Health economic evaluations of shoulder pain, colorectal cancer and scoliosis

Norsk utgave av tittel:
Helseøkonomiske evalueringer av
skuldersmerte, kolorektal cancer og skoliose

Philosophiae Doctor (PhD) Thesis

Pål Joranger

School of Economics and Business
Faculty of Social Sciences
Norwegian University of Life Sciences

Ås/Adamstuen (2015)

Thesis number 2015:16
ISSN 1804-6402
ISBN 978-82-575-1271-2

List of contents

List of contents s. 3

Acknowledgements s. 5

Abbreviations s. 7

List of original papers s. 9

Part I: Introduction s. 11

1. **The aims and the content of the thesis** s. 11

2. **On health economic evaluations** s. 12

2.1 Types of health care evaluations s. 12

2.2 Consequences, costs and perspectives s. 14

2.3 Approaches for organizing and analyzing data s. 17

2.3.1 Patient level data

2.3.2 Decision analytic modelling

2.4 Uncertainty in health economic evaluations s. 20

2.4.1 Definition of uncertainty

2.4.2 How to analyze and handle uncertainty

2.5 CEA, COI and CMA s. 25

2.5.1 Cost-effectiveness analysis (CEA)

2.5.2 Cost of illness analysis (COI)

2.5.3 Cost-minimization analysis (CMA)

2.6 Health economic evaluations applied for policy purposes s. 29

2.6.1 What can Full Health Economic Evaluations (FEEs) be used for in the health sector?

2.6.2 The actual implementation of FEEs

2.6.3 Barriers and incentives in the application of FEEs for policy purpose

3. **Summary and discussion of the papers** s. 37

3.1 Summarizes and critically assess the papers s. 37

3.1.1 Paper I: Costs of shoulder pain and resource use in primary health care: A cost-of-illness study in Sweden

3.1.2 Paper II: Modeling and Validating the Cost and Clinical Pathway of Colorectal Cancer

3.1.3 Paper III: Cost and survival of colorectal cancer and consequences of changing treatment algorithms: A model approach

3.1.4 Paper IV: A health economic evaluation of screening and treatment in patients with adolescent idiopathic scoliosis

3.2 Comparing of the applications and the methods used in the four papers s. 43

3.3 Generalization of the papers results to other settings s. 45

3.3.1 Factors hampering the generalization and what to do about it

3.3.2 Generalization of the results in the four papers

4. **Conclusion** s. 49

4.1 Contributions s. 49

4.2 Key conclusions s. 49

4.3 Policy implications s. 50

Reference list s. 51

Part II: The papers s. 57

Paper I: Costs of shoulder pain and resource use in primary health care: A cost-of-illness study in Sweden. s. 59

Paper II: Modeling and Validating the Cost and Clinical Pathway of Colorectal Cancer. s. 73

Paper III: Cost and survival of colorectal cancer and consequences of changing treatment algorithms: A model approach. s. 113

Paper IV: A health economic evaluation of screening and treatment in patients with adolescent idiopathic scoliosis. s. 147

Acknowledgements

I am very grateful to all of the individuals who supported me throughout this work. First, a special thanks to my main supervisor, Ståle Navrud from the School of Economics and Business at the Norwegian University of Life Sciences (HH-NMBU). I thank him for his flexibility during these years and for being present when advice was needed. I have truly appreciated the interesting and inspiring discussions and the valuable input on this thesis and related topics, as well as all of the helpful advice about both scientific questions and the working process throughout my PhD.

I would also like to thank my co-supervisors: Arne Oshaug at the Oslo and Akershus University College of Applied Sciences (HiOA), who provided invaluable support from the very beginning, as well as interesting discussions; Geir Hoff at The Cancer Registry of Norway and Telemark Hospital, who shared his broad knowledge about colorectal cancer (CRC) and CRC treatment, provided instructive guidance, and provided additional support during my work with the CRC model and my two articles based on this model; and Sverre Grepperud at the Department of Health Management and Health Economics at the University of Oslo (UiO), for providing support and particularly for the good discussions and advices in the last part of the work with my PhD.

Much of my PhD was spent developing, validating and using the colorectal cancer model. In addition to Hoff and Oshaug, Eline Aas, Arild Nesbakken and Halfdan Sørbye also contributed to this work. I would like to thank Eline Aas for helpful support from the very beginning of the work with this model, for good advice, for many important and inspiring discussions on modelling and health economics, for providing access to cost data based on her systematized data on colorectal cancer costs, and for inviting me into her inspiring academic environment in the Department of Health Management and Health Economics at UiO. Thank you to Arild Nesbakken for his support, for his valuable contribution to the model, and for providing access to data on colorectal cancer that were collected by he and his colleagues. His broad knowledge on this topic and his ability to understand and evaluate the model parameters were invaluable to us. Thank you to Halfdan Sørbye for support and for sharing insights into CRC in general and palliative treatment in particular; we were heavily dependent on his knowledge in these areas.

I would also like to thank Jen Ivar Brox for collaborating on the papers on scoliosis and shoulder pain and for many informative, challenging and interesting debates on the application of health economic evaluations within this area of medicine. Thank you also to Lena Virta, Raphael Dziwornu Adobor, Harald Steene and Rikard Eriksson for their collaborations and for all I have learned from them through working on these papers.

Thanks to Mette Wik and Atle Guttormsen at HH-NMBU for useful advice and interesting discussions, and many thanks to the PhD students at NMBU for creating a friendly and interesting place to study. I also thank the staff at HH-NMBU for practical support.

Several people at HiOA deserve thanks. First, my thanks go to former dean Bente Skulstad Austrått and the present head of the former Institutes of Health, Nutrition and Management, Kari Almendingen, for their financial support. In addition, many thanks to Kari Hjerpaasen and Sidsel Tveiten for contributing to my flexibility to prioritize my PhD. Thank you to all of my colleagues who have shown interest and encouraged me, and to the librarian at HiOA-Kjeller for help in finding and obtaining all of the necessary books and articles.

A warm thanks to my mother and father for their interest and support, although they never got to see the end of my PhD. Thank you to family and friends for being there, and for giving me many nice breaks from my thesis work.

Finally, to my dear wife Marit, thank you for all of your support, patience and encouragement throughout the PhD process, despite all the cross-country skiing and weekends in the cabin that we missed. Without you, this PhD would not have been possible. To our dear daughter Vilde, who moved out to study during my PhD, thank you for your interest and support over these years.

Thank you all!

Pål Joranger
Fetsund, august 2015

Abbreviations

CEAC:	Cost-effectiveness acceptability curve
CBA:	Cost-benefit analysis
CEA:	Cost-effectiveness analysis
CI:	Confidence Interval
CMA:	Cost-minimization analysis
COA:	Cost-outcome description analysis
COI:	Cost of illness
CRC:	Colorectal cancer
CrI:	Credibility Interval
CUA:	Cost-utility analysis
DR:	Distant recurrence of cancer
DSA:	Deterministic sensitivity analysis
EVPI:	Expected value of perfect information
FEE:	Full health economic evaluation (including CMA, CEA, CUA and CBA)
FM:	Friction method
FDA:	Food and Drug Administration, US
GDP:	Gross domestic product
HCM:	Human capital method
HEE:	Health economic evaluation
HTA:	Health technology assessments
HUI:	Health utility index
ICER:	Incremental cost-effectiveness ratio
ISPOR:	International Society for Pharmacoeconomics and Outcome Research
LDR:	Local and distant recurrence of cancer
LR:	Local recurrence of cancer
NICE:	National Institute for Clinical Excellence, UK
NMB:	Net monetary benefit
OUS:	Oslo Universitetssykehus (Oslo University Hospital)
PPP:	Purchasing power parity
PSA:	Probabilistic sensitivity analysis
QALY:	Quality-adjusted life-year
RCT:	Randomized controlled trials
SA:	Sensitivity analysis

SD:	Standard deviation
SG:	Standard gamble
SMDM:	Society for Medical Decision Making
TTO:	Time trade-off
WHO:	World Health Organization
WTP:	Willingness to pay

List of original papers

Paper I

Costs of shoulder pain and resource use in primary health care: A cost-of-illness study in Sweden.

Contributions of the candidate: Contributed to designing the study. Participated in the study and drafted the manuscript. Performed the statistical analysis (built the Excel-model and performed the statistical/economic analysis, including the sensitivity analysis). Read and critically revised the manuscript.

Paper II

Modeling and Validating the Cost and Clinical Pathway of Colorectal Cancer.

Contributions of the candidate: Contributed to designing the study, developing the conceptual model, and collecting data from the literature. Built and ran the simulation model, performed the Weibull analyses for estimating parameters for the Markov model, and wrote the drafts of the article. All authors were involved in the analysis and interpretation of results and in drafting and critically reviewing the manuscript.

Paper III

Cost and survival of colorectal cancer and consequences of changing treatment algorithms: A model approach.

Contributions of the candidate: Contributed to designing the study, developing the conceptual model used, and collecting data from the literature. Built and ran the simulation model, performed the Weibull analyses for estimating parameters for the Markov model and wrote the drafts of the article. All authors were involved in the analysis and interpretation of results and in drafting and critically reviewing the manuscript.

Paper IV

A health economic evaluation of screening and treatment in patients with adolescent idiopathic scoliosis.

Contributions of the candidate: Contributed to designing the study. Collected data and performed the health economic analysis. Built and ran the simulation model for the study. Was involved in the analysis and interpretation of results and in drafting and critically reviewing the manuscript.

The Introduction (Part I) is the sole work of the candidate.

Part I: Introduction

1. The aims and the content of the thesis

The thesis consists of four independent research papers being concerned with health care evaluations of the following diseases: shoulder pain, colorectal cancer, and scoliosis. The main aims of this thesis are:

i) Conduct policy-relevant health economic evaluations for: shoulder pain, colorectal cancer (CRC) and scoliosis.
ii) Develop and validate a general model for colorectal cancer.
iii) Apply and discuss the three approaches: cost of illness (COI) analysis, cost minimization analysis (CMA) and Markov models.
iv) Discuss strength and weakness of the approaches used and their applications, the generalizability of the results, and policy implications..

The thesis has two parts, where part I represents the integrative part (introduction) and part II presents the four research papers. Part I has four sections. Section 1 presents the aims of the thesis. Section 2 presents, and critically assesses on a general level, health economic evaluations and relevant approaches for organizing and analyzing the data. The main focus is on cost-of-illness analysis, cost-minimization analysis and Markov models. The section ends by discussing the application of health economic evaluations for policy purposes. Section 3 summarizes and critically assesses the results and the methods used in the four papers. Subsection 3.1 summarizes and critically assesses paper by paper, while subsection 3.2 compares the applications and methods used in the four papers. In subsection 3.3, the problems of generalization of results are introduced followed by a discussion on how the specific results from the four papers can be generalized and applied to other settings. Section 4 concludes by presenting the contributions of the thesis, key conclusions and policy implications.

2. On health economic evaluations

2.1 Types of Health care evaluations

In the health sector, there are many ways to use the available resources to improve public health. Health care evaluations can be used for choosing among competing interventions (2). To see how the approaches used in the four research papers relate to health care evaluations, I use the taxonomy suggested by Drummond et al. (2) which focuses on the following two characteristics of health care evaluations; (i) does the evaluation deal with both inputs and outputs (often called costs and consequences), and (ii) does the evaluation compare two or more alternative interventions.

In table 1 below, taken from Drummond et al. (2), the two characteristics mentioned above are combined into a table to categorize different health care evaluations. Analyses belonging to cell 1A, 1B and 2 do not compare alternative interventions. Drummond et al. (2) denote such evaluations as "descriptions" of the costs or the outcome of one single intervention. Cell 1B is called *cost descriptions* because only costs are taken into account and *cost of illness* studies (COI) belong to this category (2). In cell 2 both the cost and the output is described and the analysis is called *cost-outcome description* analysis (here called COA). In cell 3A and 3B, we find evaluations that compare alternative interventions either according to cost or according to consequences. Randomized controlled trials are examples of evaluations that belong in cell 3A, since alternatives are compared according to their health consequences (efficacy or effectiveness). Cell 3B represents evaluations that compare two or more alternatives with respect to costs only.

