Then there are costs to other sectors, sometimes less apparent but some programmes could consume resources from other public bodies or the voluntary sector. Another category is the cost to the patients and their families from giving up work time for seeking treatment, transportation costs, and out-of-pocket outlays. Finally, there is the cost of productivity loss, for the patient in the form of lost wages, and for the employer due to work not done.

While the societal perspective is always relevant, it is also the most time consuming and data demanding. Furthermore, if for example the consideration of patients costs simply confirm the result of just considering the direct costs to the production of the treatment, it might not be worthwhile including with respect to the decision [23].

For this thesis, the perspective taken is that of the Norwegian health care service. This means that costs are limited to the direct production of the treatment alternatives. The reason that the analysis employs such a limited scope is that resolving the initial hurdle of the lack of available data would have forced extensive use of assumptions and scenario analyses. While this is not necessarily problematic in and of it self, it would have entailed allocating much time away from other parts of the analysis.

The goal of a cost-utility analysis is to provide a measure of the cost-effectiveness of e.g. a new alternative relative to existing treatment. This measure is the incremental cost-effectiveness ratio (ICER), the incremental costs divided by the incremental effectiveness, here the incremental QALYs.

$$\frac{Cost_{new} - Cost_{exisiting}}{QALYs_{new} - QALYs_{exisiting}} = \frac{\Delta C}{\Delta Q}$$

When plotted in a cost-effectiveness plane, the ICER will fall in one of four quadrants. The origin of the cost-effectiveness plane is the position of the existing alternative. The y-axis illustrates the incremental costs of the new alternative, running south to north. The x-axis illustrates the incremental QALYs, running west to east. If the ICER is in the north-west quadrant, the new treatment is said to be dominated by the existing because it has higher costs and fewer QALYs.

An ICER in the south-east quadrant means the new alternative dominates, giving more QALYs for less cost [25]. The other two quadrants have an extra layer of decision attached in that most decision-makers would accept a reasonably higher cost if the effectiveness is higher, otherwise only equally costly or cost-saving new alternatives would be implemented. In the CUA, this acceptability is the willingness-to-pay (WTP) threshold value of cost per additional year in perfect health (QALY).

3.3 Decision analytic modelling

Where an economic evaluation performed alongside a clinical trial has access to patient level data of both the effects and resource consumption, a decision model usually does not. Decision models often rely fully on published literature for informing the effectiveness and costs of the alternatives [23].

In a decision analysis framework, events take place and consequences arise from those events in the form of changes to health and costs. A model developed to show these events and consequences indicates the likelihood of an event taking place in the future by probabilities [25]. We use these probabilities to calculate the expected values of alternatives in terms of costs and health outcomes. With a decison analytic model we wish to identify and compare the expected costs and expected utility of the new alternative versus the comparator,

$$E(Cost) = (p_1 C_1) + ... + (p_k C_k)$$

$$E(Utility) = (p_1 U_1) + ... + (p_k U_k)$$

The structure of the model in this analysis shows the experience of an average patient from assumed homogeneous patient cohorts, differentiated by age and gender. Since this analysis has a life-time perspective and many possible transitions, a state-transition model was developed.

The state-transition model, commonly known as the Markov model operates with a set of health states which the patient can be in at a point in time [23]. Individuals can transition from one state to another each time cycle, (e.g. by a

progressing disease, treatment, or death) depending on what type of state one is in at that time. Recurring states allow for not transitioning anywhere. Tunnel-states have individuals enter one cycle and leave the next, e.g. receiving surgery. Absorbing states have no transitioning option. Individuals transition between states by mutually exclusive, and collectively exhaustive transition probabilities [25].

Transition probabilities can either be assumed to be constant in time, or time-dependent. In either case they must all express the probability of transitioning within the same cycle length. In this model the cycle length is one year, hence all transition probabilities are annual. Time-dependent transition probabilites can be either dependent on time in the model, for example background mortality probability, or they can be dependent on time in a health state, such as the probability of transitioning from a successful primary knee arthroplasty to undergoing a revision of the prosthesis. These probabilities are comparatively more challenging to correctly represent without the patient-level data.

3.4 Representing time to events

Some of the most important transition probabilities in this model are centred on representing how much time passes until an event takes place. How long can patients keep their native knee? How long does a prosthesis, or a revised prosthesis last until it needs to be replaced? Are there differences in time to events given the observed patient heterogeneity? To answer these questions without access to the observations themselves, we can use information from survival analyses already performed and published.

The published literature relevant for the time-dependent transition probabilities frequently show the results of survival analysis with Kaplan-Meier curves. These curves represent non-parametric estimates of the survival function. This function is defined as the probability that a failure event T occurs after t [26],

$$S(t) = P(T > t) \tag{3.1}$$

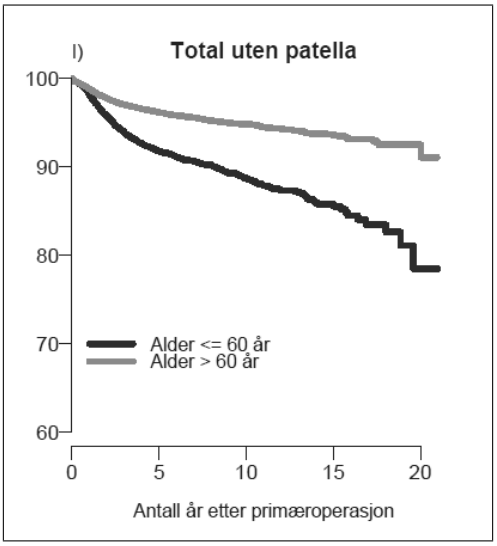


Figure 3.1: Survival of primary TKAs in Norway 1994-2014

$$\frac{\partial S(t)}{\partial t} < 0 \tag{3.2}$$

The Kaplan-Meier curve shows the proportion of study individuals at any point in time at risk for the event, but still event-free. In fig. 3.1 an example of important time-to-event information for this analysis is shown as Kaplan-Meier curves, as found in the literature. This is the survival of primary TKAs (the share of prostheses that are still intact) in Norway during 1994-2014 from the Norwegian Arthroplasty Register's 2015 annual report [5]. When only the curves are published and access to the input data is restricted, a commonly used approach is to attempt to recreate the raw data that went into producing the curves [27].

The data behind the relevant survival curves for this analysis was read off with special software ("DigitizeIt" [28]) that work by digitising the points on the curves through specifying the range of the axes. Digitised information could be used to calculate transition probabilities directly from the observations for the time of analysis, however, when extrapolation in time beyond the observations is necessary like for this model, this is not sufficient. By having information about the initial number of individuals at risk at time 0, one can approximate the raw data by using an algorithm to calculate the number at risk, and number of events at each time-step. These two pieces of information is the minimum requirement

for recreating the survival data. A central feature of survival data is that not all events are necessarily failure events; some of the observations that "go missing" are due to censoring; we do not know what happened, we only know when the individuals were last observed event-free [26]. Since the patients in this analysis are older, it is likely that some of the events are censoring due to death. By not having censoring information when recreating the data, one have to make the very strong assumption that there is no censoring [27]. Obviously, this means that the recreated data is not accurate. That said, this illustrates the reality of practical decision modelling; developing the model around the data access is generally believed to not be optimal, hence approximations like these are needed. This issue is also a reminder that the model is not intended to be an exact mathematical description of the world, it is a tool to guide the decision towards the most likely best alternative treatment.

A technical explanation of how the data was mapped and declared as survival data can be found in appendix A. All survival analyses were done in Stata (Stata Statistical Sofware: Release 14. College Station, TX: StataCorp Lp). Figure 3.2 shows the result of the process of recreating the raw data behind the curves in fig. 3.1 on page 16 as Kaplan-Meier curves. As visible, there is a lot less precision so the steps appear larger, but the overall trend is arguably well represented.

While a good face validation, the Kaplan-Meier curves have limited use for this research question. To make predictions beyond the observations required using the recreated survival data for estimating parametric models.

The hazard function, or hazard rate, gives the risk of a failure event, and takes into account how this risk changes in time; the probability of failure per time-step [26]. For the hazard rate[1] at a point in time t, the function can be expressed as the change in risk over the interval in time from t to $(t + \Delta t)$, when the interval becomes very small. Because time is a continuous variable, the probability of a particular value of time is thus equal to 0. By letting Δt approach 0 however, we

[1]Hazard rate and function are used interchangeably in the literature, however in this thesis, the hazard function refers to the general expression, and the hazard rate refers to the outcome of that function.

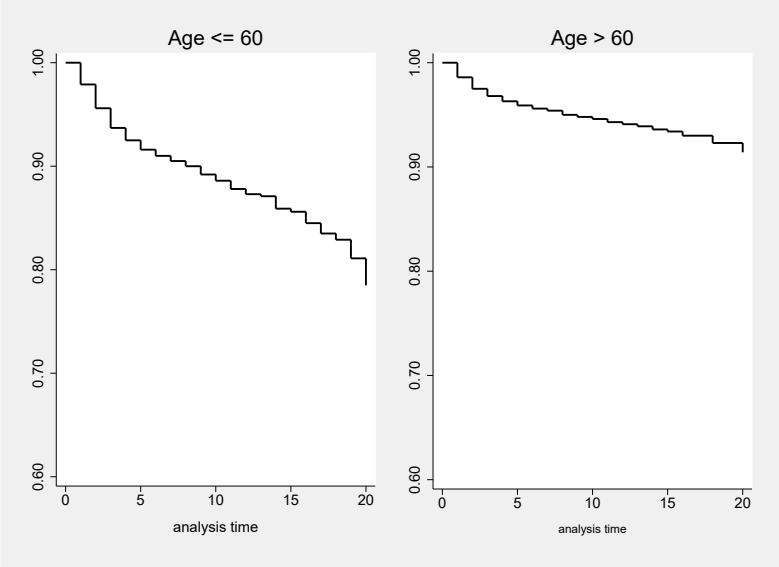


Figure 3.2: Recreated Kaplan-Meier for survival of primary TKA

get close to the hazard rate at the specific time-point t,

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t < T \leq (t + \Delta t) | T > t)}{\Delta t} \tag{3.3}$$

In other words, the hazard rate is the probability that the event T happens is in some time interval between t and (t + Δt), conditional on that the event T happens after t, over Δt [26]. The last part is very important, because this tells us that the hazard rate is not a probability, since for an infinitely small Δt, the value can exceed 1.

The Stata application `streg` was used to run regressions of the survival data under assumptions of different distributions of the mathematical functions for the hazard rate. These models were tested and compared for goodness of fit to the observed data. There was no "one fits all" distribution, so the whole range of distributions in the `streg` package was employed to inform the transition probabilities in the Markov model. Graphical evaluations and tests of the goodness of fit of each distribution for each model in the survival analysis can be found in appendix B.

The probability of transitioning from being at risk at time *t-u* to having a

failure event at time t can be expressed in terms of the cumulative hazard,

$$H(t) = \int_0^t h(u)\delta u \tag{3.4}$$

the accumulation of hazard from 0 up to t. Rewriting the survival function in terms of the cummulative hazard gives [25],

$$S(t) = e\{-H(t)\} \tag{3.5}$$

Using this we can express the discrete probabilities for transitioning from at risk to failure between the start and end of an interval of e.g. cycle *0* to *1*,

$$tp(t_0) = 1 - \frac{S(t_1)}{S(t_0)} = 1 - e\{\frac{-H(t_1)}{-H(t_0)}\} \tag{3.6}$$

or simpler,

$$tp(t_0) = 1 - e\{H(t_0) - H(t_1)\} \tag{3.7}$$

This is the function that was used to calculate time-dependent transition probabilities for all time-dependent parameters in the model.

Table 3.1 shows how the estimated hazard rates from the regressions should be used to obtain the survival function. The third column shows how the parameterization of the surival functions' scale should be calculated from the output. The fourth column shows the ancillary shape parameters of the distributions[2]. The least flexible distributional assumption is the exponential, which has a failure hazard that is constant in time, consequently giving constant transition probabilities.

[2]While the names of these distributions might conjure up images of very specifically shaped curves, it should be noted that the shapes are highly sensitive to the value of the ancillary parameters, and in most cases there were very little difference between them.

Table 3.1: Parametric survival functions under different distributional assumptions

Distribution	**Survival function, S(t)**	**Parameterization**	**Ancillary parameter**
Exponential	$e(-\lambda t)$	$\lambda = e(\beta x)$	
Weibull	$e(-\lambda\ t^{\gamma})$	$\lambda = e(\beta x)$	γ
Log-logistic	$(1+(\lambda t)^{1/\gamma})^{-1}$	$\lambda = e(-\beta x)$	γ
Lognormal*	$1- \Phi[(\ln(t)-\mu)/\sigma]$	$\mu = \beta x$	σ
Gompertz	$e[-\lambda\ \gamma^{-1}(e^{t}-1)]$	$\lambda = e(\beta x)$	γ

* Φ is the standard normal cummulative distribution function specified with $x = \ln(t)$

3.5 Accounting for model uncertainty

The results from an economic evaluation are uncertain. Especially so in decision analyses where the information is drawn from many different sources and choices must be made to structure the information. The uncertainty of all input parameters to a model stems from three sources. First order uncertainty concerns the concept of variability. From patient data there is always variations between the patients in terms of events happening, or not [25]. This is then accounted for by the variation around the mean parameter estimates.

Individual patient characteristics can be age, gender, risk factors, or prior diseases and can give systematically different outcomes. This is referred to as heterogeneity [25]. In this model, heterogeneity is dealt with by running the model for different patient cohorts specified by age and gender, the most detailed level of the input data.

Second-order uncertainty, or parameter uncertainty concerns how certain we are in the estimated values of the input parameters. Since the parameter values directly determines the results of the model we must analyse how sensitive the results are to variations in the parameter values.

The most robust way to explore the effect of variations in parameter values

on the results is by way of a method for probabilistic sensitivity analysis (PSA) called Monte Carlo simulation [29]. Here, each input parameter is allowed to vary randomly over their probability distributions, specific to each individual parameter. In other words, completely random values from their distributions are drawn. Then, the model's main outcomes, the estimated costs and QALYs, are recorded. Because of the random variation, the outcomes will be slightly different with each run of the model. *How different* is determined by the size of the variance around the mean of the input parameters, and ultimately the *joint variation* that results from having all parameters randomly drawn from their distributions. Repeating this process a large number of times produces a large sample of the outcomes based on the joint variations over the probability distributions of the input parameters.

For this model the outcome results from the PSA were calculated into ICERs, and shown as a scatter of probabilistic ICERs around the models deterministic ICER.

PSA was also used to calculate the net-monetary benefits of the two treatment strategies, shown as cost-effectiveness acceptability curves. This was done to indicate which of the strategies would have the highest probability of being the cost-effective alternative at varying willingness-to-pay thresholds per QALY.

Finally, the PSA results were used to calculate the upper monetary value of reducing uncertainty of the model; the expected value of perfect information.

To make the model input-parameters probabilistic appropriate probability distributions were assigned. The most commonly used distributions in decision models are the beta, dirichlet, gamma, and the normal [29]. For the model's utility parameters, the beta distribution was assigned. This is because it is constrained to the interval 0-1 [25]. The parameters were made probabilistic and beta distributed by specifying the α and β values. Since only the mean (μ) and variance was reported, the values were adapted by "the method of moments" [25],

$$\alpha = \frac{\mu^2 \cdot (1-\mu)}{s^2 - \mu} \qquad \beta = \alpha \cdot \frac{1-\mu}{\mu} \tag{3.8}$$

For multinomial parameters, i.e. those divided into a number of categories one cannot assume they are independently distributed. For example, constant transition probabilities from one state to four others. In these cases the dirichlet distribution

was assigned. For each parameter x_j in the category we draw a random value from the gamma distribution, Gamma(α_j,β_j). The α is scaled to an integer equivalent of the deterministic probability, and β equals the sum of all probabilities in the category, 1 – also scaled to an intenger number to keep the "aspect ratio" . This is done for all of the x's in the category (j to k). The randomly drawn new values for each health state in the category are then divided by the sum of the other randomly drawn new values, giving the probability π_j when all health states in the category are allowed to vary on the gamma distribution,

$$\pi_j = \frac{\alpha_j}{\sum_{j=1}^{k} \alpha_j} \tag{3.9}$$

For data known to be highly skewed, like costs, the gamma distribution is a typical choice. The gamma distributions α and β were also parametrised with the mean and variance by the method of moments,

$$\alpha = \frac{\mu^2}{s^2} \qquad \beta = \frac{s^2}{\mu} \tag{3.10}$$

For parameters resulting from regressions, such as the parametrisation of hazard and the ancillary shape parameters of the survival regressions we cannot assume they are independent and fit individual distributions [25]. To make these parameters probabilistic we use Cholesky decomposition. Here we would like to generate a correlated column vector of random variables, (x = y + Tz), from the regression output column vector, y.

$$y = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

The Cholesky equation requires two additional components; random draws from the standard normal distribution, z, and a decomposed version of the covariance matrix of the variables T, to connect the shape parameter to the scale parameter. Decomposing the variance-covariance matrix V into T,

$$V = \begin{bmatrix} var(x_1) & cov(x_1, x_2) \\ cov(x_1, x_2) & var(x_2) \end{bmatrix}$$

where T is a lower triangular matrix that would, if multiplied with its transpose produce V,

$$T = \begin{bmatrix} a & 0 \\ b & c \end{bmatrix} = \begin{bmatrix} \sqrt{var(x_1)} & 0 \\ [cov(x_1, x_2)]/a & \sqrt{var(x_2) - b^2} \end{bmatrix}$$

Letting Excel perform random draws from the standard normal distribution[3] we can generate the column vector z. The correlated, dependent parameters become,

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} (\beta_1 + a \cdot z_1) \\ (\beta_2 + b \cdot z_1 + c \cdot z_2) \end{bmatrix}$$

In addition to the parameter uncertainty, uncertainty about model structure was investigated by "what if analysis"; showing the effects of alternative choices when developing the model on the main outcomes of the analysis. The structural uncertainty analysis is given in chapter eight, while the PSA is included in chapter six.

[3] =*NORMINV(RAND(),0,1)*

Chapter 4

The model

A schematic presentation of the Markov model can be seen in fig. 4.1 on the following page. The health states that have a solid frame are recurring states, indicating a possibility of staying in the health state from one cycle to the next. States depicted with dashed frames are tunnel states, where staying is not an option. Two states are shown with a single arrow coming in, which illustrate the options of transitioning to them from multiple states. The red, and white states are absorbing. These have the interpretation of there being no transitioning-out option, (except for to the dead state). The included health states and model structure was validated by three clinical experts; a physiotherapist, a physician, and an orthopaedic surgeon.[1] This chapter concerns how the model functions; which states, and which transitions are assumed to describe the experience of an average patient with severe knee OA in the different cohorts. The chapter following this one includes the values for health states or transition probabilities, the inputs to populating the model.

[1]The initial model structure was presented to dr. Signe Flottorp and PhD. Stijn Van de Velde of the Norwegian Institute of Public Health (NIPH) on the 13th of January. The model was subsequently revised from suggestions. The revised model structure was on the 19th of January presented to orthopaedic surgeon at Rikshospitalet, dr. Albert Paus who is a specialist of revision surgery of hip and knee arthroplasty. Adjustments of the post-surgical states was made from the feedback. Final model structure was presented to the NIPH on the 10th of February.