Drummond et al. (2) is of the opinion that not all cells of table 1 fulfill the requirements for being full economic evaluations. According to Drummond et al., only cost-effectiveness analysis (CEA), cost-utility analysis (CUA) and cost-benefit analysis (CBA) can be denoted as full economic evaluations comparing alternative interventions since they include both costs and consequences (see cell 4). The three types of analyses differ with respect to which units outcomes are measured in. CEA measures the consequences as physical units, such as a drop in blood pressure, cases detected, or life years saved. This analysis is particularly useful if the consequences of the alternatives compared are measured in the same physical unit. CUA measures all consequences in a generic health-related unit ("health state preference score" or "utility weights"), and is typically based upon individual preferences. The most common measure of consequences in CUA is the quality-adjusted life-year (QALY). The use of this approach implies that all interventions can be meaningfully compared. CBA measures the consequences of interventions in monetary units, based on individual preferences, to make them commensurable with the costs. CBA is based on economic welfare theory applying the principle of Potential Pareto improvement (Kaldor-Hicks) as a value criterion (3,4,5). A particular intervention is socially desirable (i.e. represents a potential Pareto improvement) if the sum of all benefits that follow from an intervention exceed the sum of all costs of the same intervention. Such a decision rule does not consider the distributional impact of the intervention.

Table 1*. Types of Health Care Evaluations; according to Drummond et al. (2)).*

Are both cost (inputs) and consequences (outputs) of the alternatives examined?

Is there comparison of two or more alternatives?		No		Yes
		Examines only consequences	Examines only costs	
	No	**1A Partial evaluation** Outcome description	**1B** Cost description	**2 Partial evaluation** Cost-outcome description
	Yes	**3A Partial evaluation** Efficacy or effectiveness evaluation	**3B** Cost analysis	**4 Full economic evaluation** Cost-effectiveness analysis Cost-utility analysis Cost-benefit analysis

Source: (2)

CEA and CUA are suitable for comparing alternatives and maximize achievement of a defined objective within a given budget (1, 2). This is because increments in the relevant budget require assessment of the opportunity cost that is likely to fall outside the health care sector (2). If CEA or CUA is used to tell whether an alternative is worthwhile or not, we have to make a reference to an external standard like a threshold cost-effectiveness ratio (1).

If a relevant threshold is not known, then the decision makers, in addition to evaluating the alternatives against each other, can also find out if any of them are worth the costs of the interventions, by using cost-benefit analysis (CBA) (1). CBA can then also assess whether the health budget should be increased to accommodate the new alternative (2).

Cost-minimization analyses (CMA) only include costs since the consequences are assumed to be the same across the interventions considered. Drummond et al. (2) do not explicitly locate CMA in any of the cells of table 1. However, based on the discussions in Drummond et al. (2) and Brigg et al. (6), it seems that the relevant cells are 3B or 4, or both, depending on how the consequences are handled.

In the following I use the concept of Health Care Evaluations about all types of evaluations (analyses) that fit into table 1, while I use Health Economic Evaluations" (HEEs) about all health care evaluations that

take costs into account (2). Finally, I use Full Economic Evaluations (FEE) about all HEE that simultaneously include both consequences and costs. I have also chosen to categorize CMA as a FEE; see section 2.2 for the arguments behind this decision.

2.2 Consequences, costs and perspectives

Health economic evaluations may include health effects (health outcomes) and/or costs (inputs) of an intervention. *Health effects* can be measured in physical units such as life years, number of recurrences, blood pressure and number of injuries. In addition, health effects can be measured by quality of life instruments. One example is a QALY which captures improvements in both quality (morbidity) and quantity (mortality) from an intervention (7). To measure the quality element by using direct elicitation from participants, the three most common methods are rating scale, time trade-off (TTO) or standard gamble (SG) (8, 9). Because these methods can be complex and time consuming, pre-scored multi-attribute health status classification questionnaires have been developed (9). Examples of questionnaires include the EuroQoL Group, 15D, Short Form 6D and the Health Utilities Index (HUI). These questionnaires are generic instruments, allowing them to be used for many different health states. Many classification systems exist, and some systems have been validated for only certain types of health states, such as heart disease or diabetes mellitus (10). The methods chosen can to a large extent influence the estimated quality of health outcome (2).

Drummond at al. (2) define *costs* as the consumption of resources in association with planning, implementing and maintaining the intervention. Such costs can be imposed on the patient and the patient's family, on the health care sector, or on society. In addition, health care interventions may also reduce future costs (e.g. lower costs of care). Some literature defines such effects as consequences while others define them as "saved" costs. The literature also typically distinguishes between direct costs and indirect costs (2, 11-13). The *direct costs* include resources consumed (costs) or saved (benefits) by the intervention (2). Important parts of these costs will often be the time used by physicians, nurses, and other providers of health- and non-health service. The *indirect costs* include production losses or gains due to a change in morbidity and mortality rates.

Two approaches used for estimating costs are gross-costing (top-down) and micro-costing (bottom-up) (14). Of these, micro-costing is somehow considered as the "golden standard" (12). It spells out the production and cost function for the service analyzed. According to Wordsworth et al. (15), the approach should be considered for treatment where: i) the cost of the staff or overheads are important, ii) there is extensive sharing of staff or facilities between treatments or patient groups, or iii) where health care costing systems do not routinely allocate costs to the intervention level. Under these circumstances, the

bottom-up approach could increase consistency and transparency and hence comparability of costs. However, this approach is relatively costly and time consuming.

A typical process for estimating costs involves three distinct steps (2, 3, 14): i) identifying the relevant cost-items, ii) measuring the use of the cost-items (quantities), and iii) placing a value on one unit of these items (prices). When *identifying*, we have to find all relevant cost-items, both inside and outside of the health sector, if a social perspective is used. This step can be seen as a study of the production function of the intervention (14), and a comprehensive list of cost-items should be prepared so the analysts better can decide on what items to include in the next step (3). The choice of using micro- or gross-costing will influence the identification process because the micro-costing breaks down the cost-items into small components while gross-costing implies larger components (14). In both cases, we need to both *measure* and *value* the components.

When *measuring* components in gross-costing, national average figures such as reimbursement rates can be used, while for micro-costing we can e.g. break down a surgery to the use of equipment, medicine, and hours per surgeon or nurse. An appropriate time horizon has to be used to cover all relevant costs (3). Data from local settings can be used if the aim is to support a local decision, while national registry data may be more appropriate for national health policy. One should put most effort into collecting precise data for parameters with the greatest impact on the final results. Possible sources of information include prospective studies (like randomized controlled trials, pragmatic trials, and observational study), registries, international scientific publications, and expert opinions. For a more detailed cost study, special patient surveys or patient diaries can be used (14).

After measuring the cost components we need to *value* the cost-items in monetary terms. Ideally in an economic perspective, the costing should be based on the principle of *opportunity cost*, which in our context means the value of the foregone benefits of using those resources (either used or lost due to illness) in the best, alternative way (13, 14). However, as the estimation of opportunity cost of every resource used would be very demanding and time consuming, in practice approximations are used.

Some of the cost-items can be valued in competitive markets like some types of transportation services, cars, fuel, computers and food. The market price for these items can be good proxies for the opportunity costs. Since many of the resources used in and produced more exclusively for the health sector are not priced in a transparent and competitive markets (14), the analysts often have to value these items in separate cost analyses.

One way to cost *productivity losses* is to apply the human capital method (HCM) (1), which is an example of using the market price as a proxy. Here the value of employment for society is set equal to the gross wage of a worker (the marginal productivity). By using this approach, the costs of leisure time and work time can be estimated. For patients outside the labor market, the "replacement value approach" can be used,

which for example means that the housewife's production is valued by the corresponding value of such work in the market. Alternatively, based on the "opportunity cost approach", the value of unpaid work is assumed to be at least as much as the net wage rate that the same person would attain in the market place.

Informal care, i.e. care provided by informal caregivers such as patient's family, friends, acquaintances or neighbors, can also be estimated by the replacement value approach or by using opportunity cost (1, 12).

An alternative approach to HCM for measuring the value of work time is the *friction method* (FM)(16, 17), which focuses mainly on the valuation of lost time from paid work caused by illness (18). The developer (17) of the method argued that HCM is based on unrealistic assumptions about wage flexibility and labor markets clearing. Instead, there will exist a pool of unemployed people which can replace the sick person, and there will only be a loss of productivity during a "friction" period. HCM estimates are often many times higher than the estimates obtained from the FM (17, 18).

The inclusion of some types of cost in economic evaluations is being discussed, e.g. productivity losses caused by sick leave and morbidity-related reduced productivity during work hours (13, 19, 20). Another question is whether non-related health care costs should be included (16, 21). If evaluating a treatment that saves patients from cancer death, should we then only include the cost of the cancer treatment or should we also include the cost of other non-cancer-related care in the health sector (21)?

The choice of which costs to include will also depends on the *perspective taken*. Different decision makers care about different interests, such as the interests of a certain hospital, the health care sector, the patient and the patient's relatives, the employer, or the whole society. Some evaluations choose to include the interest of the health care sector and the patient, while other evaluations may choose to include the interest of the whole society. The choice of perspective will have an impact on the value of an intervention (e.g. cost-effect ratios) and might change the ranking of competing projects. Because the choice of which costs to include can be important to the conclusion about cost-effectiveness, the decision maker could participate in the process of choosing which costs to include, or at least ensure that they understand which potentially important costs are excluded. For some evaluations, the process of deciding which costs to include is not only based on sound principles, but rather on which costs are possible to estimate within the time and resources available.

2.3 Approaches for organizing and analyzing data.

The techniques for organizing and analyzing the data for health economic evaluations can be divided into two categories: i) evaluations based on patient-level data, and ii) decision-analytic modelling evaluations (9, 22). In the presentation below, the decision-analytic modelling is emphasized.

2.3.1 Patient level data

Evaluations using patient-level data are often based on a randomized controlled trial (RCT), where the researcher asks cost-related questions about the use of time, the use of services from the health sector and transportation expenses, and estimates the health effect of the relevant intervention (2, 9, 23). RCTs continue to provide an important source of data for HEEs, but they have several potential limitations/shortcomings such as short-term follow-up, partial nature of the comparisons undertaken (restricted number of alternative interventions), use of intermediate health outcomes, and unrepresentative patients (often without co morbidity etc), clinicians and locations (24).

2.3.2 Decision analytic modelling

Decision-analytic modelling is based on statistical decision theory and shares common theoretical origins with both expected utility theory and Bayesian statistics (2). Decision analytical modelling can be defined as a systematic approach to decision making under uncertainty (25). With its set of methods, it can satisfy the following objectives of any economic evaluations (2): (i) *structure* reflecting the possible prognosis and the effects of the interventions evaluated, (ii) by an analytic framework the *evidence* relevant to the study can be brought to bear, (iii) provide an *evaluation* by translating the evidence into estimates of cost and effects of the relevant alternatives, and identify the best alternative by using the appropriate decision rule for the relevant HEEs (f ex CEA, CUA and CBA), (iv) facilitate an assessment of *uncertainty and variability*; and finally (v) through uncertainty analysis assess priorities for future research.

Decision-analytical modelling is, as mentioned above, closely associated with Bayesian statistics. The Bayesian and frequentistic approaches are two competing philosophies of statistics (26). The frequentists are the dominant group, but the Bayesian approach has received increased attention within health evaluations. The frequentists represent classical statistics, while the Bayesian approach holds that unknown properties of the population have probability distributions about which we can have subjective beliefs. According to the Bayesian approach, to estimate the parameter value and its distribution for the model (and the Probability Sensitivity Analysis (PSA)), we can merge our own newly collected data (*likelihood function*) with existing information or beliefs that we have about the parameter (*prior distributions*). This strategy will generate a parameter value and a probability distribution (*posterior distribution*) that we can use in our model (27). The likelihood function and the prior distributions can be weighted according to their credibility. In this way, the Bayesian approach helps us to merge all we know about a parameter, and

avoid using only the most recently collected data. The approach also has tractable properties for the decision maker like estimating the probability that one intervention is better than another, rather than using hypothesis testing (27).