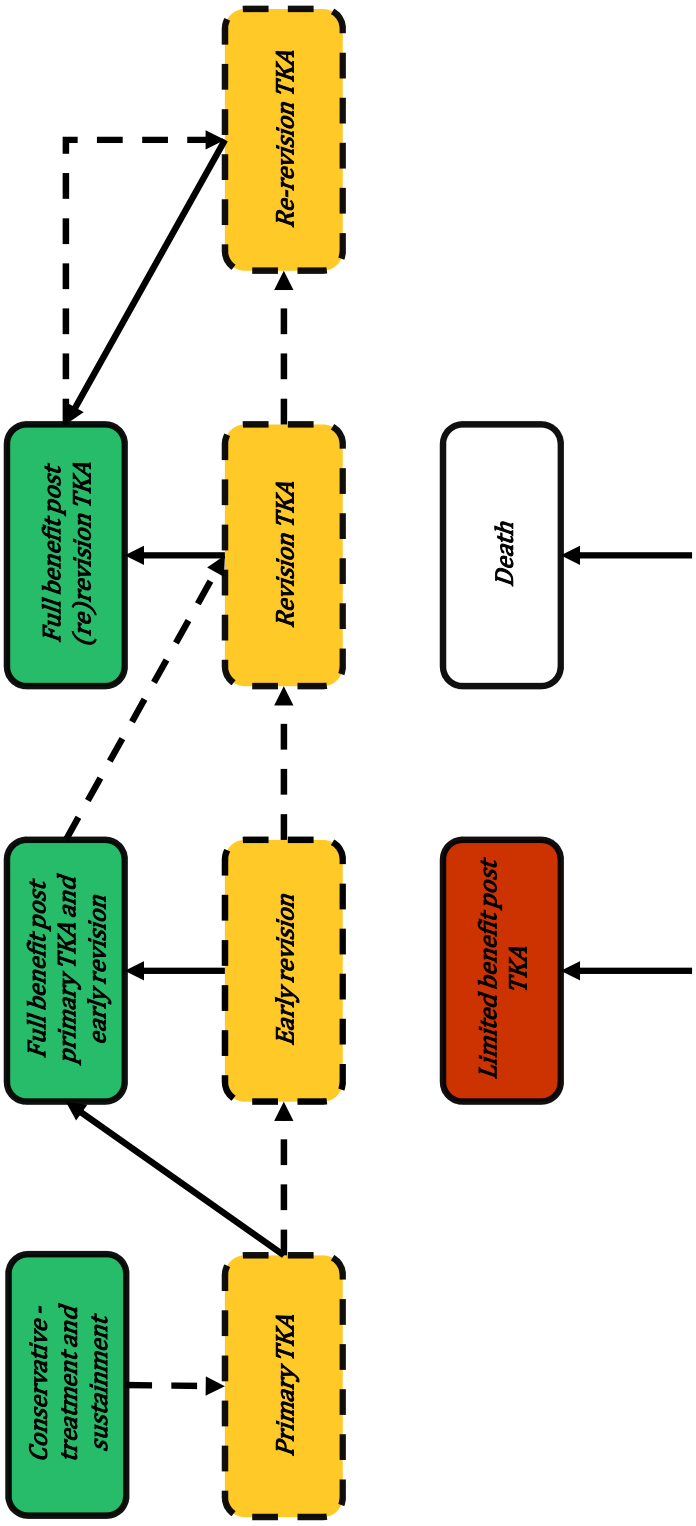


Figure 4.1: Schematics of the Markov model

The leftmost states (horizontally held) are alternative starting points of the model. Recurring states have the possibility of not transitioning onwards, tunnel states are for one cycle only.

4.1 Explaining the model and health states

All modelling work for this analysis was done using Microsoft Excel. The model has nine unique health states and two treatment strategies.

The non-surgical conservative programme consists of the interventions outlined in the background chapter: neuromuscular exercise, education, dietary advice if overweight, fitting of insoles, and usage of an NSAID, analgesic, and proton-pump inhibitor, if needed. The interventions are modelled as being delivered as a supervised package for the first three months, followed by self-management with telephone contact with professionals for adherence motivation. Some of the patients in the cohorts who started on the conservative programme will continue being in this state, sustaining, the next cycle. If so, they are not enrolled into the exercise, educational and dietary supervised sessions again. The assumption is that they will be equipped for self-management, but will still have telephone follow-up for adherence. While some continue, some will not experience any improvement of symptoms, and will need to undergo primary total knee arthroplasty. This is then a failure of the native knee.

As the Markov model is memoryless, once the knee fails and the patients undergo primary TKA, the model cannot distinguish these patients from those who started directly in the surgical strategy in the utilities experienced. However, there are distinct differences between the same cohorts on the different strategies in some key probabilities related to time in other states and time in the model. For example, for a cohort age 60 in the initial cycle starting out with the conservative programme, some will rather quickly experience a failed native knee and transition to the surgery. Others, will be able to keep it for many years, and thus enter the surgical state at a much higher age. This has implications for the *total mortality* of the patient, which will be discussed in more detail further on. Entering the surgical state at a later time than their identical cohorts in the competing strategy will also reduce these patients' risk of revision surgery according to the Norwegian registry data [5]. The reason being that on average older patients move about less than younger patients, and hence there is less direct strain on the primary prosthesis. Thus, while the model cannot tell the patients from the two strategies apart once they enter the states in terms of the utility, it can distinguish the different risk

profiles.

When starting directly with primary TKA, transitions are possible to one of the three subsequent states; full benefit, early revision, and limited benefit. Most patients transition to the full benefit state, indicated by their satisfaction with the result [30]. Some are not satisfied with the result and transition to the limited benefit state. This state holds those that are unsatisfied, and not wanting to undergo a revision. A third possibility after the primary TKA is to need an early revision. Early revisions are mainly due to infections, instability, loose components or other complications arising from the primary surgery [31]. Early revisions are modelled as its own state because they are not due to "wear and tear", a hallmark of the late revision[2]. Good satisfaction post early revision is assumed to lead to the full benefit post-primary TKA state. As with post-primary TKA, there is the option of not wanting to undergo more surgery when the result is not satisfactory, leading to the possible transition to the limited benefit state, following an early revision.

Those transitioning to a full benefit state post primary TKA (or post early revision) are assumed to not be in further need of health care services for their knee until a possible failure of the prosthesis leading to a (late) revision. In practice there is no limit to how many revisions a patient can have as long as the prosthesis can be fixed in the bone. Nevertheless, a cap was set in the model at one revision of a revised TKA, excluding early revisions.

For the death state there are two different probability sets used, depending on what type of state is causing the death. For non-surgical states the probability of dying is assumed to be equal to the 2014 Norwegian population age-dependent mortality rates, converted to probabilities. These probabilities are in the following referred to as *background mortality*, and are to be interpreted as caused by anything other than death directly attributable to the treatment for the osteoarthritic knee. For the surgical states there is a small, but nevertheless present, age-dependent excess mortality rate observed in the 30 days following primary hip and knee arthroplasty [32]. The probability of death in a surgical state is therefore the

[2]In figure fig. 4.1 on page 26, the arrow exiting early revision going to revision is in fact describing undergoing another early revision, the arrow is there for graphical coherency among tunnel states

combination of the background mortality and the excess post-operative mortality, i.e. the *total mortality*.

4.2 A formal look at the model

One can view the model as a system of health states defined as endogenous variables. The health states are determined by, or are functions of, exogenous variables, and other endogenous variables (other health states in previous cycles). The equations below shows how the model was built and structured in Excel, and this section aims to show the flow of patients (the cohort simulation) through the model's health states more formally.

As one cycle represents one year, the interpretation of the surgical tunnel states are therefore, initially undergoing the surgery and spending the rest of the year in a post-operative recovery situation before transitioning on. The simulation runs until the cohort is 100 years of age, at which point the assumption is that so few are still alive that modelling the experience of the remaining will not have much influence on the final results.

The logic is this: the states are, except the initial cycle, all functions of the previous cycle. More precise; any future behaviour of a Markov process X_{t+1} is only determined by the current state X_t, it is not influenced by additional information about past behaviour, X_{t-1} [33]. The implication is that we do not explicitly model the exiting from states, we model the entering from the other states the previous cycle. After cycle 0, the cohort size is fixed and new individuals do not enter into the model.

For the non-surgical conservative programme, cycle 0 and 1 are special because the conservative programme is structured differently for the initial and all subsequent cycles. Equations (4.1) - (4.4) account for these cycles of the conservative programme.

$$C_1 + T_1^{P,c} + D_1 = C_0 \tag{4.1}$$

$$C_1 = C_0[1 - \tau^B - \rho^c] \tag{4.2}$$

$$T_1^{P,c} = \rho^c C_0 \tag{4.3}$$

$$D_1 = C_0\tau^B \tag{4.4}$$

where,

$C_{0,1}$ is conservative treatment, intial cycle or subsequent sustainment.

$T_1^{P,c}$ is primary TKA following a failed knee.

ρ^C=f(t), (exogenous) is the probability of failure of the native knee, a function of time in the model .

τ^B=f(t,a,g), (exogenous) is the probability of background mortality, a function of time in the model, t (cycle), a (age at baseline + t), and g (gender).

The first equation shows us that the model is a closed system. The second equation shows the proportion of the initial cohort sustaining in the conservative programme, the third shows the proportion undergoing primary surgery, and the fourth shows the proportion dying. From cycle k, (from cycle 2 to a = 100), each cycle is still a function of the events in the previous cycle, (k-1), only. As cycle 1 opened up the possibility of undergoing primary TKA, the model becomes a system of the states;

$$C_k + T_k^P + ER_k + FB_k^P + LB_k + T_k^R + FB_k^R + T_k^{RR} + D_k = C_0 \tag{4.5}$$

such that, even though there are many more states added post cycle 1, the sum of the cycles must still always sum to the initial cohort.

$$C_k = C_{k-1}[1 - \tau^B - \rho^C] \tag{4.6}$$

$$T_k^P = C_{k-1}\rho^C \tag{4.7}$$

$$ER_k = T_{k-1}^P p(ER)^P + T_{k-1}^R p(ER)^R + T_{k-1}^{RR} p(ER)^R \tag{4.8}$$

$$\begin{aligned} FB_k^P &= FB_{k-1}^P[1 - \tau^B - \rho^P] \\ &+ T_{k-1}^P[1 - \tau^T - p(ER)^P - p(LB)^P] \\ &+ ER_{k-1}^P[1 - \tau^T - p(ER)^P - p(LB)^P] \end{aligned} \tag{4.9}$$

$$\begin{aligned} LB_k &= LB_{k-1}[1 - \tau^B] + T_{k-1}^P p(LB)^P \\ &+ T_{k-1}^R p(LB)^R + T^{RR} p(LB)^R \\ &+ ER_{k-1}[p(LB)^P + p(LB)^R] \end{aligned} \tag{4.10}$$

$$T_k^R = FB_{k-1}^P \rho^P \tag{4.11}$$

$$\begin{aligned} FB_k^R &= FB_{k-1}^R[1 - \tau^B - \rho^R] \\ &+ T_{k-1}^R[1 - \tau^T - p(ER)^R - p(LB)^R] \\ &+ T_{k-1}^{RR}[1 - \tau^T - p(ER)^R - p(LB)^R] \end{aligned} \tag{4.12}$$

$$T_k^{RR} = FB_{k-1}^R \rho^R \tag{4.13}$$

$$\begin{aligned} D_k &= D_{k-1} + \tau^B[C_{k-1} + FB_{k-1}^P + FB_{k-1}^R + LB_{k-1}] \\ &+ \tau^T[T_{k-1}^P + T_{k-1}^R + T^R R_{k-1} + ER_{k-1}] \end{aligned} \tag{4.14}$$

Exogenous variables

$$\tau^B = f(t, a, g) \tag{4.15}$$

$$\tau^T = f(t, a) \tag{4.16}$$

$$\rho^C = f(t) \tag{4.17}$$

$$\rho^P = f(t, a) \tag{4.18}$$

$$\rho^R = f(t, a, g) \tag{4.19}$$

$$p(ER)^P \tag{4.20}$$

$$p(ER)^R \tag{4.21}$$

$$p(LB)^P \tag{4.22}$$

$$p(LB)^R \tag{4.23}$$

where,

FB is the full benefit state following primary (P), revision(R), or re-revision (RR)

ER is the early revision state with same interpretation of superscrips as FB

LB is the limited benefit state

T^R is revision TKA

T^{RR} is re-revision TKA

τ^T(t,a,g) is the probability of death in the surgical states, total mortality, a function of time in the model, t, the cohorts age, a, and gender.

p(FB), p(ER), and p(LB) are the constant transition probabilities following a primary (P), revision and re-revision (R) TKA.

ρ^P(t,a,g) is the probability of revision TKA following a full benefit from a primary TKA, a function of time in the model, age, and gender.

ρ^R(t,a,g) is the probability of a revision following a revised TKA, also a function of time, age, and gender.

p(ER)^{P,R} and p(LB)^{P,R} are the likelihood of experiencing different outcomes following primary and (re)revision surgery.

For the TKA-directly strategy, the initial state is T_0^P. After the inital cycle, $T_t^P = 0, \forall\ t > 0$.

Substituting (4.6) - (4.19) into (4.5) thus gives the complete Markov trace of the model for one cycle.

The exogenous variables (4.15) - (4.19) are time-dependent transition probabilities; estimated using (3.7) in the previous chapter. To graphically see the transition between states, review the alternative Markov schematics in fig. 4.2 on the facing page that uses the same notation as this section.

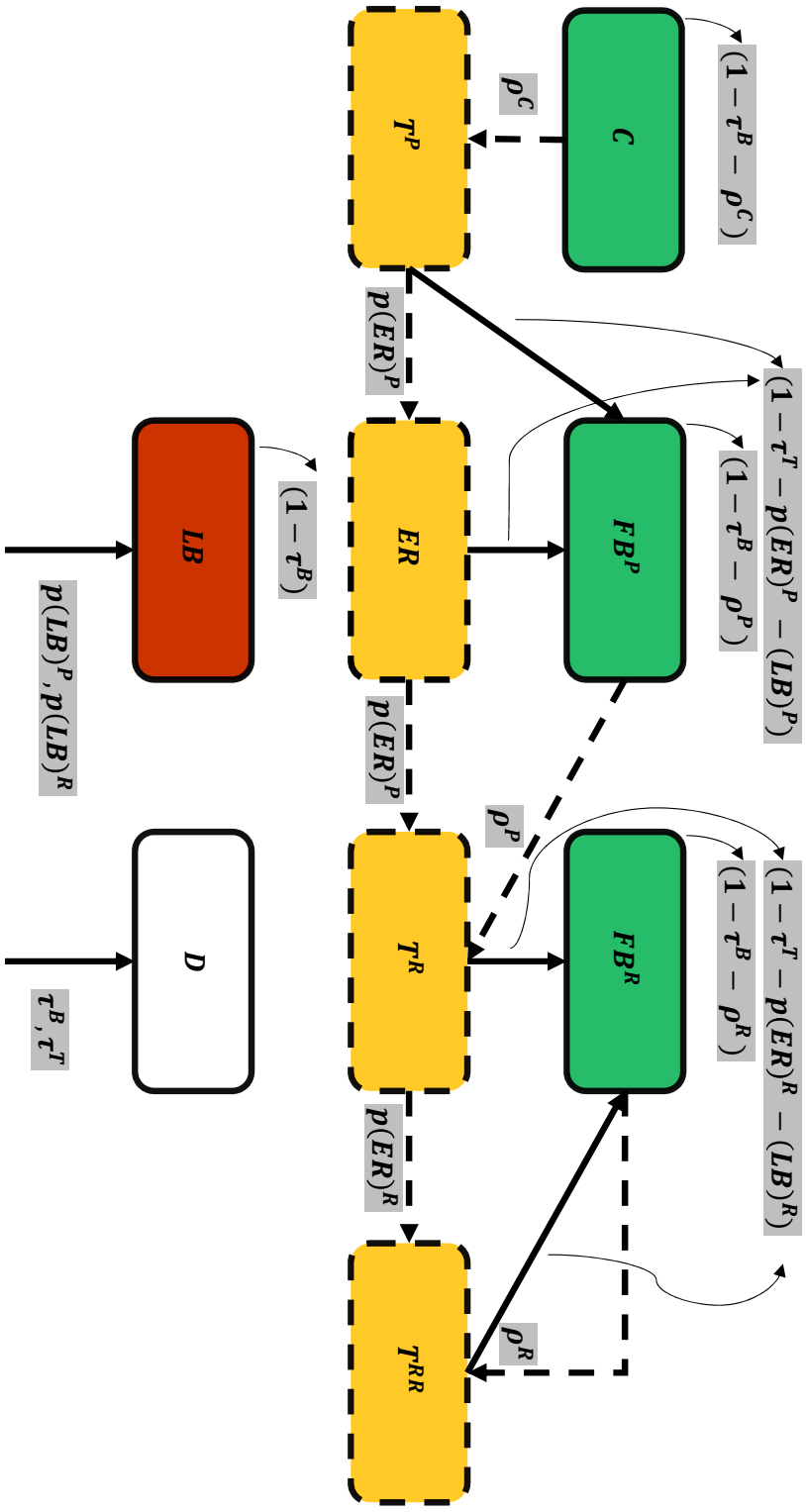


Figure 4.2: Transitions between health states

4.3 Some technicalities of the model

The model was run for eight separate cohorts of 1 000 assumed homogeneous patients. The eight cohorts were, females with age 50, 60, 70, and 80 at cycle 0 of the model, and the same age groups for males. As previously mentioned, the reason for this is to account for some possible patient heterogeneity.

In a schematic way, it makes sense to model transitions between states in discrete time-steps, such that the transitions will be occur at the end of the cycle. In reality, the transitions are of course continuous in time; they can take place at any time during a cycle [29]. To get closer to a more realistic portrayal, the recommended approach is to assume that on average, the transitions will take place half-way through the cycle. Practically this is done by removing half of the fist and last cycle (dividing by two) [34]. This procedure creates a potential problem in this analysis, seeing as the conservative programme has a continuous resource consumption structure, and the TKA-directly strategy mostly has a very large initial resource consumption in the first cycle. Halving the costs and QALYs of the first cycle would then effectively bias the analysis against the conservative programme in terms of costs. For this reason the decision was made to use half-cycle correction on cycle 1 in stead of cycle 0, and assume therefore that cycle 0 illustrates the enrolment into the strategies as a fixed investment to be borne. This approach is also suggested by Elbasha et al. [35].

Finally, for decisions that have implications long into the future, it is common to discount the future costs and effects, and to obtain a present value of future results. When discounting we need to adjust with a discount rate. This discount rate has the interpretation of the opportunity cost of investing in the programme. For effects, we can interpret it as the opportunity cost of waiting. The Norwegian guidelines for health technology assessments advices that costs and effects are discounted at the rate of 4% [36]. This effectively gives less weight to the events that take place far into the future for the expected outcomes. Again, considering the difference in timing of events of the comparators, the relative cost-effectiveness might be affected by discounting. This is further investigated in chapter eight.

Chapter 5

Populating the model

5.1 Transition probabilities

The transition probabilities in this thesis can be divided into two classes; those indicating the time to an event, and those indicating the likelihood of a certain outcome following an event. One can also think of them as time-dependent and time-independent, or constant.

The constant transition probabilities in this model concern the likelihood of experiencing one of three outcomes after an arthroplasty: full benefit, early revision, or limited benefit. These probabilities were informed by the experience of Swedish patients undergoing primary TKA in 2013. In the Swedish Knee Arthroplasty Register (SKAR) annual report of 2015 [30], the satisfaction with the one-year post-TKA visual analogue scale (VAS) scores (0-100) are given for 2 191 patients. Of these, 85% were satisfied to very satisfied with the results, while 15% were moderatly satistfied to very unsatisfied. The orthopaedic surgeon consulted with for expert opinion suggested the following calibration of that data to the model: 80% would transition to the full benefit state, 5% would go directly to early revision due to complications such as primary infections. From the remaining 15% roughly a third would "accept" their benefit and could be considered transitioning to full benefit (corresponding to the moderately satisfied patients in the SKAR report), another third would need to undergo revision surgery to e.g. investigate the source of pain, and finally the last third would have "had enough" and choose no further

Table 5.1: Constant transition probabilities post surgery

	To		
From	**Full benefit**	**Early revision**	**Limited benefit**
Primary TKA	0.85	0.10	0.05
Revised TKA	0.765	0.17	0.065
Re-revised TKA	0.765	*	0.235

*Assumed to transition to limited benefit

treatment. While these probabilities are highly dependent on expert opinion, one could argue that the Swedish satisfaction values supports them.

For the constant probabilities of the possible outcomes after (late) revision TKA, and re-revision TKA, published information available was very sparse. Hence, these transition probabilities rely solely on expert opinion. The likelihood of transitioning to the full benefit state post revision TKA is assumed in the model to decrease in relation to the equivalent post primary TKA. For early revsion and limited benefit the assumptions are increased probabilities compared to the outcomes after the primary TKA. The constant transition probabilities used in the model are shown in table 5.1. The decision was made to allow only for one re-revision, hence the probability of early revision from this state is collapsed into the probability of transitioning to the limited benefit state.

The time-dependent transition probabilities concerning time spent in a health state are, as explained in the previous chapter, defined by the estimated cummulative hazard. In table 5.4 on page 39 the parameter values that define the hazard rates of these parameters are shown. These rates were inserted into equation (3.7) in section 3.4 on page 19, to derive the transition probabilities from one cycle to the next.

The hazard rate of the patients' native knee in the conservative programme was estimated by a lognormal parametric survival model. In addition to the lognormal, the assumption of constant failure hazard with an exponential model was an option, albeit with inferior fit to the observed data compared to the lognormal. In fig. 5.1 on page 40 we observe that the lognormal model gives less optimistic predictions beyond the two-year follow-up than an assumption of constant failure hazard. A

rationale for choosing the lognormal irrespective of model fit is that the data source is an RCT. Even though RCTs are considered the gold standard of causal inference experiments, they ultimately create an artificial setting with high internal validity and unknown external validity. In other words, it could be somewhat uncertain what happens under real-world conditions. Since cohort study evidence (real-world conditions) are in the pipe-line [3], but not yet published, a more careful (less optimistic) prediction could be warranted from that perspective. The effect of instead choosing an exponential (constant) model is investigated in chapter eight. In the graph we observe that independent of distributional assumption, the prediction after less than five years in the health state about half of the cohort will have undergone a primary TKA. After 10 years, the prediction under a log-normal distributional assumption shows that more or less the entire cohort will have undergone arthroplastic surgery.

The probability of transitioning to death is a different type of time dependency since it relies not on the time in a state, but on time in the model. For the background mortality, the data was acquired from Statistics Norway's life tables for 2014 [37]. For total morality, the excess post-operative mortality was adapted from Lie et al. [32] and the rates in the life tables used for background mortality. These rates were first converted to monthly probabilities, and then to annual rates to "stretch" them out to fit the cycle length. Following this, they were converted back to rates and added on to background mortality rates, and finally calculated to annual age-dependent probabilities. The decision was made that adding the rates was appropriate since the post-surgical component was controlled for baseline mortality in the original study. The transition probabilities are given in five-year age brackets, and can be found in1 tables 5.2 and 5.3. As visible, the differences are quite minuscule.

Table 5.2: Annual transition probabilities for background mortalitiy

Age bracket	**Females**	**Males**
50-54	1.1%	1.5%
55-59	1.6%	2.3%
60-64	2.4%	3.7%
65-69	3.9%	5.7%
70-74	5.7%	8.6%
75-79	9.5%	12.6%
80-84	14.8%	17.1%
85-59	19.6%	18.3%
90-94	19.3%	14.4%
95-99	10.5%	4.9%

Adapted from Statistics Norway's 2014 Life-table rates [37].

Table 5.3: Annual transition probabilities for total mortality

Age bracket	**Females**	**Males**
50-54	1.4%	1.8%
55-59	1.9%	2.6%
60-64	3.1%	4.4%
65-69	4.6%	6.4%
70-74	7.1%	10.0%
75-79	10.8%	13.9%
80-84	16.1%	18.3%
85-59	20.8%	19.6%
90-94	20.5%	15.7%
95-99	11.8%	6.3%

Adapted from Lie et al. [32] and Statistics Norway [37].

Table 5.4: Results from survival analysis to parametrise time to failure of model components

Failure variable	**Parameters**	**Coefficient**	**Std. err.**	**Dist.**
Native	μ	0.900	0.144	
knee	σ	0.671	0.130	Lognormal
Primary	λ	0.010	0.137	
TKA, age $\leq$ 60	γ	0.011	0.012	Gompertz
Primary	λ	0.003	0.019	
TKA, age $>$ 60	γ	1.011	0.000	Log-logistic
Revised	λ	0.059	0.088	
TKA, female $\leq$ 60	γ	-0.061	0.015	Gompertz
Revised	λ	0.047	0.097	
TKA, female 60-70	γ	-0.068	0.016	Gompertz
Revised	μ	3.802	0.127	
TKA, female $\geq$ 70	σ	1.983	0.106	Lognormal
Reivised	λ	0.098	0.080	
TKA, male $\leq$ 60	γ	1.110	0.030	Weibull
Revised	μ	2.003	0.045	
TKA, male 60-70	σ	1.303	0.039	Lognormal
Revised	λ	0.074	0.078	
TKA, male $\geq$ 70	γ	-0.046	0.013	Gompertz

Model appropriateness was assessed by graphical evaluation and Akaike's Information Criterion. See Appendix B for the material.

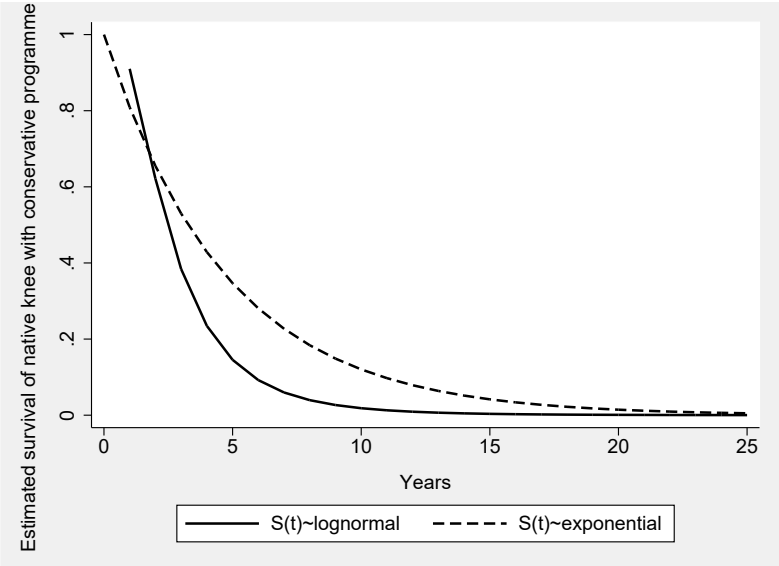


Figure 5.1: Predicted survival of the native knee in the conservative programme

5.2 Health state utility values

Literature searches for systematic reviews on the health related quality of life for patients with knee OA were performed on the 20th and 25th of February in the ORIA database accessible through the university library of the University of Oslo. Inside ORIA one can access all mainstream databases for medical journals. The objective was to find studies in the systematic reviews that fit the criteria

- Patients with moderate to severe osteoarthritis
- Reported utility values, or EQ-5D vector values (the raw descriptions)
- Measure of variance around utility values
- Population and setting similar to the Norwegian
- Prospective design
- English or Scandinavian language

The most recent systematic reviews from 2015 [38] and 2004 [39] included no studies that fit the criteria so the search strategy was adapted to focus on other economic evaluations of knee OA. Three CUAs were relevant, however none of referenced sources for their utility values that were of good enough quality. However, two were found to be informative to adapt the utility of states not described by

patients, by using their expert opinions [40, 41].

The 2015 annual report from the SKAR referenced a study that was found to fit the criteria. This study informs the baseline utility of living with severe knee OA in the model, and the utility for the full-benefit post primary TKA state. In this study, Jansson and Granath [42] collected descriptions from 365 patients with severe knee OA at the Karolinska Institute in Stockholm prior to undergoing primary TKA, and at 12 months post surgery. They used the EQ-5D and a tariff from a large British population survey.

The improvement in HRQoL for patients on the conservative treatment programme was derived from the Skou et al [6]. These utility values were the EQ-5D responses of the conservative treatment arm, and were adjusted with a Danish valuation tariff. The improvement was then added to the baseline utility value from the Swedish study. For the other health states in the model there were found no estimates of HRQoL as described by patients, and the cost-effectiveness analyses identified in the literature search all used assumptions for such states.

The disutility experienced from undergoing primary TKA and revision TKA was assumed to reduce baseline utility by 10 and 20% respectively. These adjustments were taken from the study by Solver et al. [40] where they were informed from expert opinion. Both early and later revisions were assumed to yield the same utility during the cycle of surgery and recoverey. The utility for the full benefit post revision state was assumed to increase baseline utility by 25%, also taken from the expert opinion in Solver et al. For the utility in the limited benefit post TKA state, the assumption was a 25% decrement of baseline utility, the assumption used by Losina et al. [41]. All health state utility values were assumed to be constant in time. It was not possible to give distinct utility values to the different modelled patient cohorts, thus all cohorts are assumed to experience the same average utility levels, regardless of age and gender.

5.3 Health state costs

As previously mentioned, the scope of the costs included were limited to the costs of the direct resource consumption of the two strategies.

The cost of the surgical states were assumed to equal the unit price for a

Table 5.5: Health state mean utility values

Health state	**Mean utility**	**Std dev**
Baseline severe knee osteoarthritis	0.510	0.33
Maintaining on conservative programme	0.625	0.18
Undergoing primary TKA	0.410	0.33
Undergoing (re)revision TKA	0.310	0.33
Full benefit post primary TKA	0.730	0.27
Full benefit post revision TKA	0.638	0.33
Limited benefit post primary/revision TKA	0.383	0.33
Death	0	0

diagnosis-related group (DRG) point of 41 462 NOK, weighted by the surgical DRG-weights of 3.024 for primary TKA, and 4.723 for revision TKA [43]. The DRG system is the activity based reimbursement scheme of the owner of the Norwegian hospitals, the Regional Health Authorities. DRGs are intended to cover about 50% of the total operating costs of a hospital stay, while the rest is financed by block grants from the government according to the demographical situation of the catchment area of the hospitals [43]. Since the block grants are given independently of activity, the DRGs were assumed to cover the costs of the surgeries on average.

While the costs of the surgical health states could be found directly, the costs of the conservative treatment had to be estimated from the resource use of the programme used in the RCT [6]. The relevant unit costs were identified and adjusted to 2015 prices in Norwegian kroner (NOK), and multiplied with the resource consumption of the units of the programme. The entirety of the cost estimation for the conservative programme is rather lengthy and is therefore given in appendix C.

In table 5.6 on the next page, a simplified listing of the per cycle costs of the resource demanding health states can be seen. The full benefit health states were assumed to need no resources in the primary and specialist health care service. A great deal of the variation in the costs of the conservative programme is attributable to the need for the pharmacological treatment, and participation in the weight loss programme. The need for the latter was assumed to be given

Table 5.6: Per cycle health state cost

Health state	**Mean cost (NOK)**	**Std dev**
Conservative initial programme	15 209*	**
Conservatie continuous programme	2 285*	**
Primary TKA	125 381	***
Revision TKA	195 825	***
Limited benefit	2 397	2 397

* Shown for patient in need of pharmacological treatment, and BMI $\geq$ 25.
**Because the cost varies both according to the variance of the probabilities of needing the additional costs of medication and dietary advice, and the variance around the mean amount of pain medication needed, this standard deviation is not possible to set in one value. In the model these are separate parameters allowing for the conditional probabilities to determine the cost.
*** While the (unknown) actual cost to the individual hospitals deviates from the DRGs, the model relies on the average, thus there is no variation.

by the share overweight Norwegian knee OA patients discussed earlier. This was included as its own parameter so the probabilistic sensitivity analysis could show the variation.

In the RCT, 58% of patients needed pharmacological treatment at baseline, and 41% at the end of cycle 1 [6]. The assumption for the remainder of the time in the conservative programme was that 41% of the patients would be in need. This was also included as its own parameter to inform variation in the probabilistic model.

For the limited benefit health state, the assumption is that patients receive the pharmacological components of the conservative programme. Since it was not possible to find any estimates of how long through a cycle patients would need the medication, a decision was made to set the lower assumption at 0 and the upper assumption at the whole cycle, while realistically the patients would be advised to use the pharmaceuticals only intermittently, due to the heightened risk of adverse effects [16]. The uncertainty around this estimate is thus given by the parameter informing the share of patients in need of the pharmaceuticals, and the upper and lower assumptions of how long through the cycle a given patient in need would be using the pharmaceuticals.

Chapter 6

Results

6.1 Deterministic results

Table 6.1: Expected per-patient life-time discounted direct treatment costs and QALYs gained

Modelled	**Conservative**		**TKA only**	
cohort	**Costs (NOK)**	**QALYs**	**Costs (NOK)**	**QALYs**
Female				
age 50	163 423	9.0	172 761	9.2
age 60	148 286	6.7	162 653	6.9
age 70	120 463	4.5	141 376	4.7
age 80	94 428	2.6	139 473	2.7
Male				
age 50	155 620	8.1	165 589	8.3
age 60	142 233	5.7	162 016	5.9
age 70	111 193	3.7	140 828	4.1
age 80	89 980	2.6	139 519	2.7
All cohorts	1 025 626	43	1 224 215	45
Average cohort †	128 203	5.4	153 027	5.6

† All cohorts / 8

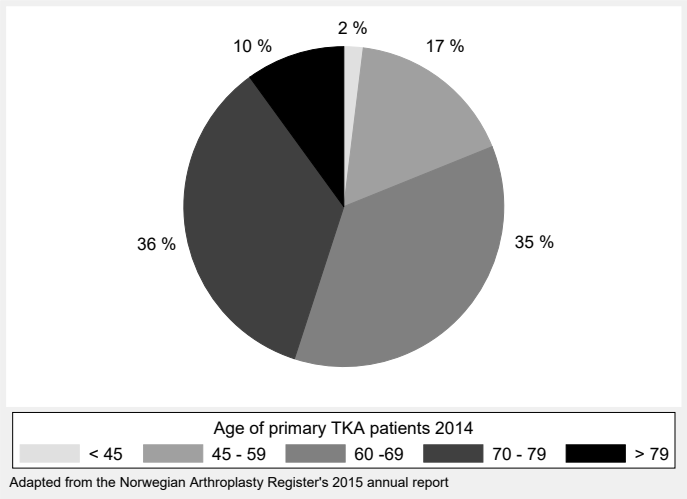


Figure 6.1: Age distribution of patients undergoing primary total knee arthroplasty in Norway 2014

The deterministic results, i.e. those obtained from the mean values of the input-parameters are given in table 6.1 on page 45. For every cohort, the conservative programme has lower expected discounted direct treatment costs. Thus it appears that the conservative programme is a cost-saving alternative compared to the TKA-only strategy when considering expected life-time direct treatment costs. Also for every cohort, the discounted quality-adjusted life-years gained is lower when cohorts are run through the conservative strategy than the TKA only strategy.

A decision-maker considering the average cohort could also consider not giving equal weight to all cohorts, seeing as some have a lot more patients than others. One way to weight could be by using the TKA-patients of 2014 [5] as a proxy for the patient group as a whole, as in fig. 6.1. Collapsing the youngest age groups together, and gender groups together, the weighted average of the incremental costs becomes – 10 819 NOK, and the weighted average of the incremental QALYs – 0.11. In the following, however, the average cohort is the unweighed average.

In appendix D, estimated clinical outcomes in terms of person-years are given. This slightly unorthodox way of showing model outputs gives an opportunity to explain some of the observed differences between the cohorts. The tables show the expected clinical consequences of the strategies for each cohort, expressed in

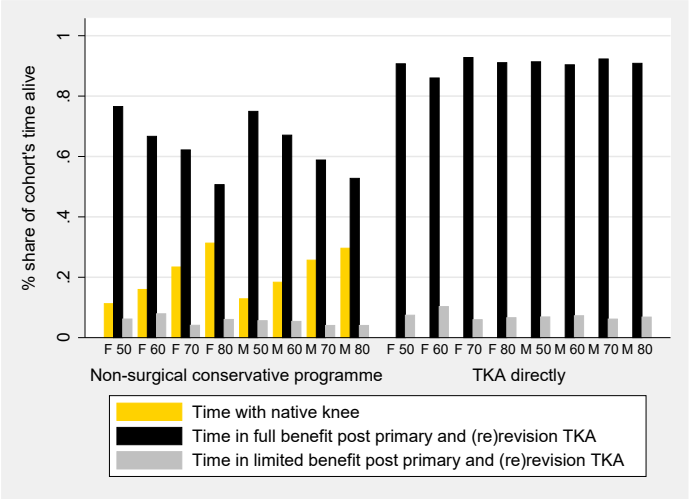


Figure 6.2: Estimated time in health states relative to time alive for the cohorts

number of person years alive, i.e. total number of person-years less person-years in the the dead cycle, over all cycles until the cohort reaches the age of 100. The clinical consequences shown are person years in the state keeping the native knee, in the full – and limited benefit states, and the dead state. In fig. 6.2 the findings are summarised graphically. Time spent in the state conserving the native knee is relatively longer (share of the cohort's person-years alive) with higher age. Interestingly, there are large differences in time spent in the full benefit state according to treatment strategy. For the conservative strategy, time in the full benefit state decreases with age, and for the TKA-only strategy, time in this state is much more stable across cohorts. For time spent in the limited benefit state, there are only small differences between the strategies. The implications of these findings for the overall cost-effectiveness of the non-surgical conservative programme is discussed further in chapter nine.

As for the cost-effectiveness of the conservative programme compared to TKA only, a good starting point is to review where in the cost-effectiveness plane the deterministic incremental cost-effectiveness ratios (ICERs) of the cohorts reside. In fig. 6.3 on page 49 we observe that all cohort ICERs are in the south-west quadrant. The interpretation should be that the relative cost-effectiveness of the conservative programme is fully dependent on the decision maker's objectives and acceptability threshold. A decision maker with purely a cost-saving objective would not care if

the QALYs gained were less, thus the conservative programme would be considered the cost-effective strategy for all cohorts. Consider a line in the chart tangent to the x-axis, where everything below would be acceptable because of the cost-saving quality. The slope of this line is the decision maker's willingness-to-pay (WTP) per QALY, call this line λ. For $\lambda > 0$, the line would rotate counter-clockwise about the origin, while the condition for cost-effectiveness stays unchanged, it must be below the line. For a $\lambda = \infty$, the line would be tangent to the y-axis. In this case the conservative programme would never be the cost-effective strategy since it cannot be below λ[1].

To assess the cost-effectiveness then, we would like to know which strategy would give the most health per invested NOK. Since ICER the does not have a good interpretation when positioned in the south-west quadrant [44], a better option is to avoid the ratio altogether. Instead we use the metric net monetary benefit [25], defined as,

$$NMB = \Delta Q \cdot \lambda - \Delta C \tag{6.1}$$

The NBM allows assessing which strategy would be considered cost-effective at different λ-values and simultaneously inform the certainty of this finding. We do this by including the results from the probabilistic sensitivity analysis into the NMB, which allows illustrating the probability that each strategy is the cost-effective alternative at the varying λ values.

[1]While Norway does not have an official value of its willingnes-to-pay per QALY one can be sure it is $0 < \lambda < \infty$. 500 000 NOK was used as a high λ-value for this analysis, however this is somewhat arbitrary and might even be considered being in the low-to-mid range in other analyses.