One of the choices that we have to make when conducting decision analytic modelling, is to choose the appropriate approach for modelling the prognosis for the disease we study (28-32). Modelling techniques comprise; (i) decision trees, (ii) Markov models, (iii) discrete-event simulation, and (iv) other approaches (25, 33). In this thesis the main focus is on Markov models.

2.3.2.1 Markov models

The role of the Markov model in health economic evaluation is to provide an analytical structure that represents key elements of a particular disease (34) which subsequently can be used in health care evaluations such as COIs and FEEs. The Markov models can be split into cohort models and patient-level simulation models (25). Each approach has different advantages and disadvantages which can affect the results in different ways. In this thesis, I will focus on the Markov model based on a cohort. Compared to decision trees, the advantages of the Markov model include the explicitly estimated timeline and the ease of handling diseases where the patient may relapse to the same health state many times. Compared to the patient-level model, it is easier to de-bug the cohort model, and the simulation can be done faster (22). This is particularly important when conducting PSA (25). Briggs et al. (25) argue that for some evaluations the patient-level simulation can give a more detailed picture of reality, but this comes at the expense of increased data requirements and computational burden.

The principal elements of the model: The Markov model includes some principal elements (1, 22, 25, 34). The model has a set of mutually exclusive and collectively exhaustive health states, and the person is always only in one of the finite number of possible states in a given period of time. The person stays in a particular health state for a certain time period called the Markov cycle length (e.g. weeks, months or years), and move from state to state according to defined probabilities known as transition probabilities. A person can either move to another state, or to the same state for another period; and move only once per cycle. The transition probabilities can either be time-dependent, or constant over time. A person only stays for one period in temporary states, and never exit absorbing states (typically the death state). Each state can have an assigned cost and utility associated with staying in that particular state. An important model assumption is the so called Markovian assumption (34). This means that the transition probability only depends on the current health state, thus ignoring in which state the person has been earlier. Thus, the individual state has no "memory" of the natural history of the disease.

Estimating survival, utility and costs: To estimate survival time we estimate the average number of cycles (the amount of time) spent in each state where the persons are alive, and get:

$$\text{Expected survival} = \sum_{s=1}^{n} t_s$$

were t_s is the time spent in state s, and n is the number of health states. Often, the quality of survival is important. Then we can associate each state with a quality factor representing the quality of life in the particular state relative to perfect health attaining the value 1 (34). This can be estimated by:

$$\text{Expected utility} = \sum_{s=1}^{n} t_s \cdot u_s$$

where u_s is the utility in health state *s* (34). Expected costs are estimated by using the following formula:

$$\text{Expected cost} = \sum_{s=1}^{n} t_s \cdot c_s$$

where c_s is the cost in health state *s*.

Transition probabilities: The transition probabilities have to be specified. Given n number of health states, and letting a_{ij} represent the probability that a person will move (transit) from state i to state j within one particular cycle, we get the n x n matrix A = a_{ij} for all the transition probabilities. By definition,

$$\sum a_{ij} = 1.$$

Transition probabilities can be derived from data in the literature, from primary data, or from registers. Some of the data are available from published papers such as the probability of a patient getting a certain treatment (e.g. the probability of prescribing adjuvant chemotherapy).

Running the Markov model: There are three methods for running/calculating the Markov model (8): (i) Fundamental matrix solution, (ii) Cohort simulation (from now on called *cohort based Markov model*), and (iii) Monte Carlo simulation (25). In the latter we randomly select people from a cohort, and each of them transits through the model at a time. In contrast, the cohort simulation simultaneously tracks the whole cohort through the model. The Markov model can also be solved by using matrix calculations (35) if the transition probabilities are constant throughout all cycles (8). If they instead are changing, we can conduct cohort simulations by for example using a spreadsheet and produce a Markov trace to show how the whole cohort moves through the model (8). In this thesis, I focus on cohort simulations (Cohort based Markov models).

As already mentioned, one important limitation of the cohort based Markov models is the Markovian assumption. The problem of no "memory" of the natural history can be solved by building time-dependency into the model. There are two kinds of time dependency. Transition probabilities can vary according to the time the person has been (i) in the model and (ii) in a particular state (25). The former imply that one or more transition probabilities changes as the cohort ages. These can be built into the model by changing them as the cohort ages. If e.g. cancer is the disease, both age-dependent changes in background mortality, recurrence rate and cancer related mortality could be relevant to build into the model. The second kind of time dependency can often also be important to build into the model. For example, both the recurrence rate and the cancer related mortality can change with the time elapsed since the first year of cancer treatment. This can be handled by using tunnel states, which means a sequence of states that is linked together so that persons in the tunnel are only coming into a state from the former state in the tunnel. This secures that the model can account for how long time the person has been in the state that the tunnel represent. Then the model can both let the transition probabilities, costs, and QALY vary with time since the person has been in a particular health state (e.g. the state when the tumor was removed by surgery). To handle the changes in transition probabilities by age and by elapsed time since a particular health state, one possibility is to combine the use of tunnel state and 3-dimensional data matrixes, as done in Joranger at al. (36).

Another limitation of cohort based Markov models is that each patient undergoes only one state transition during a single cycle. In the model we can count the membership in the different states either in the start or in the end of each cycle, while in reality people transit continuously through each cycle. To correct for this we must typically perform half-cycle corrections for survival time, QALYs and costs. In some analyses the costs have already been corrected as part of the cost estimation of the sub-models. The analysts have to make choices both with respect to using half-cycle-correction or not, and between different ways of correcting (8, 37, 38).

2.4 Uncertainty in health economic evaluations

Health economic evaluations are often comprehensive and based on a range of elements that contribute to different types of uncertainty. Thus, it is important in health economic evaluations to analyze and handle the uncertainty.

2.4.1 Definition of uncertainty

The sources of uncertainty have been categorized in many different ways (2, 3, 8, 39-44), and there are no common way to categorize them (39). Here, I have chosen to categorize the sources of uncertainty for evaluations (i.e., stochastic analysis) in a way well known from the health economic literature (2, 43), first

developed by Briggs et al. (44). For patient-level data the following categories are used: methodological uncertainty, sampling variation, extrapolation and generalizability/transferability.

For uncertainty related to decision-analytic modelling studies, I distinguish between the following sources of uncertainty (2, 43): (i) Methodological uncertainty, (ii) Parameter uncertainty, (iii) Modelling uncertainty (including: structure- and process uncertainty), and (iv) Generalizability/transferability. The same categorization is also in the Norwegian guidelines for preparing model based HEEs in the health sector (20, 45).

Methodological uncertainty is the uncertainty related to the method used to estimate the parameters for the health effect, resource use, unit cost, etc. Which methodological assumptions to choose are often discussed among experts, and this brings uncertainty to the results from HEEs. Further, the HEEs are based on data sets and a set of methodological choices, which raise the following questions: for which group of people is this evaluation relevant, or to which group of people can this evaluation be *generalized*? This process of generalization will always generate some level of uncertainty for the decision maker.

For the model-based evaluation, *parameter uncertainty* is the uncertainty related to the estimation of the parameters that measure health effects, resource use, unit costs, quality of life year (QALY) estimates, etc. This uncertainty is partially based on the sample variation, which is produced by the natural variation across the respondents, given the statistical method used, that can be described by standard errors, p-values or confidence intervals.

Modelling uncertainty is related to the uncertainty produced by the model structure and the modelling process. For a Markov model, *model structure* can e.g. be the number of health states and the path that the patients can follow between those states, allowing for time dependency or not, the length of a cycle, and the time horizon of the model. The distinction between methodological and structural uncertainty appears to be unclear, and some authors merge these two sources under the term structural uncertainty for model-based evaluations (11). *Modelling process uncertainty* occurs as a result of the many choices made by the particular analyst or team of analysts (3). For the user of the evaluations, these choices contribute to the uncertainty.

2.4.2 How to analyze and handle uncertainty

Uncertainty related to decision analytical modelling will be emphasized because most of my work in this thesis is concerned with such models. The theoretical framework for handling uncertainty in this kind of modelling is closely related to the theoretical foundation of decision analytical modelling (see section 2.3.2).

For analyzing uncertainty, we can both use deterministic and probabilistic sensitivity analysis (DSA and PSA, respectively). Both are usually based on a base case version of the model in comparison with an alternative treatment. The base case model builds upon the assumptions and the parameters that the analysts consider to be the most trustworthy. DSA and PSA will then estimate the difference between the base case estimate (result), and the estimates generated from varying some of the model elements (assumptions, structure, and parameter values).

Deterministic sensitivity analyses (DSA) generate one expected value from a simulation, while probabilistic sensitivity analyses (PSAs), since based upon inputs with stochastic properties, generate a probability distribution of possible values. DSA can be classified as (46): i) one-way SA, ii) multi-way SA (including scenario analysis and best/worst case analysis), and iii) threshold analysis (see table 2). *One-way* SA is the simplest method for performing a SA. Here we change one element in the base case model, and calculate how the result changes. Then we change this element back to the base case situation, and change another element to determine how this changes the result, and then do the same for all the assumed uncertain elements, one by one. We can then identify to which elements the model results are most sensitive when changing them, and by how much and in what direction each element changes the model results.

Table 2*. Types of SA for analyzing the different types of uncertainties (40, 43, 44).*

Type of uncertainty	Type of sensitivity analysis (SA)
Methodological	Deterministic SA: One- and multi-way SA Probabilistic SA: Scenarios
Parameter uncertainty	Deterministic SA: One-way, multi-way (incl. scenario and best/worst case), threshold and analysis of extremes Probabilistic SA: All or a selection of parameters simultaneously
Modelling uncertainty Structural Process	 Deterministic SA: One- and multi-way Probabilistic SA: Scenarios No obvious methods
Generalizability/ transferability	Deterministic SA: One-way, multi-way, threshold Probabilistic SA: Scenarios

Modified from: Briggs (43).

When interested in changing several model elements at the same time, we use a *multi-way* SA (often named a *scenario analysis*). Another type of multi-way analysis is the *best/worst case analysis*, in which we change all uncertain elements in the direction that generates the best/worst possible result. Such an approach is particularly interesting for risk adverse decision makers. For uncertain parameters, we are often interested in knowing at which level of the parameter the result changes from a gain to a loss. In that case; a *threshold analysis* can be used.

One important advantage of DSA is that such an analysis is simple to understand, and typically easy to implement. However, there are also a number of potential problems (40): (i) The estimate

of the expected value will be biased in nonlinear models. (ii) It is difficult to decide what can be considered to be an "extreme but plausible" value for some or all parameters when not knowing the distribution of the parameters. (iii) Furthermore, it is also difficult to know whether the parameter threshold value is likely or extremely unlikely when not knowing the distribution of the parameters. The use of PSA is one way to handle the problems that arise from using DSA.

Probabilistic sensitivity analysis (PSA)*:* The principle of a PSA is rather intuitive. Each parameter is assigned a probability distribution. From each of these distributions, simultaneously drawn values are entered into the model (8) and the results are then computed. This process is repeated many times (often 10 000 – 100 000 times) to generate a large number of results that themselves constitute a probability distribution for the overall result. Monte Carlo simulations are often used and this type of simulation samples from the distributions at random. From the distribution of the result, we can estimate the expected value, the credibility interval, and the probability that each evaluated alternative is cost effective. These simulations can also be used to calculate the maximum value of additional evidence (25, 40).

Because using parameter distributions, PSAs are often criticized for adding another layer of uncertainty that should also be subjected to sensitivity analysis. Briggs et al, however, argue that this is not necessary (25). The same distributions that are used in PSAs are also used to estimate parameters, and there are often a small number of candidates for the distributions for each type of parameter (25).