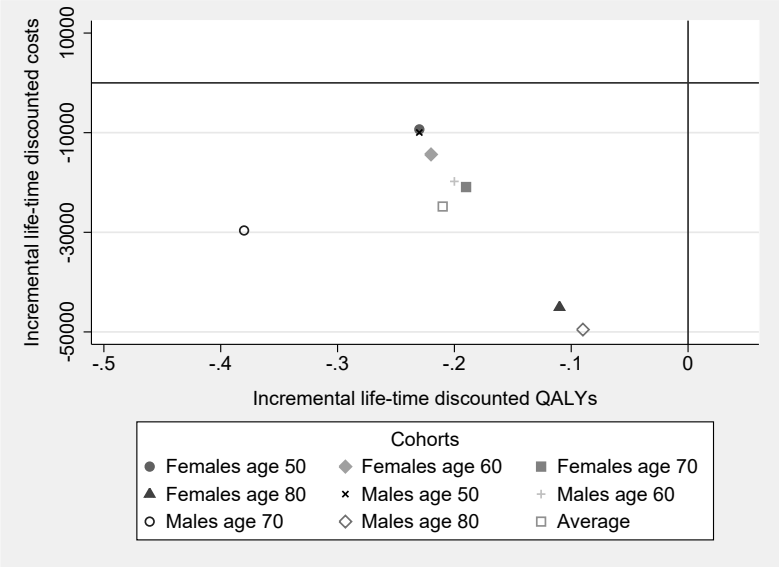


Figure 6.3: Deterministic incremental cost-effectiveness ratios

6.2 Probabilistic results

For each cohort, and the average cohort, the model was run with all parameters randomly picked from their probability distributions to test the sensitivity of the deterministic results. In fig. 6.4 on the following page the incremental costs and incremental QALYs are shown for 1 000 probabilistic runs of the model. As visible, a fair amount of the simulations of the ICER reside in the south-east quadrant. For these simulations the conservative programme is a dominating strategy. From the simulated cost and QALYs, 1 000 NMBs was calculated for both strategies as well. By varying the λ-value from 0 to 500 000 NOK, it was possible to calculate which strategy had the largest NMB, for each simulation. This value was recalculated to a binary variable for each strategy, where the value 1 indicated the case where the NMB of the strategy was the larger of the two, and 0 otherwise. Averaging the NMB of each simulation at different thresholds gave the probabilities that each strategy was cost-effective at those thresholds. By plotting these probabilities at the different λ-values we get the cost-effectiveness acceptability curves (CEACs) of the strategies. For the average cohorts' simulations, the curves are plotted in figure fig. 6.5 on page 51. Up-until a λ of 117 000 NOK, the conservative strategy had the highest probability of being cost-effective, based on the simulations. Beyond

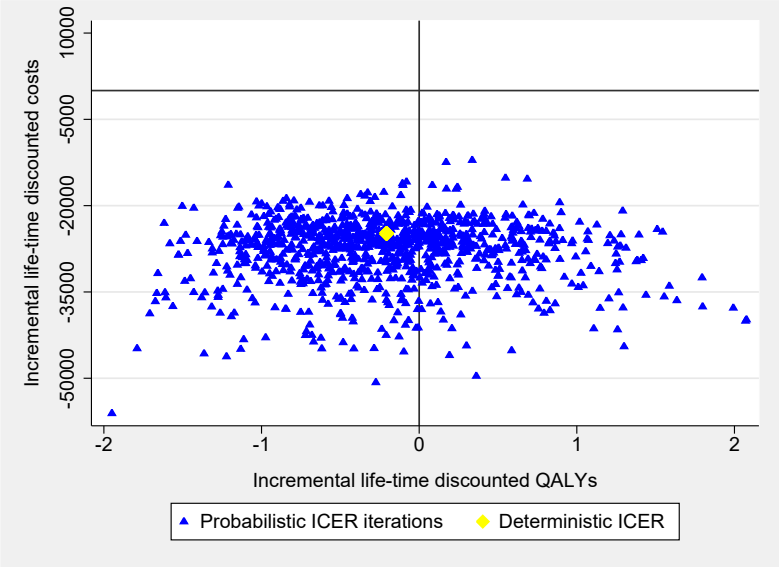


Figure 6.4: Scatter plot of the deterministic ICER with probabilistic simulations for the average cohort

this value, the TKA only strategy always had the larger probability of being cost-effective. The general rule for interpreting these curves is that the farther apart the CEACs are at a given λ, the more certain one can be that implementing the one with the highest probability, would be the cost-effective strategy at that level[2].

While it appears that the TKA-only strategy have the highest probability of being cost-effective from a relatively low WTP-threshold when looking at the patients as an average cohort, this does not represent the situation for all cohorts. Figure 6.6 shows the CEACs of the four female cohorts, and fig. 6.7 on page 52 shows the same for the male cohorts. The cohorts of 50 – and 60 year-olds at baseline for both gender are well represented by the average cohort's CEACs. For the cohort of 70 year-old males, the TKA only strategy has a markedly higher probability of being cost-effective even in the low range of willingness-to-pay, compared to the overall result. For 80 year-olds of both gender, the choice appears highly uncertain.

[2]The CEACs do not asymptote to 1 and 0 with very high λ's because even at a $\lambda = \infty$, a fair proportion of the simulations of the ICER in the south-east quadrant would still not be below the willingness to pay per QALY.

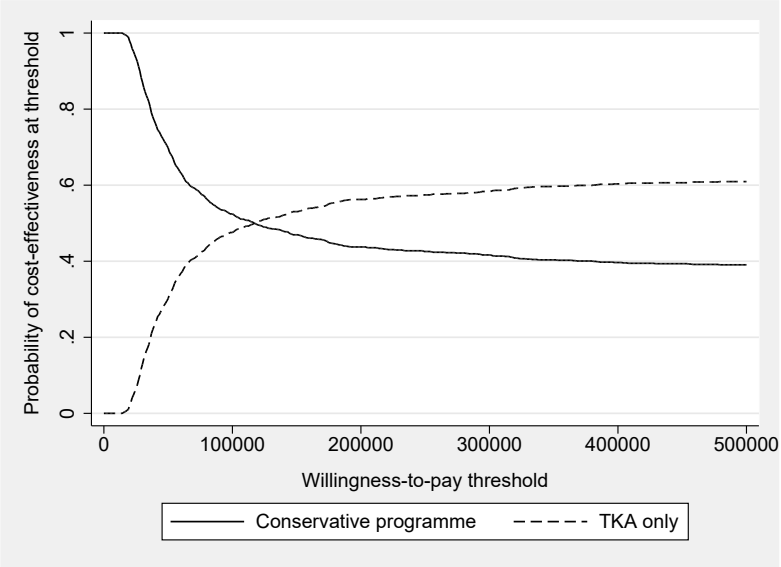


Figure 6.5: CEACs for the average cohort

High uncertainty could be interpreted as high probability of implementing the wrong strategy. This comes at the cost of health benefits forgone, and resources wasted. Further, if knowledge of what was causing the uncertainty for these cohorts was available, we could gather new information to reduce the uncertainty around the implementation decision. This leads appropriately into the concept of the value of information for this model, which is dealt with in the following chapter.

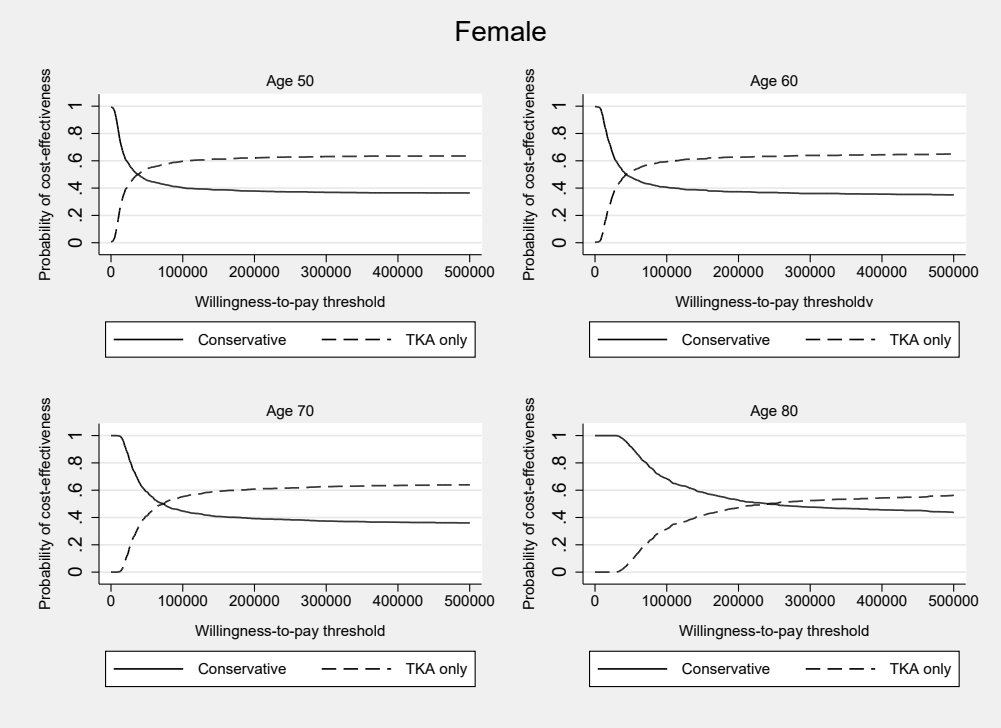


Figure 6.6: CEACs for female cohorts

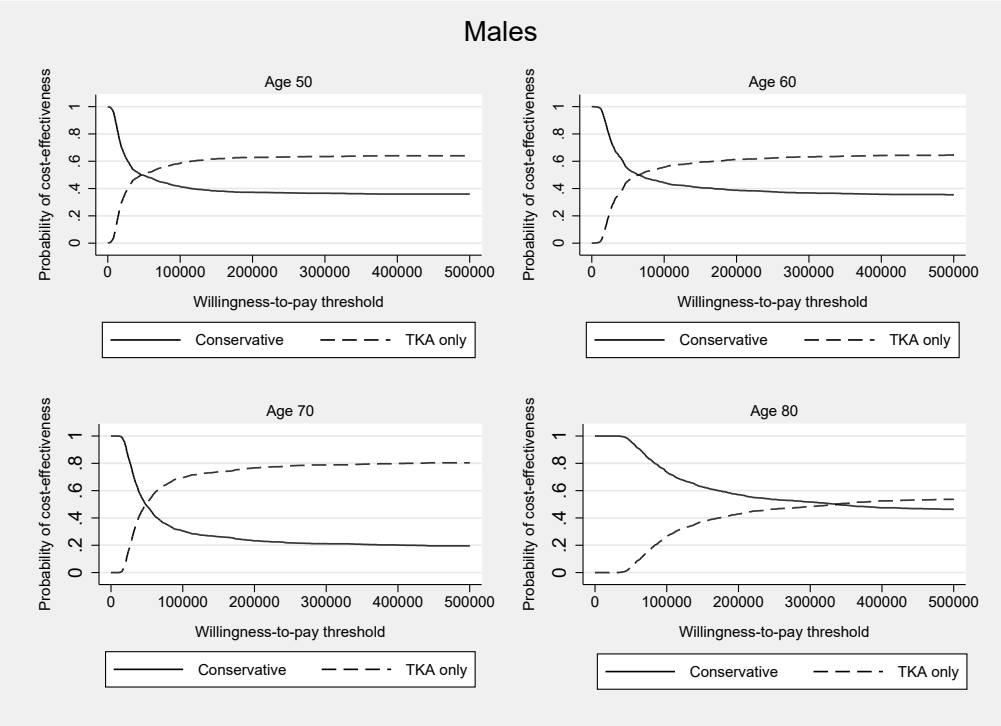


Figure 6.7: CEACs for male cohorts

Chapter 7

Value of information analysis

7.1 Expected value of perfect information

All decisions based on current information can potentially be wrong. From the CEACs, the standard decision rule would be to implement the strategy that has the highest expected net monetary benefit at the appropriate willingness-to-pay per QALY. The probability of making the wrong decision is given by the height of the other strategy's CEAC. For example in fig. 6.5 on page 51, the decision maker can decide to not implement the conservative programme at the threshold value of e.g. 300,000 NOK per QALY, where the TKA only strategy is more likely to be cost-effective for $\sim$ 60% of the simulations. In doing so, the decision maker accepts the probability of making the wrong decision of $\sim$ 40%. Additionally, if a differentiated decision based on subgroups of patients is an option, the uncertainty is even higher for the oldest groups of patients.

When uncertainty is high, an option is to conduct a value of information analysis, to see whether undertaking additional research to reduce uncertainty is worthwhile. New information that completely removes the uncertainty is termed perfect information [25]. While the decision based on current information is to select the strategy with,

$$max E_X[NMB(j, X)] \tag{7.1}$$

i.e. one of the j = 1,2 strategies with the largest expected net monetary benefit for

Table 7.1: Expected value of perfect infortion per patient, average cohort

WTP (NOK)	**Individual EVPI**
0	0
10 000	0
50 000	7 608
100 000	27 763
150 000	39 929
200 000	49 309
250 000	58 803
300 000	68 363
350 000	77 957
400 000	87 579
450 000	97 230
500 000	106 900

X, a vector of uncertain parameters in the model. Perfect information would be the same as having selected the strategy with the largest NMB for known values of X,

$$max_j[NMB(j, X)] \tag{7.2}$$

However, the true values of X will not be observed, but one can observe the largest NMB in each run of the probabilistic sensitivity analysis. By selecting the strategy with the largest NMB in each simulation , the exptation of those NMBs,

$$E_X max[NMB(j, X)] \tag{7.3}$$

can be interpreted as having perfect information in each case.

From (7.1) and (7.3) we express the upper boundary of the returns to additional research as the expected value of having perfect information,

$$\text{EVPI} = E_X max[NMB(j, X)] - max_j E_X[NMB(j, X)] \tag{7.4}$$

the difference between the NMB from perfect and current information.

Since at increasing willingness-to-pay thresholds, the probability of cost-effectiveness

of the strategies does not asymptote to either 0 or 1, the returns to new information does not taper-off after the intersect of the CEACs, as is the typical shape of EVPI-curves when the ICERs are in the north-west quadrant of the cost-effectiveness plane. What we see from table 7.1 on page 54 is rather that the EVPI increases continuously even after the intersect of the CEACs, which was at 117 000 NOK for the average cohort. Also in table 7.1, we observe that at very low willingness-to-pay thresholds, the individual EVPI is zero since new information is unlikely to change the decision. As the decision maker's threshold increases, the uncertainty and hence the value of reducing that uncertainty increase as well.

While the individual EVPI might not justify undertaking additional research, information of this kind has public good qualities since once produced, it has the potential to inform the decision for all current and future patients. One should therefore compare the EVPI for all current and future patients to the cost of reducing the uncertainty [25].

In estimating the current and future patient population, scope was limited to Norwegian patients only, and the expected life-time of the technologies' current form assumed to be 10 years. To obtain an estimate of the number of patients with moderate-to-severe osteoarthritis of the knee, the number of primary TKAs performed per year was taken as a proxy. Since 1994, the number of primary TKAs has risen by approximately 1 000 patients every five years [5]. In 2014, about 5 500 primary TKAs was performed ($\mathrm{I}_{t=0}$), so a linear prediction is that by 2025, 7 500 patients will need this surgery annually ($\mathrm{I}_{t=10}$). From this, the prediction is that population incidence rate, I_t rises with roughly 200 patients per year, and by discounting the number of future patients with the cost-discount rate of 4%, the 10-year expected future patient population is assumed to be,

$$\sum_{t=1}^{10} \frac{I_t}{(1.04)^t} \cong 55000 \tag{7.5}$$

The population EVPI is therefore given by multiplying the individual EVPI by the future use for the 55 000 current and future patients. From fig. 7.1 on the next page it is clear that, given the large current and future patient population that could benefit from additional research, this would very likely be worthwhile.

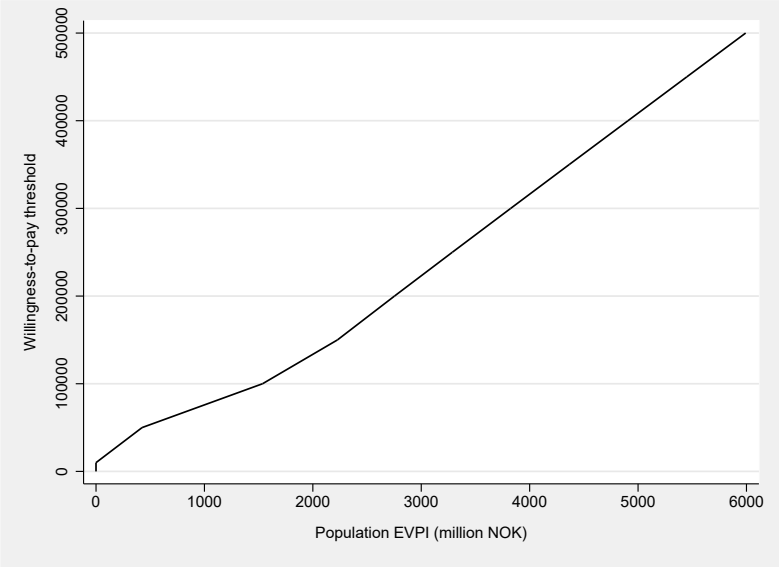


Figure 7.1: Population EVPI, million NOK

Thus, with this first hurdle passed, it should be established exactly what kind of research would enable realising the population EVPI.

7.2 Expected value of perfect information for parameters

To find out which parameters are likely to cause the uncertainty, we calculate the expected value of perfect information for parameters (EVPPI), which gives the monetary value of reducing the uncertainty around input-parameters, or parameter groups [25].

Obtaining the EVPPI practically involves running the probabilistic model in an inner and an outer loop. The inner loop consists of fixing a parameter's value at a random draw from it's probability distribution. With this parameter fixed the Monte Carlo simulation is run at the usual 1 000 iterations and obtain new net-monetary-benefits for the strategies, which is the outer loop. The inner loop was repeated 1 000 times for this model, effectively yielding 1 000 000 simulations of the model for each parameter. The EVPPI gives an insight as to which parameter has the best potential for reducing decision uncertainty, and where additional research

should be directed.

The common approach is to divide the parameters into groups that represent the type of research needed to resolve their uncertainty [25]. Here, the EVPPIs were run for the parameter groups; utilities, cost-components (i.e. the probabilites of additional costs of weight-loss advice and pharmacological treatment in the conservative programme), the direct treatment costs, failure hazards of the time-to-event parameters, and the assumed constant transition probabilities.

Given the discrepancy in the decision uncertainty for the oldest cohorts versus the others, the EVPPI estimations were all run twice. First, for the average cohort at a willingness-to-pay threshold near its most uncertain level, 100 000 NOK. Second, for the most uncertain cohort, males age 80 at baseline, with a willingness-to-pay threshold of 280 000 NOK.

A bar-chart with the results from the EVPPI for the average cohort can be found in fig. 7.2 on the following page. The single parameter group inducing the largest value from having perfect information about them is the utilities group. These parameters have a direct effect on the relative effectiveness of the two strategies. As seen from the scatter-plot in fig. 6.4 on page 50, since there is more variation in the east-west direction which is given by the denominator (the QALYs), than in the north-south direction, given by the relative costs, the result was not surprising. The EVPPI for the utilities group was 85 173 NOK at the individual level in the average cohort. This is the then the upper benefit achievable by reducing the uncertainty of this parameter group, valued in monetary terms. Scaled to its applicability for the entire current and future patient population, the value is roughly 4.7 billion NOK. For only the present patient population of 5 500 patients, the value is still 468 million NOK – arguably enough to justify resolving the uncertainty.

The other parameter groups have comparatively modest value of the perfect information about them. While they also amount to large sums at the current - and future population level, it would not be wise to allocate resource efforts from the utility parameters to resolve these instead, because they cannot possibly give the same returns on the investment.

In fig. 7.3 on page 59, the EVPPI for the same groups of parameters are seen when estimated for the most uncertain cohort, males age 80 at cycle 0. The result

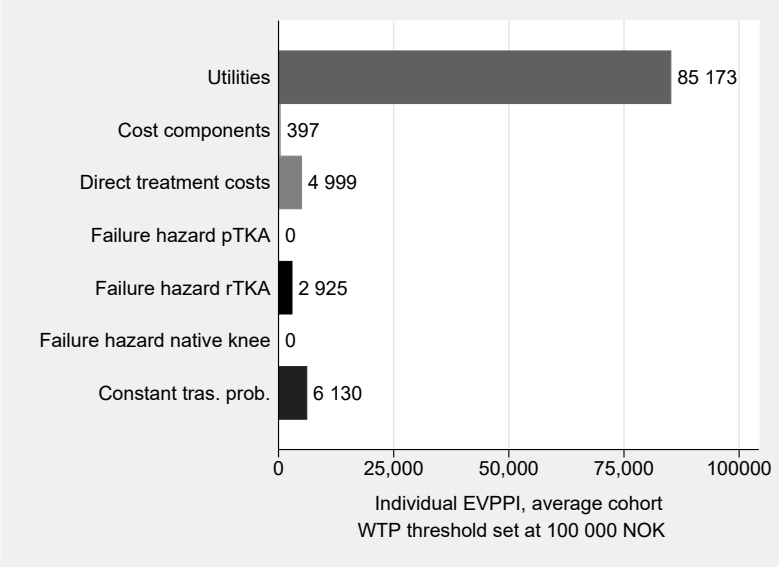


Figure 7.2: Individual EVPPI for average cohort

is informed by 7 x 1 000 inner and 1 000 outer loops of the probabilistic model. The results show that the only parameter group with a non-zero value is the utilities group. This is to be expected since at a higher threshold value, less emphasis will be given to cost fluctuations [1].

The big question left is, what would additional research on the utility values for these patients cost? This is difficult to assess. A fair assumption would be that the cost of such research is increasing in the level of precision (certainty) wanted, since a greater sample size would likely provide more efficient estimates. The following section shows how much more precise the estimates for the utility parameters would need to be to make the decision more robust.

7.3 Threshold analysis of new information

Having identified utilities as the most worthwhile parameter group for additional research we can look at the effects of more certainty around those estimates on the uncertainty of the decision. That is; while perfect information as a theoretical

[1] To see why, consider again the willingness-to-pay line, λ, running through a scatter-plot, like in fig. 6.4 on page 50. The more the line rotates westward, the more critical the relative QALYs (the product of time and utilities) become.

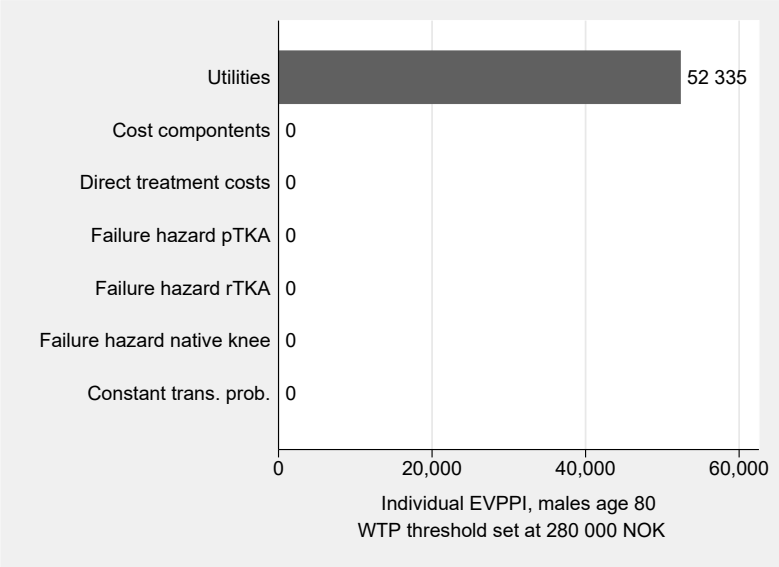


Figure 7.3: Individual EVPPI for male age 80 cohort

concept is fine, it is likely sufficient with somewhat less than perfect information to make the overall decision more robust. To review the potential of this, the standard deviations of the utility parameters were reduced to one half, one quarter, and one eight of current information. In the upper left panel of fig. 7.4 on page 61 one can see 1 000 random draws from the probability distribution of the baseline utility parameter. This is the current information.

In fig. 7.5 on page 61 one can see that halving the standard deviations of all utility parameters makes the decision somewhat more certain, however not radically different to the current uncertainty. With standard deviations equal to a quarter of the current information for the utility parameters, the decision becomes a lot more certain. Finally, with only an eight of the current standard deviation, the uncertainty and hence probability of error becomes more or less eliminated.

Large standard deviations often arise from the fact that the sample size is small. However, this need not be the only explanation. For example, the baseline utility estimate from the Karolinska Institue study had a decent sample size of 365 individuals, so it was seemingly not underpowered in terms of size. There is a possibility that the variation in the descriptions of health-related quality-of-life could be attributable to systematic differences in patient characteristics such

as the age of the patient. If there were systematic differences one would expect patients belonging to the same subgroup to be more in agreement over the mean than by ignoring the differences and combining all patients. If this is the case one could possibly reduce the standard deviation by controlling for such differences and express the estimates for the subgroups of patients more precisely. For a model such as this, which has characteristic-specific cohorts, it is a fair assumption that this could give more relevant estimates. To know whether there are such systematic difference we would need to obtain the patient level data from a previous study such as the one used here, and perform regression analysis to see the influence of possible characteristics of systematic difference. If access to the data is not possible, a likely more costly alternative would be to collect new data.

As for the utility parameter informing the experience in the conservative programme, the sample size is in fact small. One option to improve on this could be to use the information coming out of the ongoing Norwegian study [3]. In doing so, the model should also be adapted to use the outcomes from the study in terms of survival of the native knee.

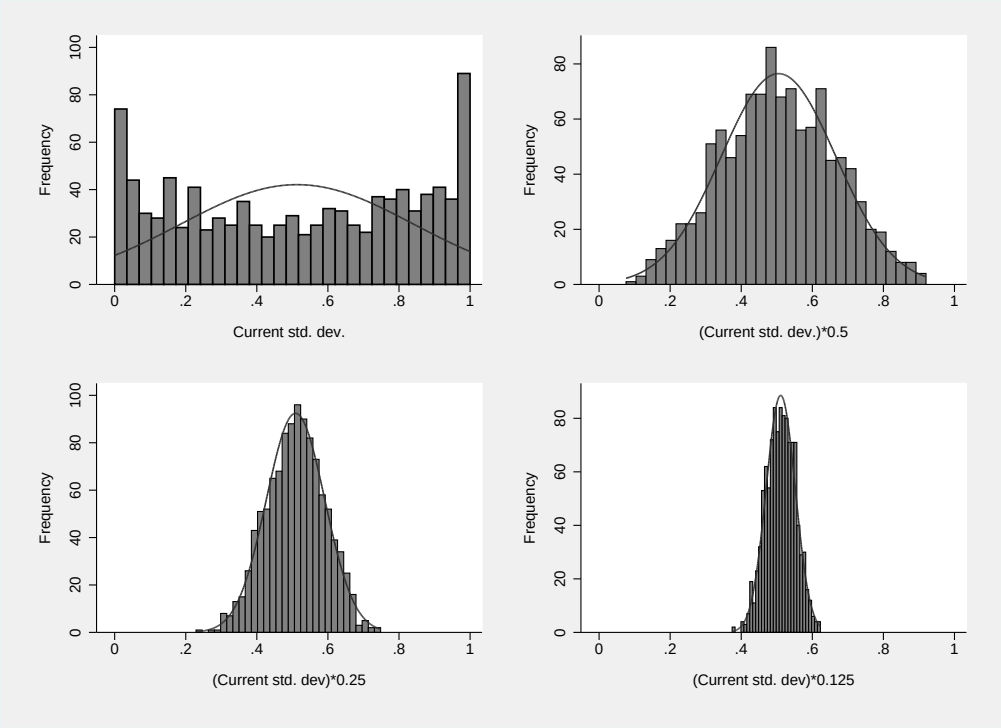


Figure 7.4: Reducing the standard deviation of the baseline utility parameter

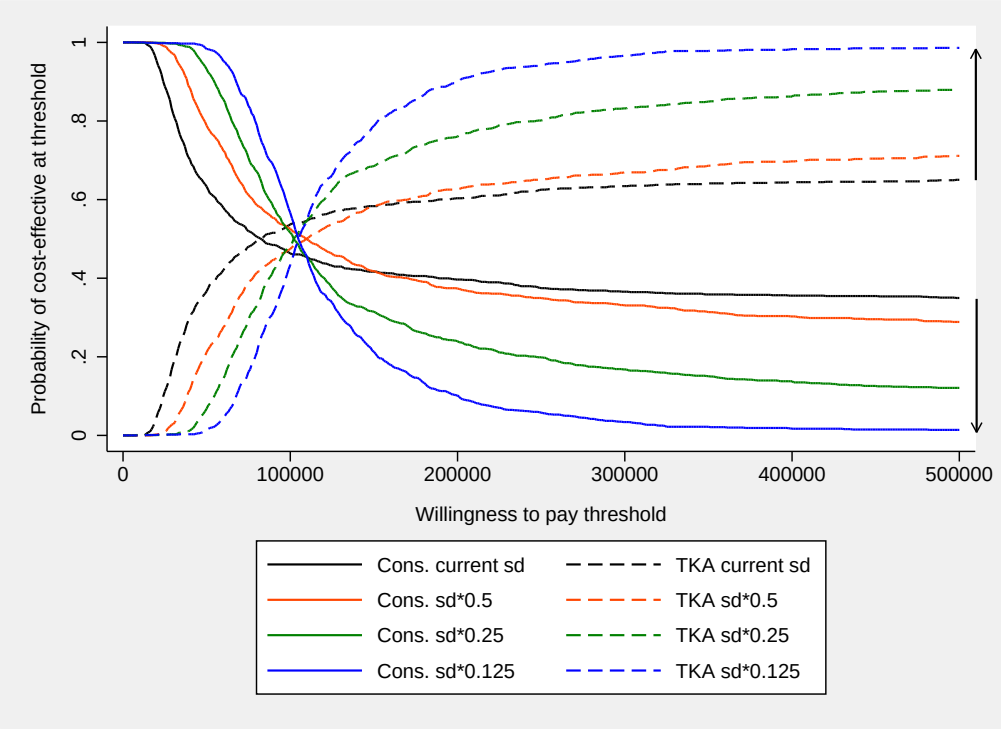


Figure 7.5: Effects of more efficient utility estimates of the CEACs for the average cohort

Chapter 8

Structural uncertainty analysis

8.1 If all transition probabilities were constant

This section concerns the choices made for the model's structure during development. This is to show how alternative assumptions would have affected the deterministic model outputs, incremental costs and QALYs.

All constant transition probabilities was interpreted as a too strong assumption during development. One point for checking the impact of this assumption is that if the analysis is to be recreated by another researcher in the future (if e.g. more/better data becomes available), less time can spent recreating it by assuming constant transition probabilities, and hence the cost of recreating it would be lower.

By using the results from the hazard functions under the assumption of exponential distributions, all time-dependent transition probabilities were made constant. The box-plots in fig. 8.1 on page 65 shows the incremental costs and QALYs under the assumptions of either time-dependent, or constant transition probabilities. For both costs and QALYs this did not have a large effect on the between-cohort variation, seeing as the median is more or less at the same position. The difference in incremental costs of the conservative programme under the two assumptions is that it becomes more cost-saving under the assumption of constant transition probabilities. This shift is not seen for the incremental QALYs, which is almost identical under the two assumptions. This means that since incremental costs became comparatively lower than incremental QALYs, assuming constant

transition probabilities is likely to make the non-surgical programme be favoured more strongly at low willingness-to-pay thresholds for all cohorts. However, since it appears still less effective, it will likely not have any impact on the decision for a decision maker with a moderate-to-high willingness-to-pay threshold.

8.2 If all cycles were given equal weight

Many of the most expensive effects of the two alternatives appear many years after the baseline cycle, for example needing a primary TKA for the non-surgical strategy. By discounting future cycles, these events carry less weight in the life-time costs and QALYs. Even though the national guidelines require discounting, it could be fruitful to see whether the different timing of events has implications, and especially so for the different cohorts. Individuals in the youngest cohorts could possibly be more likely to for example experience a revision arthroplasty (calibrated from expert opinion to have a lower probability of full benefit, and lower utility), due to their longer life-expectancy. To check this, the discount rate was set to 0%.

The results for the model deterministic outputs can be seen in the box-plots in fig. 8.2 on page 66. Giving equal weight to all cycles for each cohort did not have a great effect on the life-time QALYs gained for either of the cohorts. Equal weighting of all future costs had the implication that for the two youngest female cohorts, and the youngest male cohort, the conservative programme strategy actually had higher life-time costs than the TKA directly-strategy, which is why the top-whisker of the box-plot crosses the reference line. The implication for the decision is that the conservative programme would be a dominated strategy for these cohorts.

This underlines that the much of the cost-saving qualities of the conservative programme lies in its ability to delay the high-cost events. The implication for the structural uncertainty of the model is that since the “survival of the native knee”-parameter is absolutely crucial for the cost-advantage of the conservative programme, future replications of this model should emphasise obtaining new, updated information on it.

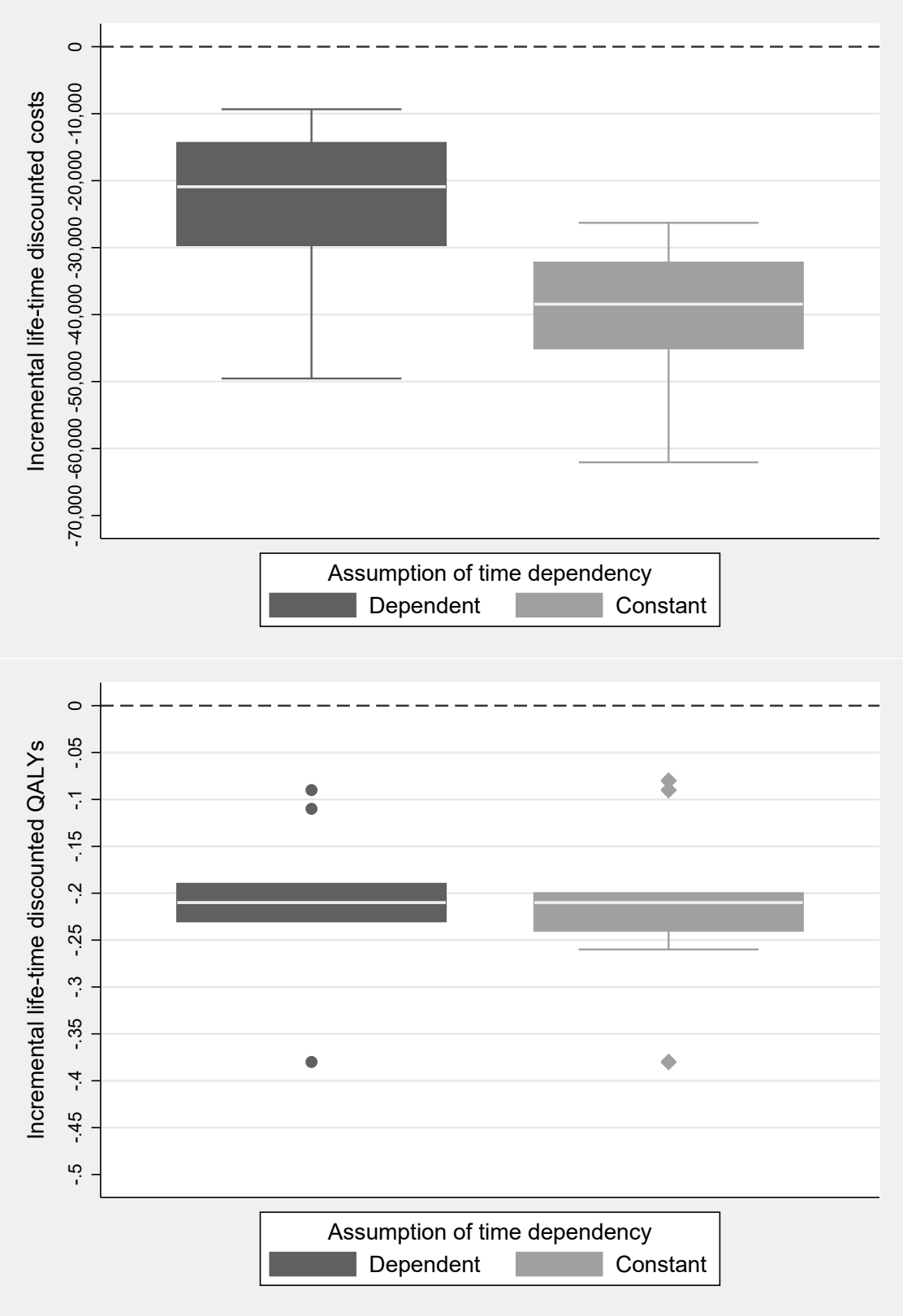


Figure 8.1: Model deterministic outputs for all cohorts under assumptions of time-dependency

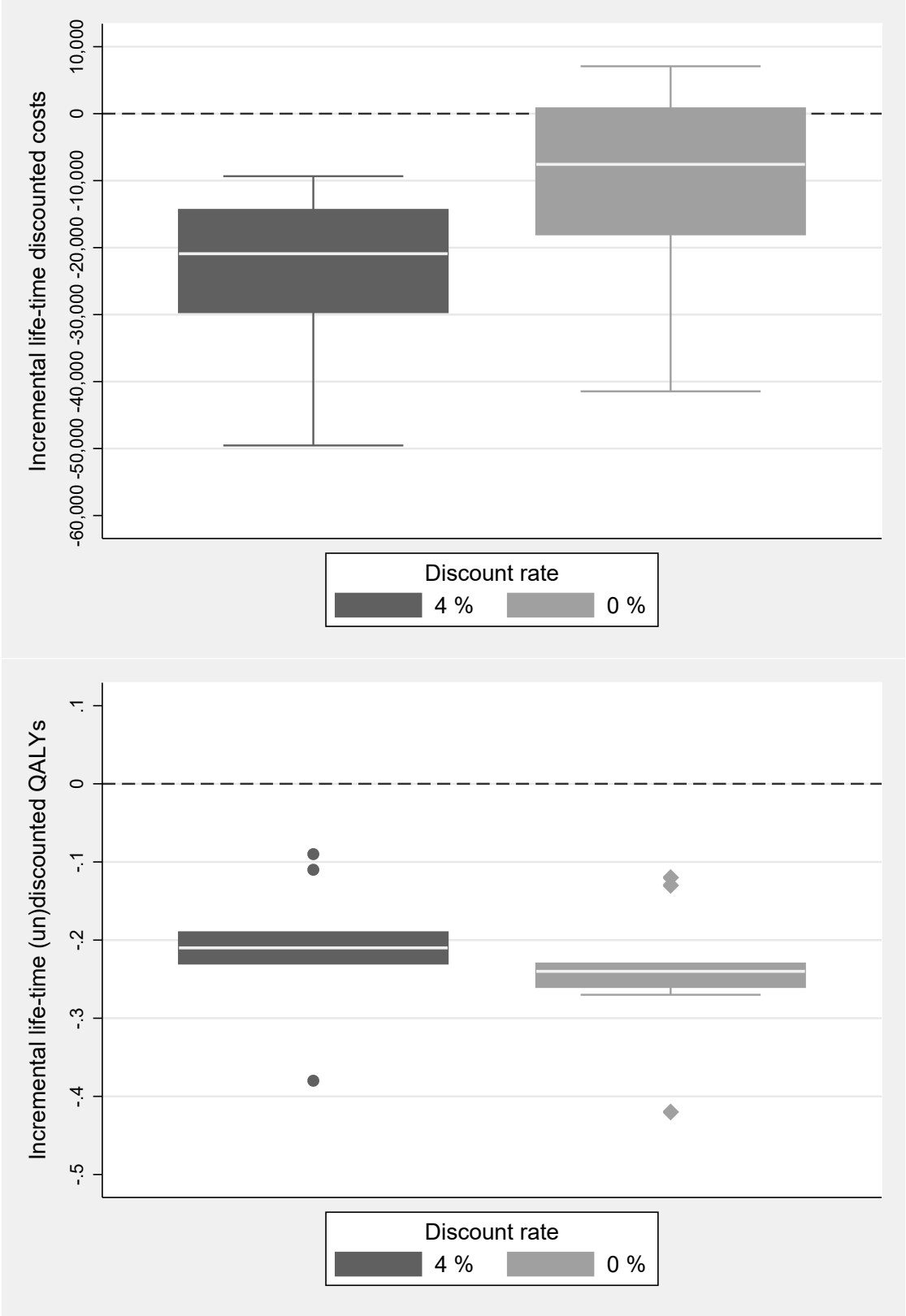


Figure 8.2: Model deterministic output for all cohorts under different discount rates

Chapter 9

Discussion

9.1 Objectives

This analysis had two objectives. First, to establish to what degree a non-surgical treatment programme could be used as conservative treatment for patients who are candidates for arthroplastic surgery for their osteoarthritis. The second objective was to assess whether a such strategy would be a cost-effective alternative to directly undergoing total knee arthroplasty.

An updated search for similar studies was performed in ORIA on the 5th of May, yielding no relevant analyses. The comparison of non-surgical treatment to TKA as in this study therefore appears to be rare. A possible explanation might be that for this patient group, the total knee arthroplasty procedure is generally accepted as the suitable clinical choice, hence deviations from this is also rare. This thesis looked at giving the appropriate non-surgical treatment in an organised programme and in the event of exhausted benefits allow TKA, and whether this is a superior strategy to have the patients directly undergo TKA. This is not supposed to be interpreted as *rationing* the TKA-procedure through the use of the non-surgical programme. The reason for this is that Norwegian standard practice already has this constraint. Rather, the thesis seek to show the cost-effectiveness of appropriate guideline-complacent non-surgical treatment, implicitly showing the *potential* that the existing non-surgical treatment has, would it be put together of interventions that osteoarthritis experts advice, and in a manner that would

promote adherence. The analysis was conducted under the assumption, supported by research, that existing non-surgical treatment is not optimal at present the time.

Non-surgical treatment for these patients was established to be possible to conserve their native knee. The relative cost-effectiveness compared to the TKA-only strategy depends on two things. First, the willingness-to-pay for a quality-adjusted life-year. Second, whether the decision is to be made for an average patient group, or whether subgroups of patients should be treated differently.