Another criticism of PSAs is the assumption of independence between parameters since some of the parameter values could, at least to some degree, be correlated (25). It is possible to build correlations into PSAs, but we often lack the necessary information concerning such correlations. Performing PSAs can be time consuming, since, for each parameter, a distribution must be defined, important properties must be estimated, and values must be programmed into the simulation model. Further, it is more time consuming to run the estimations for PSAs than for DSA, particularly so for large models. Therefore, in practical modelling, we often observe that the comprehensive Markov model with many health states and related cost models are not constructed as a PSA (31).

Handling of methodological and structural uncertainties in PSA: Gray et al. (1) argue that "Structural uncertainty is an under-researched area of uncertainty, but may contribute to even greater uncertainty than parameter uncertainty". To handle methodological and structural uncertainties within the framework of a PSA, three methods are proposed: (i) probabilistic scenarios, (ii) model averaging, and (iii) parameterization.

Probabilistic scenarios: When we are performing a PSA, we typically let all parameters change simultaneously while other elements (method and structure) remain unchanged. If so, the result of the PSA is based on the specific assumptions about method and structure, as if there was no uncertainty about these elements. However, we are often uncertain about these judgments. We should then change

the relevant method or structure and perform a new PSA simulation for each of the alternative assumptions or combinations of assumptions (40). By using such scenarios, the analysts can externalize the uncertainty related to these assumptions and remove the uncertainty from the evaluation by simulating and presenting the scenarios, leaving it to the decision maker to decide which scenario is the most credible. This could be a good solution if the decision maker is able to assess the credibility of the different scenarios. If not, it is probably better that the analysts handle such an uncertainty as a part of the analysis. For this, model averaging or parameterization can be used.

Model (or scenario) averaging: When performing model averaging, we do the following: (i) evolve the relevant alternative scenarios (as for probabilistic scenarios) and use a PSA to simulate the costs and effects of these scenarios, (ii) weight these costs and effects according to the assumed credibility of each scenario, and (iii) based on the distributions of the costs and the effects of the scenarios and the corresponding weights, estimate the new result and the new distributions for the total costs and effects (40). This new distribution is the weighted average of all of the relevant scenarios and provides the overall decision uncertainty and the consequences of this uncertainty. Parametric uncertainty, methodological uncertainty, and structural uncertainty can be included in the result. For larger models, only a certain portion of the relevant scenarios can be included. Thus, the excluded scenarios generate some undefined uncertainty. This problem can be easier to handle by using parameterization.

Parameterization: According to Claxton, nearly all cases of structural and methodological uncertainties in models can be handled as a missing parameter or an uncertain parameter (40). Such an approach implies that scenarios are special cases of a common "meta-model" where the missing parameters are taken as extreme values. One example is the situation of having two different data sources (i.e., A and B) for the recurrence of colorectal cancer after resection. Then, for example, we can (i) use A and neglect B or (ii) use B and neglect A. For both choices, we can think of this as setting a weighting parameter to some extreme value. Either A gets all the weight or B gets all the weight. However, if both data sources were reasonably reliable, an alternative could be to include both data sources in the model. This approach could be achieved by including a parameter that indicates the weight given to the two data sources, based on the degree to which those data are relevant and biased. In this way, we avoid working with a large number of scenarios, as could easily occur when using model averaging.

One of the main challenges of implementing model averaging or parameterization is the question of how to find the values and distributions for the wages or the parameters. This question is discussed in Bojke et al. (47), Claxton (40) and Jackson et al. (48).

2.5 CEA, COI and CMA

In this section, I will focus on COI and CMA since these are the main types of health economic evaluations used in the four research papers. However, some attention is also given to CEA to put COI and CMA into a broader perspective, and a kind of CEA is performed as part of paper III.

2.5.1 Cost-effectiveness analysis (CEA)

For the CEA we calculate both the costs and the health effects of a particular intervention and of one or more alternatives (the comparator). Thereafter, we calculate the differences in costs between the two options (ΔCost) and the difference in health effects between the two options (Δeffect). Finally, we calculate the ratio of the two differences (the cost difference and the effect difference), which provides us with the incremental cost-effectiveness ratio (ICER). The formulae for ICER is presented below (see equation (1)):

$$ICER = \frac{Cost_A - Cost_B}{Effect_A - Effect_B} = \frac{\Delta Cost}{\Delta Effect} \qquad (1)$$

The ICER can be located in and assessed in a cost-effectiveness plane. By using a bootstrap analysis, given a patient-level analysis (1, 9, 43), or by using a Monte Carlo simulation given a decision analytical modeling (25), a scatter-plot of points can be produced in the cost-effectiveness plane. This again can be used for making a cost-effectiveness acceptability curve (CEAC)(1, 43). By building upon Bayesian statistical methods the CEAC can be interpreted as the probability that an intervention is more cost-effective than its comparator. The CEAC can be said to represent the decision uncertainty in the CEA (2).

An alternative to the ICER is the concept of "net monetary benefit" (NMB) presented by Drummond et al. (2), Briggs et al. (25) and Glick et al. (9). NMB is also being recommended in the Norwegian guideline for priority setting (20) as one way of presenting the results from health economic evaluations. NMB is a simple re-arrangement of the cost-effectiveness decision rule: If Cost and Effect still reflects the costs and effects associated with each of the two options and R_T denote a particular threshold value (the maximum willingness to pay for a unit of Effect), then the intervention considered is deemed cost-effective if $\Delta C/\Delta E < R_T$. Rearranging the same inequality yields $\Delta E \bullet R_T - \Delta C > 0$, where $\Delta E \bullet R_T - \Delta C = NMB$. An intervention should be adopted if $NMB > 0$.

This decision rule is entirely equivalent to the standard rule that follows from using the ICER (25), but it has some advantages(1). First, we do not need to worry about equivocal interpretations of positive or negative ICERs, and second, using the linear expression of NMB, the sampling distribution will be much closer to a normal distribution than for the ICER (1). The ICER has the problem of approaching infinity if ΔE approaches zero.

2.5.2 Cost of illness analysis (COI)

The Cost of illness (COI) analysis focuses only on costs. However, the cost studies are not analyzing the cost of an intervention, but the costs associated with an illness (disease). Thus, the main aim of the COI analysis is to measure the economic burden of illness to the society (12, 13). According to Drummond et al. (2), identification of the costs and their measurement in monetary units is similar across most health economic evaluations, so all costing methods discussed in section 2.2 are also relevant for COI analyses. The analysis represents the earliest type of health economic evaluations (12), and is currently a standard analysis used by organizations like the World Bank (12) and the US National Institute of Health (49).

Some of the most important choices we have to make when doing COI analyses are (12, 13, 49): (i) the epidemiological data used (incidence versus prevalence approach), (ii) whether using a retrospective versus a prospective study for data collection, (iii) what cost components are included, and (iv) what methods to use when estimating the economic costs. An incidence study refers to the new number of cases arising during a predefined period of time, while a prevalence study refers to the total number of cases that exists in a defined period of time (for example 6 months). The prevalence approach generally gives higher costs of illness than the incidence approach. By using a prospective instead of retrospective study design, the analysts can better design the data collection according to the data needed (12). However, the prospective design can for some diseases take a long time and be more costly (12).

There has been an extensive debate about the COI analysis (11-13, 19, 49-53). Shiell et al. (19) accentuate some of the important objections against the approach. These are:
(i) The COI analysis rests on an intuitive economic logic that equates the cost of illness with the benefits of treatment, which implies that the treatment removes all illness and its consequences.
(ii) They questioned the use of HCM for estimating the amount that should be spent to save lives, and argue that it will lead to a bias towards those diseases which affect white, middle-class males in employment. They put forward loss of life years or quality-adjusted life years as possibly more relevant measures (19). Even if HCM is used in a more limited way (e.g. productivity loss) to estimate the economic impact of disease, they argue that the methods have shortcomings – e.g. that a given level of sickness-absenteeism would not end with the expected reduced production assumed for the HCM if the sick was easily replaced with labor outside the labor market. Further, Shiell et al. (19) doubt the existence of perfect labor markets, which is important to the assumption that labor costs equals the productivity.
(iii) The use of COI analysis has embedded circularity. The COI analyses could estimate illnesses that already receive large resources, which then will be more costly than under-prioritized illnesses. If the COI analyses are used for prioritizing within the health sector, this can lead to circularity.

COI analyses are only focusing on estimating the cost of illness, and authors have argued that COI analyses can contribute directly or indirectly to better health related decisions in the following ways:

(i) Information of the costs of a disease can help policymakers to decide which diseases need to be addressed (12, 54). While COI can be used as a first help to the policymakers to see which diseases need to be addressed (50), the CMA, CEA, CUA and CBA can be used for analyzing which intervention to implement based on cost-effectiveness.

(ii) Estimates of the cost of illness are produced and can be used in CEAs, CUAs and CBAs (13). This is a practice used for some health economic evaluations of colorectal cancer screening (55-57) and result from both models (36, 58) and "model-free" (59) approaches are used to estimate cost of illness.

(iii) COI analyses can show the financial impact a disease has on the heath sector.

(iv) COI analyses were the first economic evaluation approaches used in the health field (12), and much of the costing methods developed have been adopted in CMAs, CEAs, CUAs and CBAs (1-3, 14). Since most of the methods used in COI analyses, including the HCM, are also used in economic evaluation more generally, critique of the methods of costing in COI analyses will often also be a critique of the costing methods of CMA, CEA, CUA and CBA, and vice versa. Not only total cost but also intermediate costs estimated in COI analyses are of potential interest for other evaluations, and therefore should be reported if possible with confidence or credibility interval, so they could be use in other deterministic or probabilistic analyses.

2.5.3 Cost-minimization analysis (CMA)

CMA is a health care evaluation for comparing alternative interventions to find out which one is the most cost-effective. Because of the assumption that the health outcomes of the compared interventions do not significantly differ, we only need to be concerned with costs, and the most cost-effective alternative follows directly from choosing the option with the lowest costs (1).

CMAs are sometimes used when a prospective economic evaluation is being conducted alongside a clinical trial that fails to find any significant difference in the primary clinical outcome (1). However, as Briggs and O'Brian (6) argue, the failure to find a difference in a study designed and powered to test the hypothesis that the health outcome differ between two alternative intervention, cannot be interpreted as evidence of no difference. As Altman (60) stated "absence of evidence is not evidence of absence". Demonstrating equivalence in health outcomes (non-inferiority design), typically requires a much larger sample size than when testing for differences, which is the most common in RCTs.

Briggs and O'Brian also argue that the focus of the analysts should be on the joint density of the cost and effect difference, the uncertainty surrounding the ICER, and the presentation of the related cost-

effectiveness acceptability curves. To attain this, they argue that the uncertainty surrounding the difference between the health outcomes has to be included in the evaluation.

According to Briggs and O'Brian (6) it is seldom that CMA can be used as a "full" economic evaluation, and, most likely, for the same reason, Drummond et al. (2) did not explicitly locate CMA in cell 4 of table 1 (full economic evaluations). However, Drummond et al. did not place the CMA in another cell either. The problem with how to interpret the CMA also becomes apparent as they placed CMA in cell 4 ("full economic evaluation") in the previous edition (second edition) (61) of the book mentioned above (2).

Briggs and O'Brians (6) first argument above raises two questions: i) how sure should we be about "no difference" between the health outcomes, and ii) how can we determine this. They argue that by using sufficiently powered randomized controlled trials (RCT) and statistical tests, one can determine if the health outcomes of the interventions are sufficiently similar. However, is this sufficient? What if possible methodological limitations can cause bias for the randomized controlled trial (RCT)? For some RCTs, this uncertainty can be marginal, while for others it can be quite important. The importance of this uncertainty can be assessed by experts within the field, and can in principle be added to the statistical uncertainty (parameter uncertainty) to determine the total uncertainty.