9.2 Main findings

The prediction beyond the available evidence is that patients can maintain their native knee for some years, but after about ten years, most will have either experienced a failure of the native knee and undergone TKA, or have died. For this reason, high-cost events are postponed such that the conservative treatment appears to be a cost-saving alternative for all cohorts. The conservative treatment strategy also give fewer expected life-time quality-adjusted life-years for all cohorts. This causes the relative cost-effectiveness to be a function of the decision makers' willingness-to-pay threshold. At low thresholds ($<$ 117 000 NOK), the conservative treatment has the highest probability of being the cost-effective choice for the average patient. Above this threshold, TKA directly is always the cost-effective strategy. If the decision-maker is willing to make subgroup-specific decisions, the relative cost-effectiveness for the oldest cohorts is markedly different from the average. A willingness-to-pay threshold of $>$ 240 000 NOK for the cohort of females age 80, and $>$ 337 000 NOK for the cohort of males age 80, is needed to dismiss the conservative strategy as the most probable to be cost-effective. As shown in fig. 6.2 on page 47 and in Appendix D, the oldest cohorts spend the longest time relative to their time alive in the state conserving their native knee. The reason then for the high uncertainty around the decision for these groups can be traced back to the analysis' lack of cohort-specific inputs other than failure hazards and mortality risk. For given utilities, costs, and transition probabilities from conservative to primary TKA, the mortality of the cohorts take centre stage. Because the oldest female and male cohorts do not live long enough to accumulate

many years in the full benefit state, the relative QALY advantage of the TKA-only strategy is decreased. The cost-advantage of the conservative strategy for these cohorts are affected likewise; most of the cohort die before needing the surgery. Thus, the uncertainty for these cohorts are driven by the limited information about them in terms of utilities, costs, and failure of the native knee.

9.3 Strengths

The data sources employed in this thesis all have a Scandinavian origin. This geographical closeness suggests that the model's inputs are relevant for the target patient population. It could also be considered a strength that the analysis not only covers an average cohort, but also attempts to make a wider scope by include more patient characteristics that can affect the decision. This analysis has a robust investigation of uncertainty of the model, both for parameters and the structure. Further, actual suggestions on how to reduce the uncertainty is provided. With the analysis, comprehensive technical appendices are provided to reduce any confusion about how the results were obtained. The intention was to be transparent enough that the analysis could be fully re-creatable from the information given by another researcher.

9.4 Limitations

The research question is arguably narrow. Ideally one would have investigated the effects of conservative treatment for patients not yet progressed to a state where arthroplasty is an option. Attempts were initially made to open up the analysis for this at an early stage, however, due to the limited time available, the decision was made to abandon this.

This thesis has limitations in terms of the availability of data. The method of survival analysis without information on censoring cannot give more than approximated results.

More critical is the lack of detailed information on health state utilities for the modelled cohorts. This is as mentioned reflected in the larger uncertainty around

the decision for cohorts more influenced by the mortality parameters. There is also a possible limitation regarding the international transferability of the utility estimates used in this analysis. The baseline, and post primary TKA utility estimates are based on description by Swedish patients adjusted to reflect the health state preferences of a British population subjected to the time-trade off (TTO) method of valuation [45]. The one-year follow-up utility estimate for the patients in the Danish RCT was derived from the patients and adjusted using an EQ-5D tariff reflecting the health state preferences by TTO of a Danish population [46]. While the Danish and British tariffs have been shown to be highly correlated, direct comparisons of health states revealed some differences; using the same EQ-5D vector values, the Danish tariff gave values worse than death (utility < 0) for 22% of the health states, and the UK tariff 34% [46]. This would effectively lead to a higher mean value using the Danish tariff. The implication for the present analysis is that if the UK tariff been used in the RCT, the gap between the utility experienced in the full benefit state and the state for maintaining the native knee could have been larger. This would then lead to lower life-time QALYs for the conservative strategy relative to the TKA-directly strategy.

The analysis' cost perspective is limited. Only considering costs to the health care service from the supply of treatment is ignoring the possible impact of costs from rehabilitation, especially so for the oldest cohorts, and the potential offsetting effect of production gains from improvement for the working age cohorts. It should be noted that inquiries were made to the Directorate of Health with regards to the availability of data on rehabilitation for knee osteoarthritis patients. The government does not have data on the resource use for rehabilitation for these patients, and estimation is complicated due to the fact that there are many fragmented providers in both the public and private sector. On the other hand, given more time, scenario analyses could have been provided alternative estimates.

Pharmacological interventions could have been included on its own as a relevant comparator instead as just a cost-component to the non-surgical programme, to allow for the possibility that some might not want neither arthroplastic surgery nor to participate in a conservative programme. This might have given a more complete picture, although given that pharmacological interventions are known only to reduce pain and not provide any improvement in the knee, the additional

knowledge gained from including it as a comparator and its relevance for the research question is uncertain.

Finally, the validity of the assumption that patients in the limited state permanently stay on the pharmacological regimen based on that in the RCT is somewhat uncertain. Considering that long-term use of the modelled pharmaceuticals is not recommended, more information on the treatment in this state would have benefited the analysis.

9.5 Further research

As already discussed, there is a large potential gain to the current and future patient population by reducing the uncertainty of the utility-parameters. The analysis should be repeated if better information on these become available. Furthermore, when the Norwegian cohort study [3] with the similar programme gets published one will have the opportunity to calibrate the failure hazard of the native knee to reflect external validity of the results, and to employ measures of quality of life from the target population.

Chapter 10

Conclusion

Compared to undergoing total knee arthroplasty directly, a non-surgical conservative treatment programme followed by a TKA in the event of a failed native knee is a cost-saving, less effective treatment alternative for patients with severe osteoarthritis of the knee. There is, however, considerable uncertainty, and depending on the decision-makers willingness-to-pay per quality-adjusted life-year, a non-surgical conservative treatment programme could be considered a cost-effective alternative at low willingness to pay thresholds.

The results for the modelled cohorts indicate differences in age attributable to the relative expected life-length of the cohorts, however, due to the lack of cohort-specific information beyond this it suggests that the average cohort results are more informative at the present time.

References

[1] Margreth Grotle et al. "Prevalence and burden of osteoarthritis: results from a population survey in Norway." In: *The Journal of rheumatology* 35.4 (2008), pp. 677–684.

[2] Folkehelseinstituttet. *Folkehelserapporten 2014. Helsetilstanden i Norge*. 2015. URL: http://www.fhi.no/artikler/?id=113662.

[3] Nina Østerås et al. "Implementing international osteoarthritis treatment guidelines in primary health care: study protocol for the SAMBA stepped wedge cluster randomized controlled trial". In: *Implementation Science* 10.1 (2015), pp. 1–12.

[4] Nina Østerås et al. "Patient-Reported Quality of Care for Osteoarthritis: Development and Testing of the OsteoArthritis Quality Indicator Questionnaire". In: *Arthritis care & research* 65.7 (2013), pp. 1043–1051.

[5] Ove Furnes et al. *Norwegian Arthroplasty Register Report 2015 (Norwegian)*. 2015, pp. 59–101.

[6] Søren T Skou et al. "A randomized, controlled trial of total knee replacement". In: *New England Journal of Medicine* 373.17 (2015), pp. 1597–1606.

[7] Muralitharan Nair. *Fundamentals of applied pathophysiology: an essential guide for nursing students*. John Wiley & Sons, 2009, pp. 476–484.

[8] David T Felson. "Osteoarthritis of the knee". In: *New England Journal of Medicine* 354.8 (2006), pp. 841–848.

[9] Corrinda Black et al. “The clinical effectiveness of glucosamine and chondroitin supplements in slowing or arresting progression of osteoarthritis of the knee: a systematic review and economic evaluation”. In: *Health Technology Assessment* 13.52 (2009), pp. 3–16.

[10] KM Leyland et al. “The natural history of radiographic knee osteoarthritis: A fourteen-year population-based cohort study”. In: *Arthritis & Rheumatism* 64.7 (2012), pp. 2243–2251.

[11] Carina A Thorstensson et al. “Natural course of knee osteoarthritis in middle-aged subjects with knee pain: 12-year follow-up using clinical and radiographic criteria”. In: *Annals of the rheumatic diseases* 68.12 (2009), pp. 1890–1893.

[12] V Silverwood et al. “Current evidence on risk factors for knee osteoarthritis in older adults: a systematic review and meta-analysis”. In: *Osteoarthritis and Cartilage* 23.4 (2015), pp. 507–515.

[13] JH Kellegren and JS Lawrence. “Radiological assessment of osteoarthritis”. In: *Ann Rheum Dis* 16 (1957), pp. 494–501.

[14] R Altman et al. “Development of criteria for the classification and reporting of osteoarthritis: classification of osteoarthritis of the knee”. In: *Arthritis & Rheumatism* 29.8 (1986), pp. 1039–1049.

[15] Paul Dieppe and Kenneth D Brandt. “What is important in treating osteoarthritis? Whom should we treat and how should we treat them?” In: *Rheumatic Disease Clinics of North America* 29.4 (2003), pp. 687–716.

[16] Norwegian Electronic Medical Manual. *Kneleddsartose (Norwegian)*. 2001. URL: http://legehandboka.no/handboken/kliniske-kapitler/ortopedi/tilstander-og-sykdommer/kne/kneleddsartrose/.

[17] Helsedirektoratet. *Retningslinjer kneleddsartrose*. 2015. URL: https://helsedirektoratet.no/retningslinjer/ortopedi/seksjon#Kneleddsartrose%20-%20alvorlige%20plager%20%28veiledende%20frist%2012%20uker%29.

[18] J Bruce Moseley et al. “A controlled trial of arthroscopic surgery for osteoarthritis of the knee”. In: *New England Journal of Medicine* 347.2 (2002), pp. 81–88.

[19] Alexandra Kirkley et al. "A randomized trial of arthroscopic surgery for osteoarthritis of the knee". In: *New England Journal of Medicine* 359.11 (2008), pp. 1097–1107.

[20] TE McAlindon et al. "OARSI guidelines for the non-surgical management of knee osteoarthritis". In: *Osteoarthritis and Cartilage* 22.3 (2014), pp. 363–388.

[21] Gudmund Grønhaug, Nina Østerås, and Kåre Birger Hagen. "Quality of hip and knee osteoarthritis management in primary health care in a Norwegian county: a cross-sectional survey". In: *BMC health services research* 14.1 (2014), p. 1.

[22] Paul Jarle Mork, Andreas Holtermann, and Tom Ivar Lund Nilsen. "Effect of body mass index and physical exercise on risk of knee and hip osteoarthritis: longitudinal data from the Norwegian HUNT Study". In: *Journal of epidemiology and community health* 66.8 (2012), pp. 678–683.

[23] Michael F Drummond et al. *Methods for the economic evaluation of health care programmes*, pp. 9–25, 33–43, 85, 277–309.

[24] John Brazier. *Measuring and valuing health benefits for economic evaluation*, pp. 55–65, 99–118.

[25] Andrew Briggs, Mark Sculpher, and Karl Claxton. *Decision modelling for health economic evaluation*. OUP Oxford, 2006, pp. 15–37, 77–100, 122–123, 166–194.

[26] Petter Laake, Stian Lydersen, and Marit Bragelien Veierød. *Medical statistics in clinical and epidemiological research*. Gyldendal akademisk, 2012, pp. 167–198.

[27] Zhihui Liu, Benjamin Rich, and James A Hanley. "Recovering the raw data behind a non-parametric survival curve". In: *Systematic reviews* 3.1 (2014), p. 1.

[28] *Digitizer software – digitize a scanned graph or chart into (x,y)-data.* 2015. URL: http://www.digitizeit.de/.

[29] MG Myriam Hunink et al. *Decision making in health and medicine: integrating evidence and values*. Cambridge University Press, 2014.

[30] Martin Sundberg et al. *Swedish Knee Arthroplasty Register annual report 2015 (Swedish)*. 2015, pp. 60–65.

[31] Tesfaye H Leta et al. "Failure of aseptic revision total knee arthroplasties: 145 revision failures from the Norwegian Arthroplasty Register, 1994–2011". In: *Acta orthopaedica* 86.1 (2015), pp. 48–57.

[32] Stein Atle Lie et al. "Duration of the increase in early postoperative mortality after elective hip and knee replacement". In: *J Bone Joint Surg Am* 92.1 (2010), pp. 58–63.

[33] Howard M Taylor and Samuel Karlin. *An introduction to stochastic modeling*. Academic press, 2014, pp. 79–80.

[34] Uwe Siebert et al. "State-transition modeling: a report of the ISPOR-SMDM modeling good research practices task force-3". In: *Value in Health* 15.6 (2012), pp. 812–820.

[35] Elamin H Elbasha and Jagpreet Chhatwal. "Myths and Misconceptions of Within-Cycle Correction: A Guide for Modelers and Decision Makers". In: *PharmacoEconomics* 34.1 (2016), pp. 13–22.

[36] Helsedirektoratet. *Økonomisk evaluering av helsetiltak - en veileder*. 2012. URL: https://helsedirektoratet.no/retningslinjer/veileder-i-okonomisk-evaluering-av-helsetiltak.

[37] Statistics Norway. *Life tables, 2014*. 2015. URL: http://www.ssb.no/befolkning/statistikker/dode.

[38] Leonard Shan et al. "Intermediate and Long-Term Quality of Life After Total Knee Replacement". In: *J Bone Joint Surg Am* 97.2 (2015), pp. 156–168.

[39] Olivier Ethgen et al. "Health-related quality of life in total hip and total knee arthroplasty". In: *J Bone Joint Surg Am* 86.5 (2004), pp. 963–974.

[40] James Slover et al. "Cost-effectiveness of unicompartmental and total knee arthroplasty in elderly low-demand patients". In: *The Journal of Bone & Joint Surgery* 88.11 (2006), pp. 2348–2355.

[41] Elena Losina et al. "Cost-effectiveness of total knee arthroplasty in the United States: patient risk and hospital volume". In: *Archives of internal medicine* 169.12 (2009), p. 1113.

[42] Karl-Åke Jansson and Fredrik Granath. "Health-related quality of life (EQ-5D) before and after orthopedic surgery". In: *Acta orthopaedica* 82.1 (2011), pp. 82–89.

[43] *Innsatsstyrt finansiering 2015*. 2015. URL: https://helsedirektoratet.no/Lists/Publikasjoner/Attachments/775/Innsatsstyrt-finansiering-2015-regelverk-IS-2230.pdf.

[44] Andrew H Briggs et al. "Model parameter estimation and uncertainty: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6". In: *Value in Health* 15.6 (2012), pp. 835–842.

[45] Paul Dolan et al. "The time trade-off method: Results from a general population study". In: *Health economics* 5.2 (1996), pp. 141–154.

[46] Kim U Wittrup-Jensen, Jørgen Lauridsen, and Kjeld M Pedersen. *Modelling Danish EuroQol (EQ-5D) tariffs by applying the time trade-off method*. Syddansk Universitet, 2008.

[47] Patricia Guyot et al. "Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves". In: *BMC medical research methodology* 12.1 (2012), p. 9.

[48] David G Kleinbaum and Mitchel Klein. *Survival analysis: a self-learning text*. Springer Science & Business Media, 2006.

[49] Legemiddelindustrien. *Felleskatalogen*. 2015. URL: http://www.felleskatalogen.no/medisin/.

[50] Norwegian Health Economics Administration (HELFO). *Regelverk og takster for fysioterapeuter*. 2015. URL: https://helfo.no/takster/regelverk-og-takster-for-fysioterapeut.

[51] Kliniske ernæringsfysiologers forening tilsluttet forskerforbundet. *Lønnsstatistikk 2013*. 2013. URL: http://www.keff.no/no/Nyheter/Nyheter_fra_2014/L%C3%B8nnsstatistikk+for+2013.9UFRLI2n.ips.

[52] Den norske legeforening. *Fastlegetariff 2014*. 2014. URL: http://normaltariffen.legeforeningen.no/pdf/Fastlegetariff_2014.pdf.

Appendices

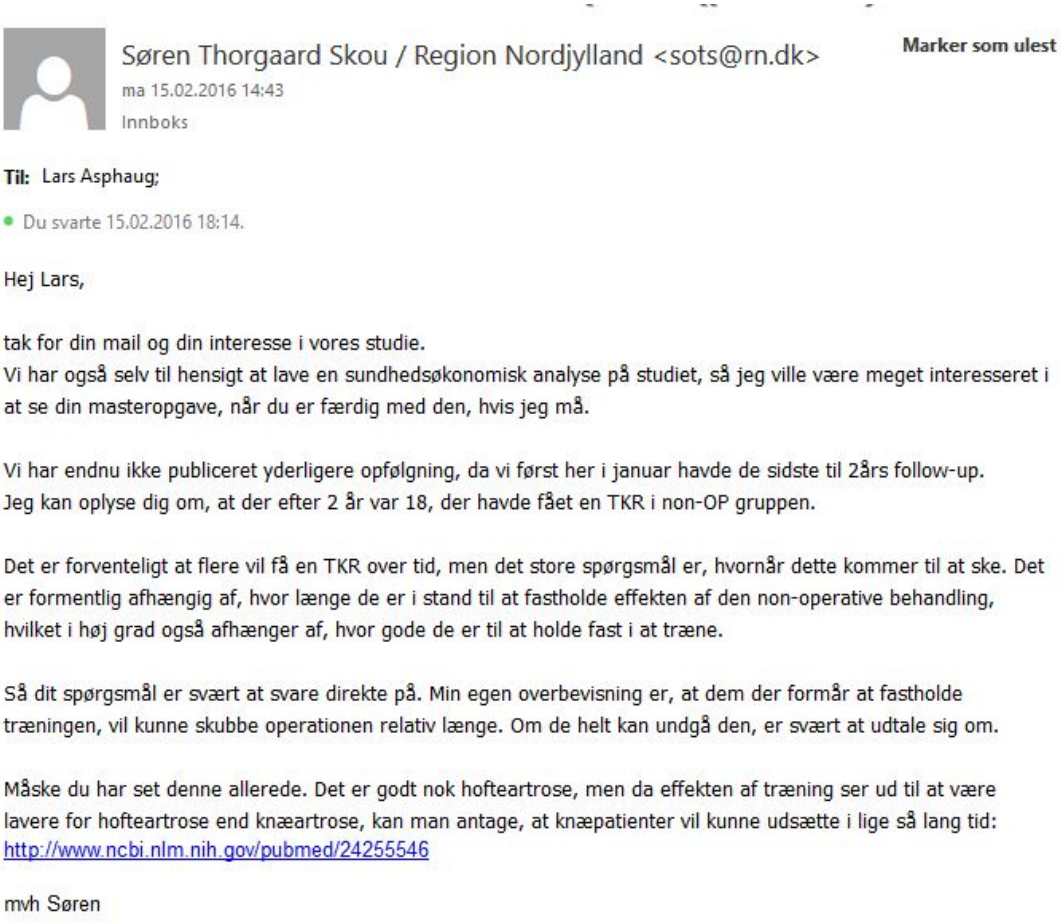
Appendix A

Survival data

Survival analysis was performed to model time to event for three parameters: failure hazard of the conservative treatment programme, revision hazard of primary TKA, and revision hazard of revised TKA. For the failure of the conservative programme, it was not possible to obtain characteristics of the patients who exited the programme and underwent primary TKA (failure). Since the follow-up was only 12 months, any survival analysis would have not been able. Yet unpublished numbers of failure was obtained by contacting the main author of the study see figure fig. A.1 on the next page. The survival of primary TKAs in Norway was taken from the 2015 annual report of the Norwegian Arthroplasty Register [5]. The only patient characteristics available was stratification by age, below and over the age of 60. The survival of revised TKAs in Norway was taken from Leta et al. [31] where patient characteristics was available for differences in gender, and age by < 60, 60-70, > 70.

Algorithm for obtaining the numbers at risk and number of failures

This section relies upon the work of Guyot et al. [47]. The amount of individuals in the data that have not yet experienced a failure event are referred to as numbers at risk. Ideally one would know the number at risk and the number of failure events at each time interval that is sensible in terms of the digitized coordinates of the published survival curves. However, as with the data used for this study, this is not



Søren Thorgaard Skou / Region Nordjylland <sots@rn.dk> **Marker som ulest**
ma 15.02.2016 14:43
Innboks

Til: Lars Asphaug;

• Du svarte 15.02.2016 18:14.