Further, what if there exist a lot of relevant RCTs but these are too heterogeneous to be summarized in a meta-analysis? Consequently, the health outcomes have to be assessed by experts. Drummond et al. (2) refer to Briggs and O'Brian arguments (6), as well as argue in accordance with the need of the opinions of experts, and write that "The only possible application of CMA is in situations where a prior view has been taken, based on previous research or professional opinion, that the two options are equivalent in terms of effectiveness". They add that one might question the basis on which this professional view has been formed. I will argue that this situation is very similar to situations we are often experiencing, implicit or explicit, when performing health economic evaluations. Analysts have to use previous research and their professional opinions to decide on (i) which methods to use during data collection and for estimation of unit costs (for example using HCM or friction methods for estimating the cost of sick leave), and (ii) which model structure to implement if a simulation model is used (for example choose to use a Markov model or a decision-tree). These two decisions represent methodological and modelling uncertainties, respectively. Usually, these decisions are made without any estimates of the resulting methodological or modelling uncertainty, and often without (or with a limited number of) estimates of the result from alternative choices regarding modelling and the methods used. When doing a PSA entirely based on parameter uncertainty, which is the most common base for PSAs, the choices made about methods and modelling are taken for granted, and assumed without estimating the additional uncertainty related to these elements.

There is reason to believe that in some cases, some of these "ignored" choices may contribute more to the total uncertainty of result than the uncertainty related to the assumption about similarities between the

interventions in question. Therefore, I will argue that, if the decision makers or experts on the relevant health outcomes say that CMA could be used, the health economists have to scrutinize the arguments to find out if one can be sufficiently sure about the presumed similarity in health outcomes. However, uncertainty related to the size of the difference between the health effects of the compared alternatives should not be focused more than the uncertainty of other decisions about modelling or methodological concerns with comparable level of uncertainty.

2.6 Health economic evaluations applied for policy purposes

Decision making about alternative use of health care resources is a critical issue for governments and administrators in all health care systems, and economic evaluations can be useful in determining the economic effectiveness and efficiency of different alternatives (62).

The literature on the application of economic evaluations for health policy purposes is mainly focusing on the use of FEEs, which is also the case for section 2.6.

2.6.1 What can Full Health Economic Evaluations (FEEs) be used for in the health sector?

To decision makers, FEEs can be used for two kinds of decisions (40, 63): (i) The FEEs focus primarily on making the right decision about *which alternative* to implement based on the current uncertainty surrounding the results. However, simultaneous to this decision, the decision maker must also (ii) decide whether to accept the level of uncertainty or to *collect more evidence* before a decision is made (40). Based on current health economic evaluations, the decision makers must typically rely on deterministic or probabilistic SA to appraise this last question, and they must do so in an informal way. However, in the last years, a growing literature on how to formalize this decision based on the concept of the expected value of perfect information has emerged (25). Based on a PSA, both the "value of evidence" and the correct expected value can be estimated. Thus, from the viewpoint of a decision maker, the

Table 3. Result from Cost Utility Analysis (CUA) for eight hypothetical interventions. The incremental cost and incremental QALYs are estimated for a given population.

Intervention	Incremental cost (mill. Euro)	Incremental effectivenes (QALYs)	ICER (Euro per QALY)
1	4	800	5 000
2	8	200	40 000
3	6,5	1 100	5 909
4	9	120	75 000
5	5	250	20 000
6	12	150	80 000
7	18	850	21 176
8	12	750	16 000
Total	74,5	4 220	

Source: Inspired by Gray et al. (1)

inherent uncertainty is important to acknowledge, both for decisions concerning whether or not to choose a new intervention, and for decisions on whether or not to ask for more information/research (1, 5).

FEEs can be used to choose *which alternative* to implement based on the current uncertainty surrounding the results (point i above). According to the guidelines for health economic evaluations, CUA are often the preferred type of FEEs to use (3, 20, 64), as it is suitable for comparing all alternatives and maximizes the objective for a given budget (1, 2). In the following I will use CUA as an example. We focus on a decision maker provided with eight independent CUAs. For each analysis, the most cost-effective alternative is compared with the second best alternative. In table 3, the eight independent interventions are presented, assuming that the interventions were given to all patients needing the treatment within the jurisdiction of the decision maker.

Table 4. *Presentation of eight independent interventions according to incremental costs, incremental effectiveness, ICER, cumulative effectiveness and cumulative costs.*

Intervention	Incremental cost (mill. Euro)	Incremental effectiveness (QALYs)	ICER (Euro per QALY)	Cumulative effectiveness (QALYs)	Cumulative cost (mill Euro)
1	4	800	5 000	800	4
3	6,5	1100	5 909	1900	10,5
8	12	750	16 000	2650	22,5
5	5	250	20 000	2900	27,5
7	18	850	21 176	3750	45,5
2	8	200	40 000	3950	53,5
4	9	120	75 000	4070	62,5
~~6~~	~~12~~	~~150~~	~~80 000~~	~~4220~~	~~74,5~~
Total	74,5	4220			

Source: Inspired by Gray et al. (1)

We assume that the decision maker has a total available annual budget equal to 70 mill €. The objective of the decision maker is now to decide which of the eight alternatives (interventions) should be implemented. This can be done by sorting the eight alternatives from the most to the least cost-effective one, and calculate their cumulative costs (see table 4) in order to identify the most cost-effective interventions that can be financed given the available budget of 70 million €.

Following the procedure presented above, implies that all interventions with the exception of intervention number 6 would be implemented (see table 4 and the left part of the graph (the solid line) presented in Figure 1).

Table 5. *The ranking of the interventions when a new intervention is available.*

Intervention	Incremental cost (mill. Euro)	Incremental effectiveness (QALYs)	ICER (Euro per QALY)	Cumulativ effectiveness (QALYs)	Cumulativ cost (mill Euro)
1	4	800	5 000	800	4
3	6,5	1 100	5 909	1 900	10,5
8	12	750	16 000	2 650	22,5
5	5	250	20 000	2 900	27,5
7	18	850	21 176	3 750	45,5
New	16	700	22 857	4 450	61,5
2	8	200	40 000	4 650	69,5
~~4~~	~~9~~	~~120~~	~~75 000~~	~~4 770~~	~~78,5~~
~~6~~	~~12~~	~~150~~	~~80 000~~	~~4 920~~	~~90,5~~
Total	90,5	4 920			

Source: Inspired by Gray et al. (1)

What would happen now if a new independent intervention became available (see table 5)? In order to maximize the health benefits within the given budget, the decision maker would now implement all interventions (inclusive the new one) with the exception of interventions 4 and 6.

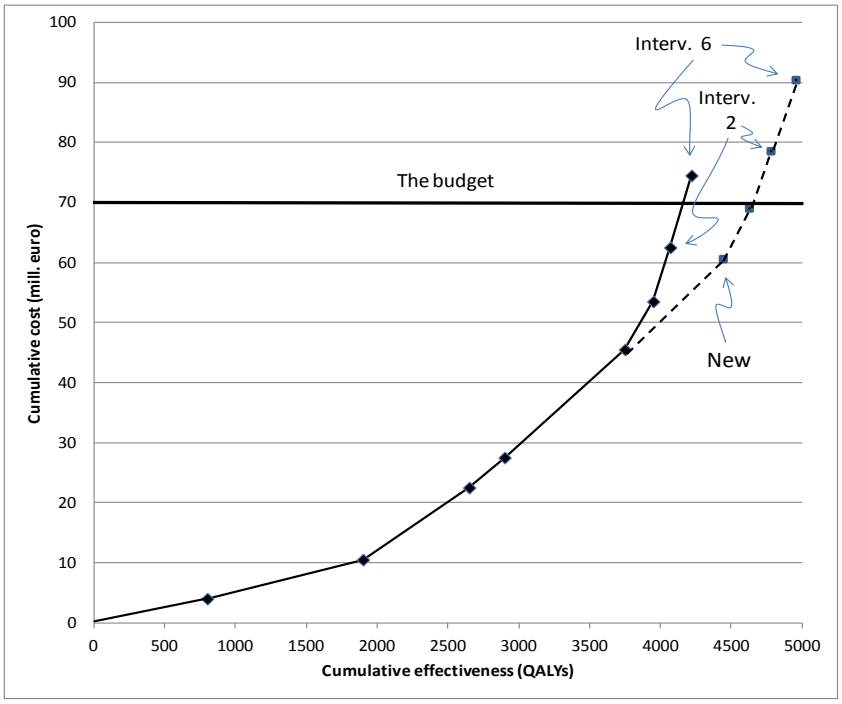

Figure 1. *The effect of using cost-effectiveness to maximize health gain. The solid line is before, and the dotted line is after the implementation of new technology. Source: Inspired by Gray et al. (1)*

From Figure 1, we observe that the introduction of the new intervention causes an outward shift in the top of the graph (see the dotted line), reflecting that the introduction (and implementation) of the new intervention increased the sum of the health benefits that can be achieved for the available budget.

This was a hypothetical example. Reality is more complex. One problem when adapting new technology can be to transfer money from the existing interventions to financing the new one. This problem constitutes one of the main criticisms of the cost-effectiveness approach (1). Further, in the above example, all relevant data for all possible interventions are available at the same time. This is not necessarily the case in practice, implying that only a limited number of possible interventions are being assessed at a time (2). One way to cope with such a problem is to use the maximum acceptable incremental cost-effectiveness ratio as a threshold for whether to adopt or not adopt a new intervention assessed by a CUA.

For a given budget, it is important to identify the correct maximum acceptable ICER. If set too high or too low, the health benefits will not be maximized for the available budget. Ideally the level should be equal to the opportunity cost of the interventions being displaced by a new and more costs-effective intervention (64). If the maximum acceptable ICER is too high, the new adopted technology could replace more cost-effective existing programs (65-67). If it is too low, the health sector would not adopt new technology that is more cost-effective than some of the existing programs that it could replace.

There are (at least) four different ways of identifying the threshold values (1): (i) the league table approach, (ii) the rule-based approach, (iii) the revealed preference approach, and (iv) the stated preference approach. *The league table approach* is already illustrated in table 3 and 4, where we assumed full information on the ICER of all the relevant interventions. By ranking the interventions by their respective ICER (table 4), the ICER of the lowest ranked intervention being included defines the threshold value. In our above example this value is 75 000 Euro pr QALY (the ICER of intervention 4 in table 5) which again reflects the opportunity cost of implementing the new intervention.

The rule-based approach refers to that the health authorities establish an explicit threshold that shall be used by analysts and decision makers (20, 64). The threshold might be somewhat arbitrary, for example, based on previous FEEs and guidelines, and/or on former practice in the health sector (1). In the UK, NICE has established the following rule (64): Below a ICER of £20 000 per QALY gained, the decision should normally be based on the cost-effectiveness estimates and the acceptability of the technology. If the ICER is in the range of £20 000 to £30 000, other factors (e.g. uncertainty the methods for estimating the QALYs and the innovative potential) should also be taken into account. When the ICER is above £30 000 the other factors must be significant if the intervention considered is to be included (64). World Health Organization (WHO) has advocated that the maximum cost-effectiveness ratio should be related to a country's gross domestic product (GDP) per capita, and suggests as a rule of thumb that an ICER less than the GDP per capita should be considered very cost-effective (1).

Revealed preferences is a method for estimating what a decision maker (or society) is willing to pay for health improvements: This is done by systematically examining the health care decisions actually made by the decision makers (revealed preferences). We can both use decisions on national level (e.g. national guidelines) and micro-level decisions (65, 66) (e.g. which group should receive a certain chemotherapy). *Stated preferences* is an alternative method for determining the willingness to pay for health improvements where decision makers or members of society are asked to report their willingness to pay. This approach can be divided into contingent valuation studies and discrete choice modeling (68).

Above I have focused on the role FEEs may have in informing decision makers in choosing between potential interventions (static efficiency). However, the use of FEEs may also have long-term effects. The systematic use of FEE will act as an incentive for the developers of new technology (innovators) to focus on cost-effective technologies (dynamic efficiency) implying that costs, in addition to effects, become important in the development phase.

In which situation and for whom can FEEs be useful? Generally, FEEs can be relevant for assessing medical devices, procedures and pharmaceuticals (69), or more specifically and according to NICE (64, 70), FEEs can be relevant for assessing new technologies within medicinal products, diagnostic techniques, medical devices, surgical procedures, therapeutic technologies other than medicinal products, screening tools, systems of care, and health promotion activities. For analyses of medicines, Simones (71) argues that FEEs can be used by policy makers to inform the allocation of scarce health care resources; health care payers can apply evidence about the value for money of medicines to inform pharmaceuticals pricing/reimbursement decisions; health care professionals can use FEEs to shed light on alternative methods for managing a specific disease; and pharmaceutical companies can use FEEs to demonstrate the value for money of their medicines.