Hej Lars,

tak for din mail og din interesse i vores studie.
Vi har også selv til hensigt at lave en sundhedsøkonomisk analyse på studiet, så jeg ville være meget interesseret i at se din masteropgave, når du er færdig med den, hvis jeg må.

Vi har endnu ikke publiceret yderligere opfølgning, da vi først her i januar havde de sidste til 2års follow-up.
Jeg kan oplyse dig om, at der efter 2 år var 18, der havde fået en TKR i non-OP gruppen.

Det er forventeligt at flere vil få en TKR over tid, men det store spørgsmål er, hvornår dette kommer til at ske. Det er formentlig afhængig af, hvor længe de er i stand til at fastholde effekten af den non-operative behandling, hvilket i høj grad også afhænger af, hvor gode de er til at holde fast i at træne.

Så dit spørgsmål er svært at svare direkte på. Min egen overbevisning er, at dem der formår at fastholde træningen, vil kunne skubbe operationen relativ længe. Om de helt kan undgå den, er svært at udtale sig om.

Måske du har set denne allerede. Det er godt nok hofteartrose, men da effekten af træning ser ud til at være lavere for hofteartrose end knæartrose, kan man antage, at knæpatienter vil kunne udsætte i lige så lang tid:
http://www.ncbi.nlm.nih.gov/pubmed/24255546

mvh Søren

Figure A.1: Correspondance with the main author of the study by Skou et al.[6]

the case – the numbers at risk were only inferable at time=0, and no information of number of failure events were available. In this case one have to make the strong assumption that there are no censored observations[47]. A censored observation is a case where an individual disappears from the number at risk for any reason. Death is a very likely type of censoring for studies on old patients spanning many years. In other words, when the information is not available we have to assume the only reason for leaving the study is a failure event.

The algorithm for obtaining the relevant numbers for survival analysis in the absence of information of actual numbers at risk with each time interval and failures each time interval was loosely based on the algorithm presented in Guyot et al., however it is to my knowledge, an original feature of this thesis.

The algorithm functions as follows:

1. Initially the column vectors of time (t) and the observed survival, S(t) are populated by the digitized values, and the initial number at risk $N(t_0)$. We have the observed survival running from t_0 to t_k, with the number at risk only empirically given at t_0. We now wish to use this to calculate the number at risk at t_1, t_2,...,t_k, and the number of failures at t_1, t_2,...,t_k.
2. Moving from t_{k-1} to t_k we have that $\Delta S(t_k)=[S(t_{k-1})-S(t_k)])/S(t_{k-1})$. We use this to calculate $D(t_k)$, the number of failures that have taken place from t_{k-1} to t_k. This is possible by multiplying with the number of individuals at risk at t_{k-1}, hence $D(t_k) = \Delta S(t_k)*N(t_{k-1})$.
3. The final piece of the puzzle is the number at risk at t_k. This is simply the number at risk the period before, less the number of failures this period. This is intuitive because to be at risk, there must be uncertainty around whether the failure will happen, thus it needs to be the final part of the algorithm. $N(t_k) = N(t_{k-1}) - D(t_k)$.

After obtaining this information, a second panel data set with individual observations was created. Each individual "observation" in the study was determined by three parameters:

1. Each individual were given a "spell" , a value indicating what time step the observation accounted for.

2. An individual specific id – an integer number to keep track of how long an individual was at risk
3. An event value (binary) where a value 0 indicated still at risk, and 1 indicated event.

When an individual had an event, it was "out" of the panel. These preparations of the data enabled using Stata to recreate the Kaplan-Meier curves and fit parametric models.

Table A.1: Algorithm for rumbers at risk and failures of the native knee in the conservative programme

Time	**Survival, S(t)**	**ΔS(t)**	**Numbers at risk, N (t)**	**Number of failures, D(t)**
0	1.00		49	0
1	0.735	0.265	36	13
2	0.633	0.139	31	5

Table A.2: Numbers at risk and failures for primary TKA, age ≤ 60

Time	**Survival, S(t)**	**ΔS(t)**	**Numbers at risk, N (t)**	**Number of failures, D(t)**
0	1.000		1000*	0
1	0.979	0.021	979	21
2	0.956	0.023	956	23
3	0.937	0.020	937	19
4	0.925	0.013	925	12
5	0.916	0.010	916	9
6	0.910	0.007	910	6
7	0.905	0.005	905	5
8	0.900	0.006	900	5
9	0.892	0.009	892	8
10	0.886	0.007	886	6
11	0.879	0.008	879	8
12	0.873	0.006	873	5
13	0.871	0.003	871	2
14	0.859	0.013	859	12
15	0.856	0.003	856	3
16	0.845	0.013	845	11
17	0.835	0.012	835	10
18	0.829	0.007	829	6
19	0.811	0.022	811	18
20	0.785	0.032	785	26

* Inferred from number of primary TKAs inserted 1994 [5]

Table A.3: Numbers at risk and failures for primary TKA, age > 60

Time	**Survival, S(t)**	**ΔS(t)**	**Numbers at risk, N (t)**	**Number of failures, D(t)**
0	1.000		1000	0
1	0.987	0.014	987	14
2	0.976	0.011	976	11
3	0.969	0.007	969	7
4	0.964	0.005	964	5
5	0.960	0.004	960	4
6	0.957	0.003	957	3
7	0.955	0.002	955	2
8	0.955	0.000	955	0
9	0.949	0.007	949	7
10	0.947	0.002	947	2
11	0.944	0.003	944	3
12	0.942	0.002	942	2
13	0.940	0.002	940	2
14	0.937	0.003	937	3
15	0.935	0.002	935	2
16	0.931	0.004	931	4
17	0.931	0.000	931	0
18	0.924	0.008	924	7
19	0.924	0.000	924	0
20	0.915	0.010	915	9

* Inferred from number of primary TKAs inserted 1994 [5]

Table A.4: Numbers at risk and failures for revised TKA, females age < 60

Time	**Survival, S(t)**	**ΔS(t)**	**Numbers at risk, N (t)**	**Number of failures, D(t)**
0	1.000		1000*	0
1	0.917	0.083	917	83
2	0.839	0.085	839	78
3	0.798	0.049	798	41
4	0.775	0.029	775	23
5	0.753	0.029	753	22
6	0.738	0.020	738	15
7	0.706	0.043	706	32
8	0.699	0.009	699	6
9	0.654	0.064	654	45
10	0.654	0.000	654	0
11	0.635	0.029	635	19
12	0.605	0.048	605	30

* Inferred from number of included observations in [31]

Table A.5: Numbers at risk and failures for revised TKA, females age 60-70

Time	**Survival, S(t)**	**ΔS(t)**	**Numbers at risk, N (t)**	**Number of failures, D(t)**
0	1.000		1000*	0
1	0.932	0.068	932	68
2	0.869	0.068	869	63
3	0.836	0.038	836	33
4	0.817	0.023	817	19
5	0.799	0.022	799	18
6	0.787	0.015	787	12
7	0.761	0.033	761	26
8	0.756	0.007	756	5
9	0.719	0.048	719	36
10	0.719	0.000	719	0
11	0.704	0.022	704	16
12	0.679	0.035	679	25

* Inferred from number of included observations in [31]

Table A.6: Numbers at risk and failures for revised TKA, females age > 70

Time	**Survival, S(t)**	**ΔS(t)**	**Numbers at risk, N (t)**	**Number of failures, D(t)**
0	1.000		1000	0
1	0.948	0.052	948	52
2	0.900	0.051	900	49
3	0.874	0.028	874	26
4	0.860	0.017	860	15
5	0.846	0.016	846	14
6	0.836	0.011	836	10
7	0.816	0.024	816	20
8	0.812	0.005	812	4
9	0.784	0.034	784	28
10	0.784	0.000	784	0
11	0.772	0.015	772	12
12	0.753	0.025	753	19

* Inferred from number of included observations in [31]

Table A.7: Numbers at risk and failures for revised TKA, males age < 60

Time	**Survival, S(t)**	**ΔS(t)**	**Numbers at risk, N (t)**	**Number of failures, D(t)**
0	1.000		1000*	0
1	0.834	0.166	834	166
2	0.678	0.186	678	155
3	0.597	0.120	597	82
4	0.550	0.078	550	46
5	0.506	0.081	506	45
6	0.475	0.060	475	30
7	0.411	0.135	411	64
8	0.398	0.031	398	13
9	0.309	0.225	309	90
10	0.309	0.000	309	0
11	0.270	0.124	270	38
12	0.210	0.225	210	61

* Inferred from number of included observations in [31]

Table A.8: Numbers at risk and failures for revised TKA, males age 60-70

Time	**Survival, S(t)**	**ΔS(t)**	**Numbers at risk, N (t)**	**Number of failures, D(t)**
0	1.000		1000	0
1	0.865	0.135	865	135
2	0.739	0.146	739	126
3	0.672	0.090	672	66
4	0.635	0.056	635	38
5	0.598	0.057	598	36
6	0.574	0.041	574	25
7	0.522	0.091	522	52
8	0.511	0.020	511	10
9	0.438	0.142	438	73
10	0.438	0.000	438	0
11	0.407	0.071	407	31
12	0.358	0.121	358	49

* Inferred from number of included observations in [31]

Table A.9: Numbers at risk and failures for revised TKA, males age > 70

Time	**Survival, S(t)**	**ΔS(t)**	**Numbers at risk, N (t)**	**Number of failures, D(t)**
0	1.000		1000	0
1	0.896	0.104	896	104
2	0.799	0.108	799	97
3	0.748	0.064	748	51
4	0.719	0.039	719	29
5	0.691	0.039	691	28
6	0.672	0.027	672	19
7	0.632	0.060	632	40
8	0.624	0.013	624	8
9	0.568	0.090	568	56
10	0.568	0.000	568	0
11	0.544	0.042	544	24
12	0.506	0.070	506	38

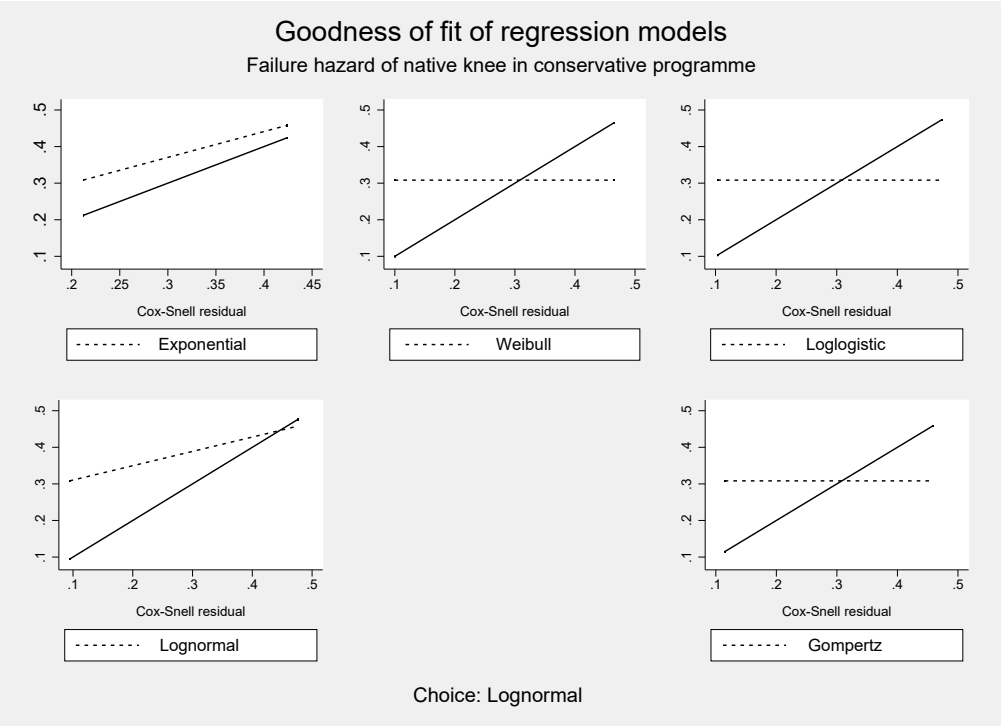
* Inferred from number of included observations in [31]

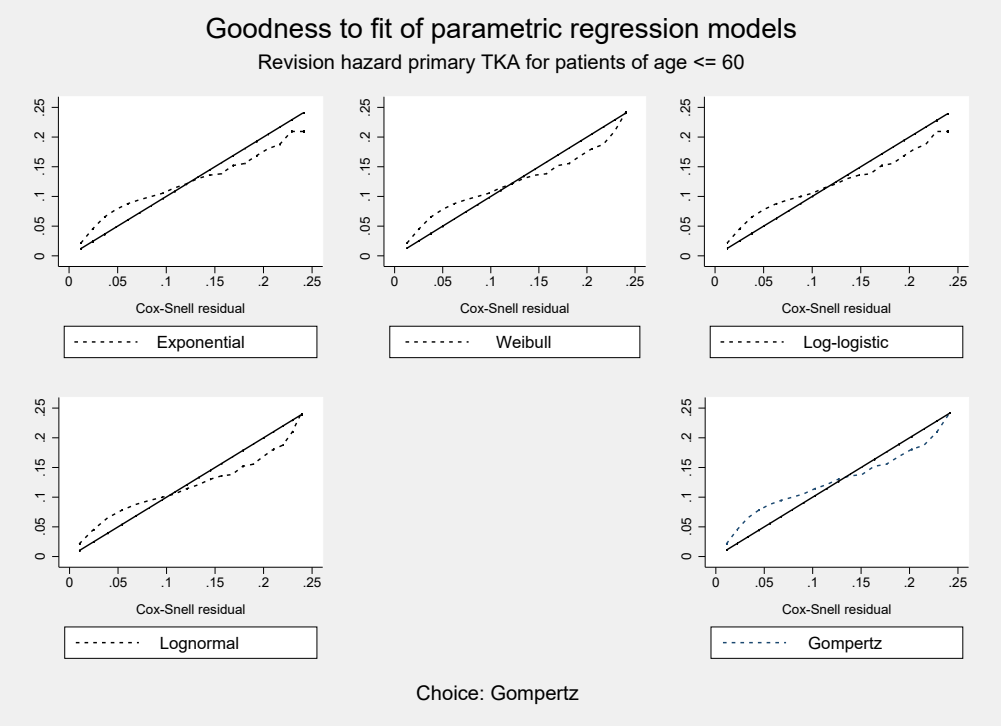
Appendix B

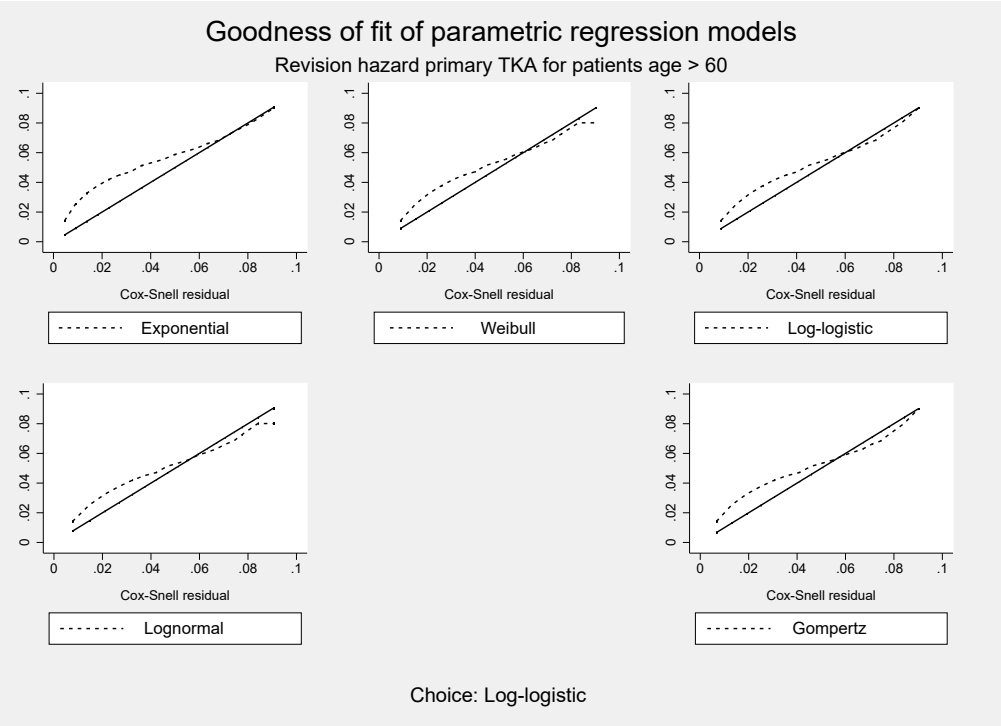
Evaluation of distributional assumptions

Evaluation of survival model appropriateness

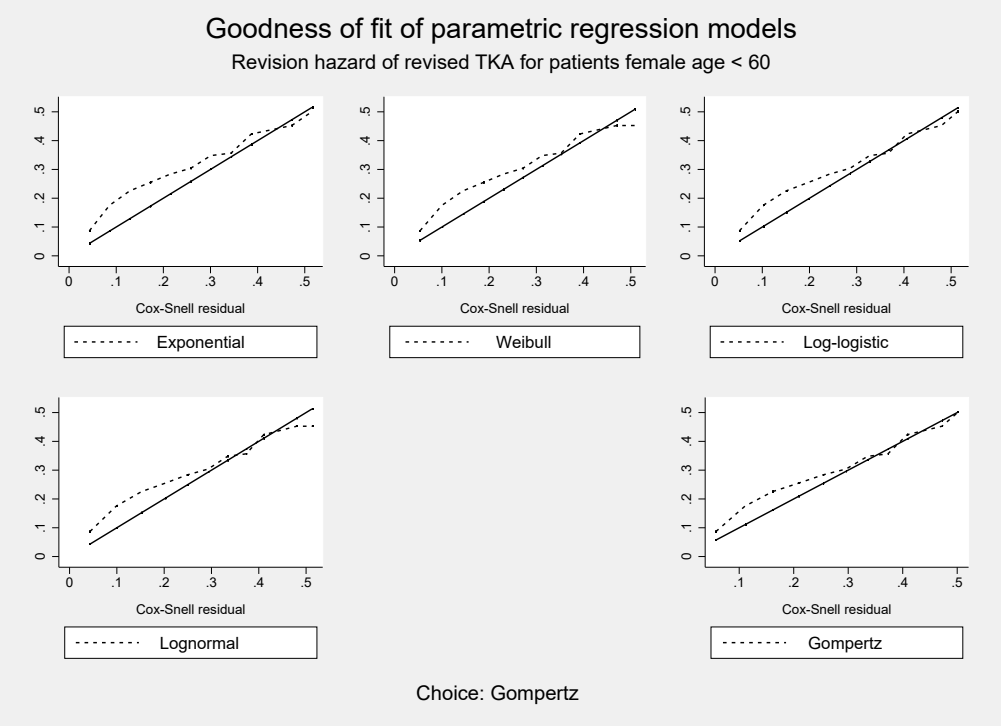
Evaluating the goodness of fit of the parametric regression models can be done by reviewing their Cox-Snell residuals. These plots tell us how good the parametric model fits a Nelson-Aalen plot of the non-parametric hazard. The closer the dashed line is to the solid line, the better the model fits the observations. For some of the plots, it was impossible to make a decision which was best based on the graphs alone. For these, the Akaike information criterion (AIC) was used. The AIC compares the models under the different distributional assumptions by using a -2log likelihood statistic [48]. The criterion suggests a better fit with a smaller AIC value. It should be noted that the model with the smallest AIC value was not always the model which was graphically most appropriate.