Usually, health economic evaluations are conducted to identify which interventions that generate most health benefits relatively to the resources consumed (1). However, cost-effectiveness is only one dimension that may be considered when deciding on how prioritize between different interventions (72). Some of the evaluations or considerations which can be relevant to the decision makers, and useful to be assessed before carrying out FEEs, are (2, 33, 73): (i) efficacy assessment, where we explore if the relevant intervention can work in a well-controlled setting (do more good than harm) when the patients fully comply. Randomized controlled trials are often used to assess this. (ii) Effectiveness assessment, which is used to find out if the intervention also will work in a real-world practice setting ("Does it work"). Here, pragmatic trials or practical clinical trials are used. (iii) Availability assessment, which is used to analyze if the intervention reaches those who need it.

Another important dimension is *equity* considerations - how costs and health gains are distributed between different groups (income, wealth, ethnicity, and health severity). Such considerations can also be included

into economic evaluation, by for example explicitly weighting health benefits and cost between targeted groups of patients or population (74-76).

2.6.2 **The actual implementation of FEEs**.

How formalized prioritization processes are, and to which extent FEEs are explicitly used in such processes, vary widely between countries (33, 62, 67, 69, 71, 77-81). A country which has formalized the process relatively extensively and given FEEs an important role is the UK (69). The appraisal of a health technology used in the UK is divided into three distinct phases: (i) scoping, (ii) assessment and (iii) appraisal (64). In the scoping process the appropriateness of the proposed remit[1] is determined by the National Institute for Health and Care Excellence (NICE), and the specific questions that each technology appraisal will address are defined. The assessment consists of two components: a systematic review of the evidence, and a health economic evaluation (64). These health technology assessments (HTAs) aim to assist those who make key decisions regarding the allocation of scarce health care resources (69) in the appraisal phase.

FEEs are of central importance to the NICE Appraisals Committee in the process of reaching their decisions on health technologies (78). NICE provides an independent, tailored CUA for each health technology which the Committee wants to assess and employs experts in health economics to support the appraisal process (64). Dakin et al. (80) modelled NICE's decisions in binary choices for or against a health care technology. They found that cost-effectiveness alone correctly predicted 82% of decisions and few other variables were significant. The chance of NICE rejection for technologies costing £27 000, £40 000 or £52 000 per quality-adjusted life-year (QALY) was 25%, 50% and 75%, respectively. Past NICE decisions appear to have been based on a higher threshold than £20 000–£30 000/QALY (80).

Also, other countries like Australia, France, Canada, the Netherlands, Sweden, Belgium, Scotland, Norway and Taiwan, use information from FEEs to support their decisions (20, 67, 69, 71). In Norway the health care system is almost fully financed by general taxation, meaning that all citizens are covered by a National Insurance Scheme (67). FEEs are compulsory when assessing new prescription-only medicines for the reimbursement scheme. The Norwegian Medicines Agency is responsible for setting maximum prices on these medicines, and evaluates and decides whether or not a medicine should be reimbursed by the National Insurance Scheme. Preparing a CUA to inform these decision processes is required (45, 67). The Norwegian Directorate of Health recommends a reference value for costs per QALY of NOK 500 000 in 2005-NOK (equivalent to 59 000 €), and insists that this is not to be interpreted as a threshold (81). However, interviews of stakeholders concerning outpatient pharmaceuticals, confirmed that this value is a strong indication of the Norwegian health system's willingness-to-pay, and claims that cost-effectiveness

[1] A "remit" is a brief given to NICE by the Department of Health and Welsh Assembly Government when a technology is referred to NICE for appraisal.

ratios of NOK 800 000 (94 000 €) or higher would "immediately be rejected" (81). In Sweden SEK 900 000 (97 000 €) could, rarely be accepted. This suggests that cost-effectiveness is an explicit reimbursement criterion in Norway and Sweden (81).

In contrast to the countries mentioned above, in the US and Japan, FEEs has not had much of an impact on prioritization decisions (69, 82). US Centers for Medicare and Medicaid Services have a policy that cost-effectiveness is not considered in national coverage determinations; and although Chambers at al. (83) identified a number of instances where cost-effectiveness evidence was cited in national coverage determinations, they found no clear evidence for the use of an implicit threshold.

There are also differences in the degree of coverage of drug expenditures between UK and US. Mason et al. (84) analyzed coverage decisions by five decision-making bodies in the US and UK on all anticancer drugs approved by the US Food and Drug Administration (FDA) from 2004 to 2008. In the US 100% of drugs were covered, mostly without restriction, while the UK bodies made positive coverage decisions for less than half of the licensed drugs (NICE 39% and Scottish Medicines Consortium 43%), and applied considerably more restrictions than the US bodies. This study gives a reasonable indication of the impact of undertaking HTAs with an economic component - greater restrictions and longer time before the coverage decisions are made (69, 84).

2.6.3 Barriers and incentives in the application of FEEs for policy purpose

A range of *barriers* to using economic evaluations for policy purposes have been identified (62, 77, 85-89). Hoffmann et al. (77) conducted a study of European health care decision makers from 9 countries and used both standard questionnaires (887 respondents), personal interviews (53) and 10 focus groups. They found that only a small percentage had undergone training in health economics, and the majority of respondents had only a poor knowledge of CBA, CEA or CUA. As barriers in the use of study result they found the following, ranked in order of importance: (i) difficulty in moving resources from one sector/budget to another, (ii) sponsorship of studies (by the industry etc.) biases the results, (iii) budgets are too tight to free resources to adopt new therapies, (iv) savings are anticipated and not real, (v) economic studies make too many assumptions. The assertion in point (ii) is supported by a study done by Bell et al. (90), which shows that FEEs funded by industry were more likely to report ICERs under the threshold level, to have lower methodological quality and were published in journals with lower impact factors. Similarly, in a retrospective pairwise comparison, Miners et al. (91) found that estimated ICERs in analyses submitted by manufacturers to the technology appraisal programs of the NICE were on average significantly lower than those submitted by academics. Innvær et al. (85) did a systematic review of 24 interview studies (18 were from outside Europe) with health policy-makers and their use of research evidence at a national, regional or organizational level; and found that the most commonly reported barriers were absence of personal contact between researchers and policy-makers (11 of 24 studies), lack

of relevance or timeliness of research (9/24), mutual mistrust between policy-makers and researchers (8/24), power and budget struggles (7/24), and poor quality of research (6/24).

Hoffmann et al. (77) reported also on *incentives* (encouraging factors) for the use of result from economic evaluations and found the following, ranked in order of importance: a need to better explain the practical relevance of the results (actual cost savings etc), more training in health economics, more comparability of studies, more flexible health care budgets, and easier access to studies (e.g. publications in widely read journals). Innvær et al. (85), however, found that the most commonly reported facilitators for using economic evaluations were personal contact, timely relevance, and the inclusion of summaries with policy recommendations.

The differences in results from Hoffmann et al. (77) and Innvær et al. (85) can partly be caused by the differences in the countries studied and the methods used. However, some common results seem to be the barriers related to budget struggles, mistrust and the quality of the research, and incentives related to explaining better how to apply the results from the economic evaluations.

Drummond et al. (33) propose some key principles for the improved conduct of health technology assessment (HTA) for resource allocation decisions, and organized it in four categories: (i) the structure of HTA programs, (ii) methods of HTA, (iii) processes for conducting HTA, and (iv) the use of HTA in decision making. For the category "methods of HTA", they argued that HTAs should incorporate appropriate methods for assessing costs and benefits; wide range of evidence and outcomes should be considered; a full societal perspective should be considered; explicitly characterize uncertainty surrounding estimates; and consider and address issues of generalizability and transferability. For the category "use of HTA in decision making" they argue that the HTA should be timely; the HTA findings need to be communicated appropriately and adjusted to the different decision makers; and the link between the finding from the HTA and decision making processes needs to be transparent and clearly defined.

Some argue for a wider view on the role of health economic evaluations, and claim that so far the health economists have tended to focus on direct or instrumental use of FEEs, which may overlook the longer-term influence of the health economists on health care resource allocation and then underestimate the opportunities to have greater impact in the future (72). "Research communities – particular those engaged in economic evaluation – can be part of a debate which gradually re-frames prioritization debates so that over time the issue of scarcity can be more explicitly addressed" (72).

3. Summary and discussion of the papers

3.1 Summarizes and critically assess the papers

3.1.1 Paper I: Costs of shoulder pain and resource use in primary health care - A cost-of-illness study in Sweden

The aim was to assess the costs associated with primary health care and loss of productivity for the patients with shoulder pain in Sweden.

Methods: We performed a cost-of-illness (COI) study. Based on a prospective bottom up approach, data were collected for six months from patient records at three primary health care centers in two municipalities in Sweden. Since this was a prevalence based COI study, it was suitable for estimating the annual cost for the group of patients. One reason for doing a prevalence study was the significant costs (time and resources) associated with undertaking a data collection. Because both the human capital and the friction cost approaches are considered to be adequate methods, and the choice of method have important impacts on the result, both approaches were used for estimating productivity loss.

Contributions: This paper contributes in three ways: i) The analysis was the first COI analysis performed on shoulder pain in Sweden, ii) it provided usable data, iii) it generated expected cost of illness estimates (with confidence interval) that could become input into cost-effectiveness analyses of different treatment strategies in primary health care, and iv) in the analyses of uncertainty, we combined a DSA with statistical analyses under different assumptions. This type of analysis is rarely conducted.

Results: A total of 204 (103 women) patients were registered. We found that 20% of the patients were responsible for 91% of the total costs and for 44% of the health care costs. The mean health care cost per patient was €326 (standard deviation (SD) 389) during the 6 months, and physiotherapy treatment accounted for 60% of this cost. Of the total costs, sick leave accounted for 84% of the cost (using the HCA). The mean annual total cost was €4 139 per patient. The costs for sick leave have a strong influence on total costs. Thus, interventions that can reduce long periods of sick leave are warranted. Health care interventions should focus on returning people to the workforce, with special attention to the small group that generates the highest costs.

Uncertainty: We reported the 95% confidence intervals (CI) for the base case scenario for total costs and for health care costs in table 6 (€1 283-2 856 and €273-380, respectively). These intervals reflect the variability across patients and occurs because patients use services, such as x-ray and PT consultations, with different frequencies. This analysis parallels the patient-level analysis.

The SA was performed by using the sample variation related to the unchanged frequencies of health service use and the changed cost level per unit, which was assumed to be the same for all patients. For each tested parameter value or assumption, we computed both the new expected costs and the related 95% CI. To estimate the methodological uncertainty, we used scenario analysis and found that the result was most sensitive to the method for estimating the unit cost for one day of sick leave (i.e., either the human capital method or the friction cost method). To test for uncertainty related to generalizability, a multivariate linear regression analysis was used to explore how gender, age and municipality, as independent variables, predicted total costs and health service cost (see Chapter 3.3.2). These factors did not influence the total costs or the health service costs.

Discussion: A limitation of the study was that we had information about sick leave periods prescribed by the general practitioner, but did not know if patients were actually absent from work all that time or elsewhere contributed with other kinds of production. Further, we had no information about short-term sick leave, if patients had sick leave prescribed by the orthopaedic surgeon post-operatively, or if patients outside the labor market have reduced their production in informal sector. Another limitation was that the cost for medication was probably underestimated because we had no information on the consumption of drugs or of the medication paid out of pocket. However, medication had a minor contribution to the total cost. Generalization to other settings might be difficult, and will depend e.g. on how diagnostic codes are used, the treatment procedure, and the cost per unit.

Strengths of the study were that: (i) we captured almost all patients consulting with all types of shoulder pain during a six month period, (ii) both health care costs and production losses were included in the COI, and (iii) the uncertainty analysis tested for some elements regarding generalizability, parameter uncertainty and methodological uncertainty.

Since the COI analysis in Paper I does not estimate costs per patient for lifetime, the cost estimate from this study cannot be compared with the costs of an intervention that cures the patient, with the aim of estimating cost-effectiveness of the intervention. To attain this, we could use a Markov model to simulate the whole life span. We could use data about the average length and cost per period of illness, and data for how often the patients on average have such illness periods. By such a model we could analyze the cost-effectiveness of interventions that e.g. cure the shoulder pain once and for all, reduce the number of recurrences, shortens the average period of illness, or combinations of these outcomes.

3.1.2 Paper II: Modeling and Validating the Cost and Clinical Pathway of Colorectal Cancer

The aim: Paper II and its appendices comprise the main paper of this thesis, and the purpose of this paper was to contribute to modelling the colorectal cancer (CRC) cost and survival by presenting a transparent model and validating it.

Contributions: Paper II makes three contributions: (i) The paper develops and presents a general (multi-applicable) model for estimating CRC costs and survival. To the best of our knowledge, this is the first general CRC model to estimate both treatment costs and survival. (ii) The paper validates the model. (iii) The paper contributes to a more general understanding of validating models within health economic evaluations. ISPOR-SMDM has given recommendations for validating models (56), but to date, relatively few such models have been systematically validated. Kim and Thompson (59) write that: "*Health economic decision models are based on specific assumptions relating to model structure and parameter estimation. Validation of these models is recommended as an indicator of reliability, but is not commonly reported.*"

Methods regarding the model: We built a semi-Markov model with 70 health states and tracked age and time since specific health states (using tunnels and a three-dimensional data matrix). Instead of using a decision-tree, a Markov model was built because the timing of events was important (estimate survival time, discounting of cost, etc.), and important events can happen many times (recurrence of CRC). The model parameters are based on an observational study at Oslo University Hospital (OUS) (with 2 049 CRC patients), the National Patient Register, the literature and expert opinions. The model follows patients diagnosed with CRC from the age of 70 years until death or the age of 100 years. The health care payers' perspective is used.

The model is relatively complex, with 70 health states and many alternative paths. Each health state has its own economic model, and age and time since specific health states are tracked by using tunnels and a three-dimensional data matrix. A number of statistical analyses (e.g. survival analysis using Weibull) were performed, and separate computations were used to transform the output from the statistical analysis into input for the model. Weibull regression was used to estimate assumed changes in the transition probabilities as cycles elapsed after a certain event such as e.g. primary CRC treatment (92). Data from many different sources were used, some of which were assumed to be uncertain, such as expert opinions.

Methods regarding the validation: The structure and complexity of the model and the variety of data sources imply that we faced many types of uncertainty connected to parameters, methods and modelling. Therefore, we needed to perform a validation to determine whether we could trust the model. The model was validated for face, internal, cross and external validity.

Results from the validation: The validation revealed a satisfactory match with other models and with empirical estimates for both cost and survival time, without any preceding calibration of the model. In a cross validation, ten-year overall survival weighted for stage and estimated by the model, differed by 11.5 days (i.e. 0.38 months) compared with the OUS data. The cost of our model was 0,3% lower than the prediction by an Irish model (i.e. 3.0, 1.3, -3.6 and 1.2% lower in the Irish model for stages I, II, III and IV, respectively).

For external validation we compared relative survival estimated by the model with patients monitored by the Cancer Registry of Norway which contains a complete set of data for CRC patients in Norway. The model predicts 3.9% higher relative survival than national data during the first year, and 0.9, 5.6 and 5.6% lower relative survival five, ten and fifteen years after diagnosis, respectively.

In another external validation, the model was compared to empirically estimated ("model-free") total costs based on a Norwegian population study (National Patient Register). Taking into account difference in assumptions, the model estimate was 3.1% higher than the model-free estimate.

Discussion: Some of the elements that contribute to uncertainty in the results from the CRC model are: i) the cycles in the model were set to one year, which restricts the preciseness of the model to some extent; ii) because some of the OUS data used in the model are relatively old (range from 1993-2010), long-term survival is lower in the model, which can be explained by the older and less effective treatments; iii) in the model, we used a cohort of patients who were diagnosed at the age of 70 years, which may have resulted in a higher survival rate than if we had used the average age in the OUS sample; iv) the palliative sub-model was suitable for exploring treatment paths and costs, but there was no explicit built-in time dimension; therefore, an approximation was used to disperse the costs over time; and v) there appears to be a lack of data concerning the resource use related to treatment for local and distant recurrence, separately or combined, primarily because the relevant registers are not organized to estimate this parameter. The amount of uncertainty contributed by these elements, and possible solutions are discussed.

The production costs were not included in the estimation of CRC costs. For decisions makers this can be important to consider in order to get a complete estimate of CRC costs to society. For a person who is on sick leave for one year because of CRC, the production loss to society could be 2-3 times the costs of lifelong CRC treatment, according to the human capital methods. According to Yabroff et al. (93) the lack of including production cost analyses of colorectal cancer costs seems to be a widespread practice.

A limitation for the validation was that external validations can be applied to some components of the model or to the model as a whole (94), while our external validation was only applied to the model as a whole (survival and cost). Generally, the model complexity could be a drawback for decision makers to fully understand all the mechanisms of the model.

3.1.3 Paper III: Cost and survival of colorectal cancer and consequences of changing treatment algorithms: A model approach

The aims and contributions: While we in paper II presented and validated the CRC model, Paper III uses the model for estimations and makes three original contributions to the existing literature: (i) we demonstrate the usefulness of a new and recently developed and validated CRC model, (ii) we present results from

CRC cost and cost-effectiveness estimations, and (iii) we provide some insight into the uncertainty of such models and of CRC cost estimation in general.

Methods: We used different health economic evaluations like a COI analysis of colorectal cancer cost, cost or outcome description of certain elements in the analysis, a COA, and a kind of CEA. The cost description includes (i) cost analysis of different scenarios of palliative chemotherapy concerning changes in treatment and unit costs like cost of drug per dose, and (ii) the cost of certain CRC treatments and treatment related to past medical history. For a given progress in treatment (progress in surgery) and prevention (primary and secondary) we conducted COAs. To estimate the cost-effectiveness of providing treatment for colorectal cancer more generally, we performed a kind of CEA where cost and health outcomes for treated person were compared with the same for untreated persons with colorectal cancer – a treatment versus no-treatment CEA (see also section 3.5). All the analyses were based on the semi-Markov model from Paper II.

Results: The model was flexible and capable of modifying one-by-one or simultaneously many aspects of CRC treatment costs, such as prices, type and intensity of treatment and follow-up, recurrence rates, and CRC and non-CRC mortality. The cost for an average CRC patient was €41 550 (€23 390-€61 400, depending on the disease stage at diagnosis). A 20% cost change for purchasing palliative drugs had only a minor effect on the average CRC costs (<2%), while the altered use (who should receive therapy, the kind of therapy given, etc.) of palliative chemotherapy increased the cost by up to 29%. A 5% reduction in recurrence for stages I-III would reduce the health care cost by €2 280 per patient (5.5%) and an increase in the overall survival by 0.80 year per patient. Applying the suggested threshold for a QALY gained, the willingness-to-invest in a 5% reduction in recurrence rate would be €61 306 per CRC patient.

Uncertainty: One-way and multi-way (scenario) DSA were used extensively to explore the parameter, methodological and modelling uncertainty. The uncertainty in the costs result appeared to be more sensitive to future change in treatment than the uncertainty produced by the statistical estimation of parameters. Examples of important future changes include palliative chemotherapy treatment regimes and the introduction of general screening. The sensitivity to future change is especially important for preventive measures with a long latency period between the intervention and the expected health effects. One main weakness of the uncertainty analysis is the lack of a PSA for handling the parameter uncertainty.

Discussion: In addition to the limitations mentioned for Paper II above, the model did not account for: (i) health service costs like treatment provided/prescribed by general practitioners and care at nursing homes, (ii) costs generated by informal care done by relatives and friends, and (iii) productivity changes due to reduced productivity at work, sick leave and time used for treatment. Further, in some of the analyses there is a potential problem to not include public consumption not related to the CRC treatment (e.g. education or treatment for diseases not related to CRC) and private consumption, when someone lives

longer due to CRC treatment. This is as mentioned a subject for discussion (2, 14, 21). Thus we base the analysis on the current guidelines in Norway (20).

The main strength of the analyses performed in Paper III was that the model used has been thoroughly validated. A strength of the model used is its flexibility. The Markov model's estimates of health care cost for the four different stages of CRC presented in paper III, could be regarded as a COI analysis based on an incidence approach. Both the total cost reported, and the cost reported for different cost components, can be used in other evaluations like CEAs, CUAs and CBAs (which estimate the cost effectiveness of interventions that reduce the chance of getting CRC (screening, life style etc.)). However, in our case, the Markov model is general, and thus can be modified in order to perform the mentioned evaluations and a range of other CRC related analysis.

3.1.4 Paper IV: A health economic evaluation of screening and treatment in patients with adolescent idiopathic scoliosis

The aim: The aim was to compare the estimated costs in screening and non-screening scenarios in a CMA. *Contributions*: This paper made the following contributions: (i) To the best of our knowledge, few health economic evaluations have compared scoliosis screening with non-screening (61). The need for this type of evaluation is also indicated in an information statement by the scoliosis research society international task force (62): "…*there is scientific evidence to support the value of scoliosis screening with respect to technical efficacy, clinical, program and treatment effectiveness, but there is insufficient evidence to make a statement with respect to cost effectiveness*." (ii) Many relevant factors can be assumed to differ among countries, and the evaluation indicates which factors determine whether or not screening is cost-efficient. (iii) Closely related to point ii, the scenario analysis and tornado diagram based on the PSA indicate which factors are important for controlling the uncertainty. This result can guide future research to reduce the uncertainty of screening evaluations.

Methods: We used a cost minimization analysis, and assumed equivalent outcomes for health-related quality of life, and compared only relative costs in screening and non-screening settings. We included costs and administrative data from hospitals in combination with market prices to estimate costs in screening, bracing and surgical treatment. Screened children are treated (with bracing or surgery) more often than non-screened. This gap defines the non-screening scenarios, and we used reduced treatment rates of 90%, 80%, 70% for the non-screened compared to screened, and in addition the scenario using the actual treatment percent in Norway 2012 was called "non-screening Norway". The data were based on screening and treatment costs in primary health care and hospital care settings in Norway and Hong Kong. Out-of-pocket expenses and productivity losses among parents accompanying or caring for their child are also included. The incremental cost was defined as positive when a non-screening scenario was more expensive than screening. The analysis of uncertainty was based on the PSA, where all the parameters were given distributions. The PSA provided distributions and Credibility Intervals (CrIs) for the

incremental cost estimations in all scenarios. To determine which parameter contributes the most to the parameter uncertainty, we executed a tornado diagram analysis comparing a screening scenario with a non-screening scenario with an 80% treatment rate. We used a scenario analysis to explore the *modelling uncertainty* of changing parameter values that are dependent on medical practice or decisions varying between countries.

Results: The cost of screening per child was € 8.4 (95% CrI 6.6 to10.6), € 10 350 (8 690 to 12 180) per patient braced, and € 45 880 (39 040 to 55 400) per child operated. The screening was done in a relatively inexpensive way, performed by community nurses and physical therapists at the schools. The incremental cost per child in a non-screening scenario with a 90% treatment rate was € 13.3 (1 to 27), increasing from € 1.3 (-8 to 11) to € 27.6 (14 to 44) as surgical rates increased relative to bracing from 40% to 80%. For the 80% treatment rate non-screening scenario, the incremental cost was € 5.5 (-6 to 18) when screening all children and € 11.3 (2 to 22) when screening girls only. For the non-screening Norwegian scenario, the incremental cost per child was € 0.1(-14 to 16). The cost of surgery was dominating in the non-screening scenario, while the cost of bracing was dominating in the screening scenario. The economic gain of screening increases when the screening leads to higher rates of bracing and reduced surgical rates.