Goodness of fit of regression models
Failure hazard of native knee in conservative programme
Cox-Snell residual
Exponential
Weibull
Loglogistic
Lognormal
Gompertz
Choice: Lognormal

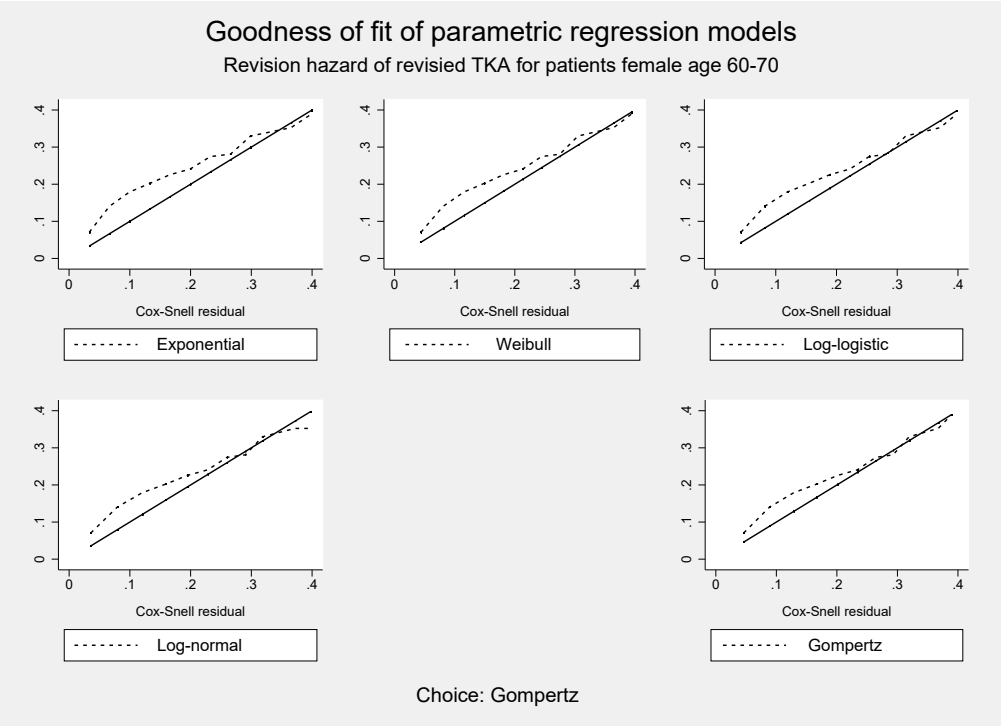

Goodness to fit of parametric regression models
Revision hazard primary TKA for patients of age <= 60
Cox-Snell residual
Exponential
Weibull
Log-logistic
Lognormal
Gompertz
Choice: Gompertz



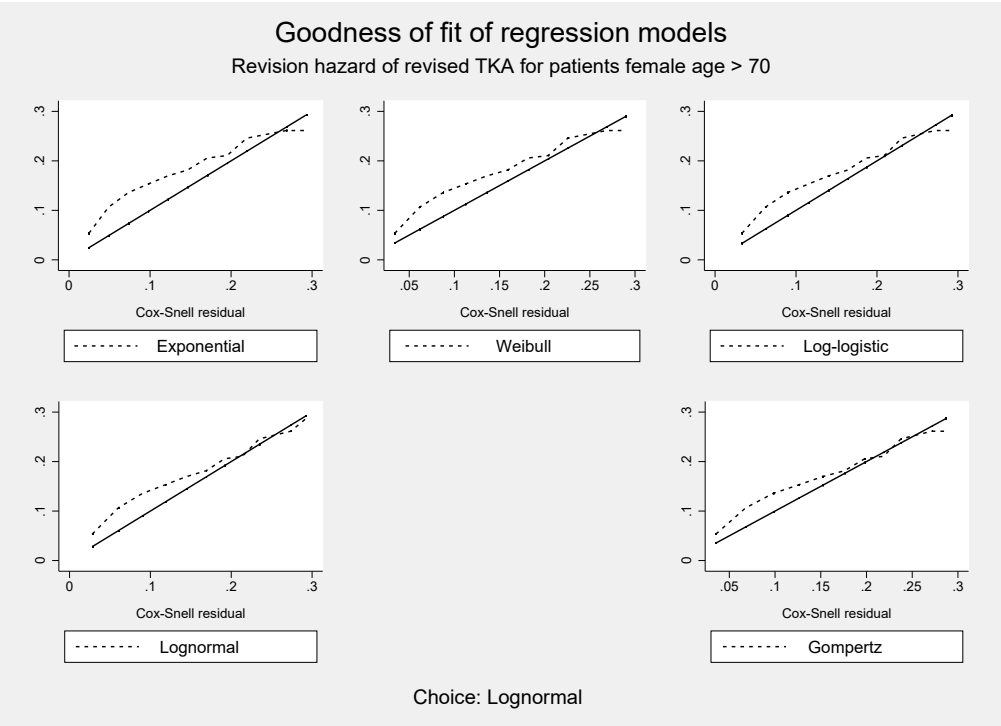
Choice: Log-logistic



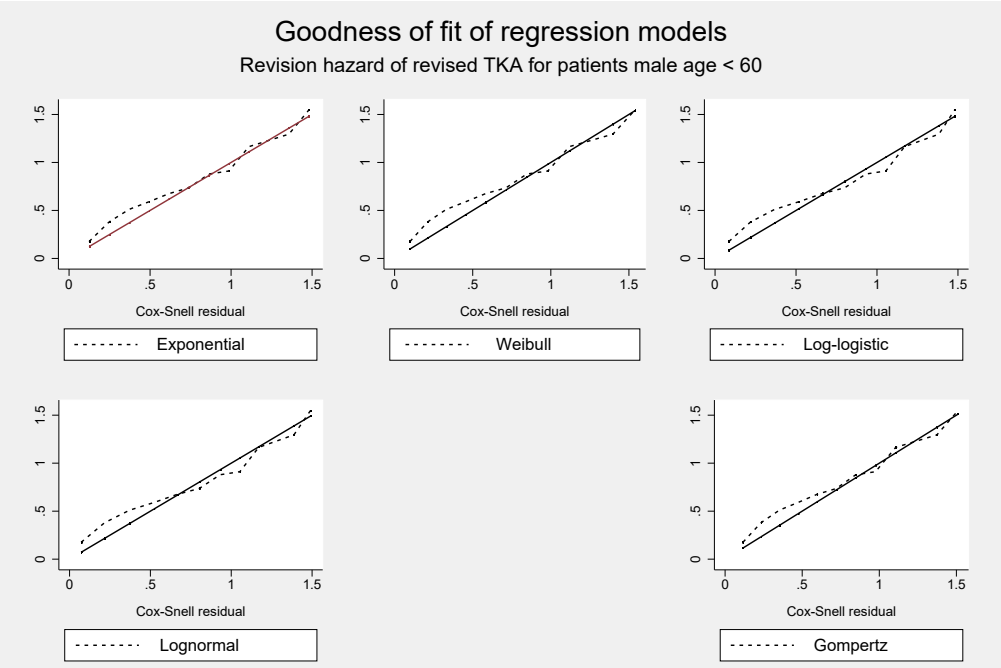
Choice: Gompertz



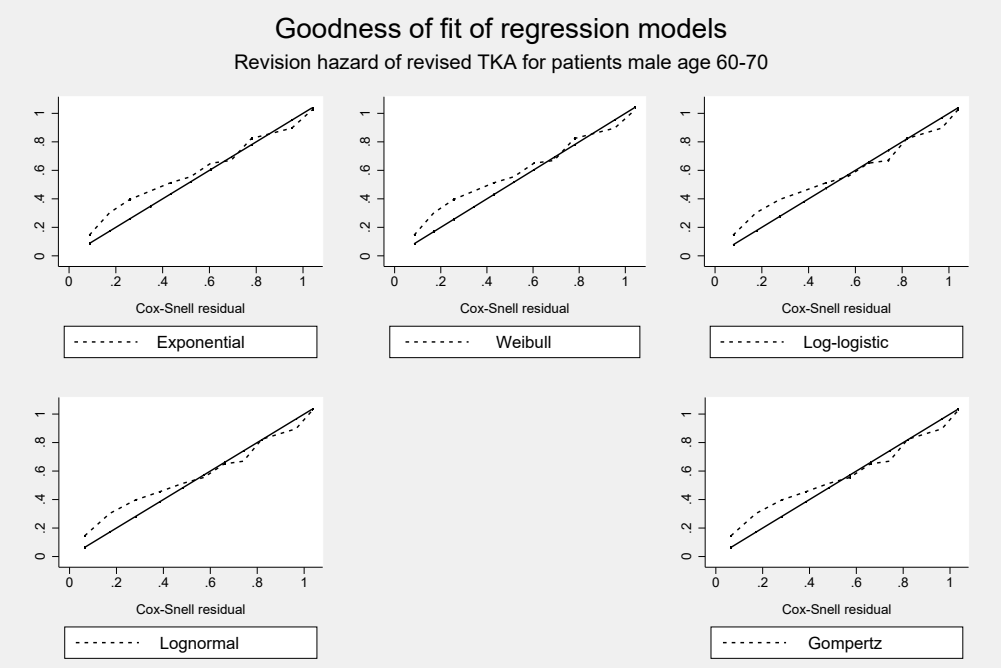
Choice: Gompertz



Choice: Lognormal



Choice: Weibull



Choice: Lognormal

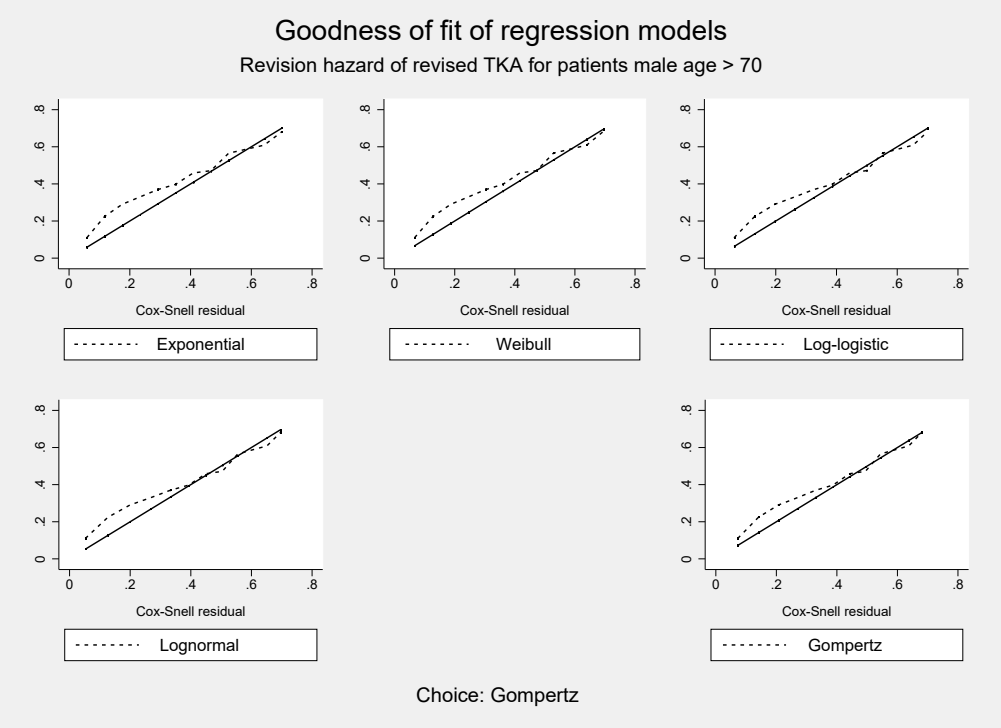

Goodness of fit of regression models
Revision hazard of revised TKA for patients male age > 70
0
.2
.4
.6
.8
Cox-Snell residual
Exponential
Weibull
Log-logistic
Lognormal
Gompertz
Choice: Gompertz

Table B.1: Akaike's Information Criterion values

Failure variable	**Expo-nential**	**Weibull**	**Log-logistic**	**Log-normal**	**Gompertz**
Native knee	87	79	78	76	82
Primary TKA age ≤ 60	1506	1507	1509	1507	1506
Primary TKA age > 60	814	810	805	810	811
Revised TKA female ≤ 60	2296	2294	2281	2281	2251
Revised TKA female 60-70	2041	2037	2028	2004	2025
Revised TKA female > 70	1733	1728	1723	1704	1717
Revised TKA male ≤ 60	2906	2897	2876	2823	2906
Revised TKA male 60-70	2833	2835	2809	2765	2832
Revised TKA male > 70	2572	2572	2554	2518	2562

Appendix C

Estimated resource use and unit costs

Identification

The process of estimating the costs of the conservative treatment programmes started with the identification of the resource use. The resource use of the five interventions or components to the programme needed to be split into two bulks. One consisted of the initial resource to participate in the organised programme. Table C.1 shows the identified resource consumption the initial phase of the programme, excluding medication. The other consisted of the follow-up support and the possible continuation of the need for pharmaceuticals to relieve pain. Table C.2 shows the resource consumption of the continuous cycles of the programme, excluding medication.

The assumption for the use of pain medication is that patients would get prescriptions for three months at the time to use the medication when needed. For the cycles of the programme, a high estimate would then be to need four contacts with the general practitioner (GP) to get a prescription, and a low estimate would be zero. This is the procedure used to assess the resource consumption for the pharmaceuticals as well – the high estimate being the highest need of medication during the cycle, and the low estimate being no need. One unit of a pharmaceutical is one package of it that sufficiently covers the daily prescribed dosage. The num-

Table C.1: Resource use conservative programme inital 12 weeks

Intervention	**Main resource unit**	**Other resource unit**	**Total time (hours)**	**Proffe-sional**
Exercise sessions	Time	None	24	Physio-therapist
Educational sessions			4	Physio-therapist
			2	Dietician
	Time	None	2	OA-patient*
Dietary advice	Time	None	4	Dietician
Othropeadic shoe insoles	Time	Formthotics insoles	0.75	Physio-therapist
Pain relief medication (see own table)	Pharmaceuticals	Time		GP

* The osteoarthritis patient referred to is not the patient undergoing the treatment, but one that informs the target patient.

ber of units that suits each pharmaceutical was estimated from the Norwegian pharmaceutical industy's catalogue of pharmaceuticals [49]. Estimated resource consumption for patients in need of pain medication is given in table table C.3 on the following page.

Unit costs

The unit cost of the physiotherapist was assumed to equal the double of the rates set by the Norwegian Health Economics Administration (HELFO) to cover social costs and out-of-pocket payments [50]. This is the recommended procedure by the guidelines for economic evaluations [36]. The rates varies according to whether the physiotherapist session is group or individual based. For the supervised exercise programme in the initial cycle, the exercise is group based. The fitting of

Table C.2: Resource use conservative programme continuous cycles

Intervention	Main resource unit	Other resource unit	Total time (hours)	Proffe-sional
Telephone-follow-up	Time	None	4*	Physio-therapist
Telephone-follow-up	Time	None	1**	Dietician
Pain relief medication (see own table)	Pharmaceuticals	Time		GP

* 20 minute contact once per month. ** 20 minute contact three times per year

Table C.3: Estimates of resource use per cycle for patients in need of pain medication

Resource	Unit need, low	Unit need, high	Assumed mean (sd)
GP contact	0	4	2 (2)
Acetaminophen 1 g/(4x/day)	0	16	8 (8)
Ibuprofen 400 mg (3x/day)	0	12	6 (6)
Pantoprazol 20 mg (1x/day)	0	26	13 (13)

orthopaedic insoles is individual. The telephone follow-up has its own rate and is assumed to have a duration of 20 minutes.

The dieticians unit cost was estimated as the hourly wage from wage statistics published by Norwegian Association of Dietitians Affiliated With The Norwegian Association of Researchers for 2013 [51]. These were given as yearly wages, and was adjusted to a normal work year of 1950 hours, adjusted for the inflation between 2013 and 2015 and social costs of 40%. The high and low estimates were included for variation.

For the osteoarthritis patient supplementing in the educational programme

the unit cost was estimated to be the productivity loss of an hour of an average Norwegian employee in 2015, valued at forgone earnings with social costs. This was estimated by adjusting the average yearly earnings to a work year of 1950 hours, and 40% social costs.

The unit cost of the GP consultation was taken from the normal tariff for general practitioners in 2014-2015 [52] and doubled as par with the guidelines.

The cost of the orthopaedic insoles were estimated from the market prices in online stores and physiotherapists institutes. Online searches were performed in February 2016 and the low and high end of the market price was included for variation.

The pharmaceuticals unit costs were estimated from the pharmaceutical industry's catalogue of pharmaceuticals [49].

All unit costs of the conservative programme are showcased in table C.4 on the next page.

Cycle implementation

The total cost per cycle is conditional on the probability of needing the pain medication and the additional costs this has (and its variation), and conditional on the probability of having BMI $\geq$ 25 kg/m^2. The probability of needing pain medication was taken from Skou et al. and was given as 58% of the patients at baseline, and 41% at 12 months follow-up. By fitting α– and β values using the known standard deviations, the probabilities was fitted to the beta distribution and included as individual parameters. The probability of being at least overweight is taken from Mork et al. [22] and was given as 72.1%. It was fitted to a beta distribution using the method of moments approach and included as its own parameter.

Table C.4: Unit costs

Resource	**Unit cost low**	**Unit cost high**	**Assumed mean (sd)**
Physiotherapist rate, group			342
Physiotherapist rate, individual			452
Physiotherapist rate, telephone			130
Dietician hourly wage	354	396	375 (20.5)
GP consulation			286
Orthopaedic insoles	499	750	625 (125.5)
Osteoarthritis patient			372
Acetaminophen			77.7
Ibuprofen			87.6
Pantoprazole			52.1

Appendix D

Estimated clinical consequences

This appendix shows how the estimated clinical consequences for the cohorts in terms of how many person-years they spend in selected health states of the model. While this is possibly a slightly unconventional way of using a decision model, it can be a helpful look into why the results are the way they are, especially for the between-cohort differences. Tables D.1 and D.2 shows how many person-years each cohort spends in the states maintaining the native knee in the conservative programme, full benefit post primary and (re)revision TKA, limited benefit post primary and (re)revision TKA, and in the death state. The fourth column shows how these years relative to the cohorts total person-years alive. This number was calculated as the total number of person years less person years spent dead, until the cohort reaches the age of 100. These numbers inform the bar-chart in fig. 6.2 on page 47.

Table D.1: Estimated clinical consequences in person years for all subgroups

Cohort	Strategy	Person years with consequence	Relative to person years alive *
		Keep native knee	
Female			
age 50	Conservative	2 459	11.3%
60	Conservative	2 323	16.0%
70	Conservative	2 049	23.4%
80	Conservative	1517	31.3%
Male			
age 50	Conservative	2 416	12.9%
60	Conservative	2 213	18.4%
70	Conservative	1 855	25.7%
80	Conservative	1 422	29.6%
		Full benefit post TKA †	
Female			
age 50	Conservative	16 720	76.5%
	TKA only	19 801	90.7%
60	Conservative	9 701	66.7%
	TKA only	12 484	86.0%
70	Conservative	5 444	62.2%
	TKA only	8 080	92.8%
80	Conservative	2 456	50.7%
	TKA only	4 382	91.1%
Male			
age 50	Conservative	14 076	74.9%
	TKA only	17 147	91.4%
60	Conservative	8 080	67.1%
	TKA only	10 856	90.4%
70	Conservative	4 251	58.8%
	TKA only	6 630	92.3%
80	Conservative	2 532	52.8%
	TKA only	4 329	90.9%

*Person years alive are defined as total number person years (see below) less time in the dead state.

Cohort	Strategy	Person years with consequence	Relative to person years alive*
		Limited benefit post TKA †	
Female			
age 50	Conservative	1 349	6.2%
	TKA only	1 623	7.4%
60	Conservative	1 150	7.9%
	TKA only	1 493	10.3%
70	Conservative	359	4.1%
	TKA only	518	6.0%
80	Conservative	183	6.0%
	TKA only	318	6.6%
Male			
age 50	Conservative	1 054	5.6%
	TKA only	1 288	6.9%
60	Conservative	648	5.4%
	TKA only	872	7.3%
70	Conservative	292	4.0%
	TKA only	442	6.2%
80	Conservative	194	4.0%
	TKA only	325	6.8%
		Death	**Relative to total person years****
Female			
age 50	Conservative	28 155	56.3%
	TKA only	28 169	56.3%
60	Conservative	25 452	63.6%
	TKA only	25 484	63.7%
70	Conservative	21 244	70.8%
	TKA only	21 291	71.0%
80	Conservative	15 154	75.8%
	TKA only	15 189	75.9%
Male			
age 50	Conservative	31 219	62.4%
	TKA only	31 232	62.5%
60	Conservative	27 952	69.9%
	TKA only	27 978	69.9%
70	Conservative	22 774	75.9%
	TKA only	22 817	76.1%
80	Conservative	15 201	76.0%
	TKA only	15 235	76.2%

** The person years are relative to 50 000 , 40 000, 30 000, and 20 000 total person years respectively depending on being in age group 50, 60, 70, or 80. † Person years in both full-benefit post primary and revised TKA.