Discussion: A limitation of this paper could be the assumption that the prevalence of scoliosis is the same in Hong Kong and Norway. An important assumption made was that the screened and non-screened scenarios have the same health outcomes for the treated children. This was based on previous research and professional opinion, which was the reason for choosing CMA as the evaluation approach. A limitation for the PSA performed was that the probability distributions used were partially based on expert opinions, both from the medical expert in the group (part of the resource use data) and from the economists (with respect to the unit costs used). Another limitation can be that for all scenarios except "Non-screening Norway", by doing scenario analysis much of the uncertainty was not handled by the analysts but left (externalized) for the decision makers.

3.2 Comparing of the applications and the methods used in the four papers

The application of Paper I is to shoulder pain, Paper II and III to colorectal cancer and Paper IV to scoliosis (table 6).

Paper I was a COI analysis, while Paper IV was a CMA. Paper II was presenting and validating a model for doing COAs, and by some modest adjustment, doing HEEs (table 6). In Paper III the model for Paper II was used for doing different health economic evaluations like a COI analysis of colorectal cancer, cost or outcome description of certain elements in the analysis, COA, and a kind of CEA (see also section 3.3). The objective of Paper III was not to do health economic evaluations of certain new treatments, but

to provide to decision makers some relevant cost and survival estimates concerning CRC treatment that we argued could be relevant for future decisions on CRC research, treatment and prevention (72, 95).

Table 6 *summarizes the application of each paper, and method used and costs included.*

	Type of health economic evaluation	**Appli-cation**	**Costs included**	**Method for data collection and systematization**	**Handling of/testing for uncertainty**
Paper I	COI analysis	Shoulder pain	H. care cost Productivity	Patient-level-data	Variance-based CI One-/multi-way DSA
Paper II	Present and vali-date a CRC model (for doing COAs, FEEs, etc.)	Colo-rectal cancer	H. care cost	Decision analytic (using Markov and decision-three)	Face-, internal-, cross- and external validation
Paper III	COA (incl. COI analysis) and CEA	Colo-rectal cancer	H. care cost	Decision analytic (using Markov and decision-three)	Validated model One-/multiway DSA
Paper IV	CMA	Scoliosis	H. care cost Productivity Informal care	Decision analytic	PSA Scenario analysis

COI: cost of illness. COA: cost and outcome descriptions analysis. CEA: cost-effectiveness analysis. CMA: cost-minimization analysis. H. care cost: Health care cost. Decision analytic: decision analytical model. DSA: deterministic sensitivity analysis. PSA: probabilistic sensitivity analysis. FEE: full health economic evaluation.

For all papers, health care cost was included. Additionally, productivity loss caused by sick leave was included in the Paper I, and productivity loss (absence from work), support (transportation, company, etc) and informal care at home were included in paper IV.

Paper I was based on patient-level data, while the three other papers used decision analytic models. In papers II and III a semi-Markov model with included decision trees were used. A relatively comprehensive decision tree was developed for palliative chemotherapy, while more limited trees were used as part of the separate cost models for the primary treatments.

The COI analysis used in Paper I estimated the cost for 6 months for all persons with shoulder pain in a given population, whether the period of shoulder pain and the related treatment had ended or not. In contrast, the CMA (Paper IV) and the Markov model (Paper II and III) estimated the health care cost per average patients as long as the disease was expected to cause disease related costs. For the Markov model the time horizon was from 70 to 100 years old. For the CMA, the horizon was from the children were 11 to17 years.

Concerning the handling of uncertainty, Paper I had patient-level data, so we used the relevant variances to estimate the confidence intervals for the mean values. These intervals indicate the uncertainty caused by difference in treatment intensity and number of days of sick leave. Additionally, one- and multi-way DSA were used to test for uncertainty related to the unit cost. For each DSA the related confidence intervals were estimated based on the variability in treatment intensity and sick leave. For paper II we tested how well the model behaved by validating the model and test for face-, internal-, cross- and external validity (94). In Paper III we thus based our analyses on the validated model, and in addition we tested for how the results would change if the assumption like the treatment frequency, recurrence rate or cost of drug, would be changed in the future. In Paper IV we handled uncertainty by using a PSA.

3.3 Generalization of the results to other settings

Decision makers from one jurisdiction often need to conduct an evaluation which is similar to one already done in another jurisdiction. By exchanging the results of a particular assessment, the decision makers could avoid unnecessary duplication of efforts (24). This raises the issue of the potential for generalization of the data, which the NICE (64) define as "The extent to which the results of a study conducted in a particular patient population and/or a specific context will apply for another population and/or in a different context." The problems of generalization can be relevant both between countries, different regions inside a country (because e.g. different incidence rate of the illness), between different centers for treatment (which may have different cancer treatment practice in the use of expensive drug), and between subgroups (24). In this section we mainly discuss generalizability of results from one country to another.

3.3.1 Factors hampering the generalization and what to do about it

Some evaluations have results well adapted to a particular decision that has to be made, while in other situations, the decision maker only has access to evaluations performed in other settings that are somehow different from the one relevant for the decision maker. Often, it can be hard to assess if the analysis can be generalized to the relevant setting or not. The following differences between two settings can contribute to reduced generalizability (96): (i) Different demography and epidemiology of disease because

of difference between countries regarding lifestyle, age distribution, level of prosperity, etc. These can for example affect the baseline mortality, the incidence rate of the disease, and the physical and psychological response to a given treatment. Generally, *relative* treatment effectiveness is often assumed to be interchangeable across countries, while baseline event rates are not (24). (ii) Different levels of health care resources and clinical practice. Then, the same disease can be diagnosed, treated (different surgery or medications etc.) and followed up differently in two settings, which in turn can affect mortality (both background- and disease specific mortality), morbidity and treatment costs. (iii) Differences in incentives for health care professionals and institutions. (iv) Different relative prices or costs. For example, the relative unit cost for surgery, nurses or medications can differ as a result of different structure between countries regarding the relevant markets (97). Drummond et al. (24) argue that location-specific estimates would be required for cost and resources. (v) Population specific values. For evaluations like CUA and CBA, which include value judgments by the population, the values of the health outcomes can via QALY- or WTP-estimates, differ between countries. Consequently the total result (e.g. cost per QALY) can also differ. However, available evidence suggests little systematic variation in mean individual preferences between countries (24).

Sculpher et al. (98) reviewed generalizability in HEEs and provide recommendations both for HEEs based on patient-level-data and for evaluations using decision analytic modelling. In order to contribute to generalizability they recommended the following for decision analytic modelling: First, the evaluation should be clear about the decision problem, the decision-makers and the jurisdiction(s). Second, the analytical approach, model structure and data used should be appropriate to the relevant decision maker(s). Third, for parameters with several data sources, the data should be pooled so that uncertainty concerning precision and heterogeneity is reflected in the model (e.g. using standard meta-analysis). Fourth, "It is important to distinguish parameter uncertainty from variability or heterogeneity, where the latter is concerned with how parameter estimates vary across 'contexts' (98). Fifth, where data are incorporated as random variables, PSA is the appropriate means of handling parameter uncertainty. Sixth and finally, if targeting more than one jurisdiction, variability in results between locations should be assessed, for example by using SA or scenario analysis.

3.3.2 Generalization of the results in the four papers

Some of the challenges regarding generalization of the four papers are that Norway and Sweden are high-income countries, so the costs estimated could be unfit for generalization to medium or low-income countries. If the relevant relative prices are not disturbed, income level can normally be adjusted for by using purchasing power parity (PPP). However, also relative prices could be different, if for example the unit cost (wage rate) of physiotherapists, nurses or medical practitioner relative to other prices is different in Norway or Sweden compared to other countries, there will still be a problem after adjustment for PPP,

like Just et al. (97) found for economic evaluations of dialysis treatment modalities. Compared to middle- or low-income countries, we can expect labor to be relatively more costly in high-income countries like Norway and Sweden, while the differences could be smaller for resources (goods) traded in global markets like equipment and medicine.

For paper I, the countries we will generalize to should also have the same prevalence of the illness, which will depend on e.g. age distribution, type of industry (e.g. the load on shoulders is greater in some industries than others) and lifestyle. Scandinavian and some other Northern European countries appear to be similar to Sweden concerning these characteristics. In our multivariate regression analysis, we tested for the effect of age on total cost or health care cost, and found no significant effects. We also tested if there were cost differences due to the place (geographical) of treatment, but found no significant differences, which is positive regarding generalization within Sweden.

Papers II and III use the same CRC-model, so they are discussed together. The model is mainly based on Norwegian data, and then based on the characteristics of the Norwegian population like incidence rate for CRC, CRC-recurrence and -mortality, and background mortality. All these characteristics could differ between countries, but all of them can be adjusted for by the model. Further, we can expect the use of resources for diagnostics, surgery, chemotherapy, radiation and follow-up to differ between Norway and other countries (93). Also, the unit price can differ between countries, as we found in paper II when comparing the model with an Irish study modelling CRC treatment costs (36). In the CRC-model there are detailed and separate cost models for each health state; and particularly detailed for the year with primary treatment, the year after recurrence, and the years with palliative chemotherapy. Resource use, unit costs, and compliance can be adjusted for in these cost models. If we have data for the other settings, the flexibility of the model makes it relatively easy to adjust to fit a wide variety of settings, like other countries or subgroups of the population.

Because the analyses conducted in Paper III are based on Norwegian populations and treatment regimes, the results can probably best be generalized to Scandinavian and other North-European countries because of similarities in lifestyle, age distribution, treatment regimes, background mortality and incentives in the health sector. However, even among these countries we can experience important differences, for example because of differences regarding palliative chemotherapy which potentially can alter total costs in important ways (se paper III, Appendix 1).

For Paper IV, differences in demography and epidemiology of the disease can be a problem. One important assumption made is that the prevalence and natural history of scoliosis is the same in Hong Kong and in Norway. If our comparison between screening (Hong Kong) and "non-screening Norway" should be relevant to another country, the country has to be like Norway in the situation of non-screening, and like Hong Kong in the situation of screening. One argument for using the analysis from Hong Kong (99) is that it is the largest reported longitudinal study of screening cohorts. Whether or not it

is relevant for other countries to assume similarity with non-screening Norway, is debatable. Therefore we analyzed three more non-screening scenarios with different rate of treatment (surgery or bracing) in table 2 in paper IV, and also changed the ratio of brace/surgery in non-screening scenarios. We presented both expected incremental costs and CrI for all the combinations. Then the decision makers can use the estimates based on assumptions most relevant to their own country.

Also, the difference in screening practice between countries could be a problem for generalization. The assumed way of screening, based on a real intervention at Norwegian schools, had the cost of € 8.4 per child, which seems to be relatively inexpensive. Further, we see for the different scenarios that the incremental cost is € - 2.3 to 13.3 for all children, and € 4.3 to 18.4 for girls only. If for example the child on her way to and from screening, has to use separate transportation and/or be followed by a parents (time cost), then all scenarios - even for girls only – could result in a negative expected incremental cost. Screening would be even more expensive if the child visited a medical practitioner only for doing the screening.

For the decision maker, I will argue that the discussion above illustrates that the uncertainty of generalizing results from one setting to another, can often be hard to handle. In the papers deterministic one-way and multi-way SA (paper I and III) and probabilistic scenario SA (paper IV) have been used to get a better picture of this uncertainty. Instead of doing SAs of the evaluations done for country A, so decision makers better can generalize the result from country A to country B, one could instead use the same model and change all relevant parameters to fit country B better. This would be a far more informative solution, but assumes that the model and all input parameters are available to those who would use it for country B. In paper II (the CRC model) and paper IV (scoliosis model) we presented the models (conceptually and mathematically) and the input parameters, so detailed that we hoped the models were possible to reproduce and use in other settings. However, our experience so far is that (particularly for the CRC model) it is probably a need for even more detailed descriptions if the simulation models should be rebuilt by others. A possible solution to this kind of problems could to be to publish the whole simulation model together with a manual, but there is no tradition for this among health economists, partly because of the problems with ownership and incentives to develop new simulation models (94, 100).

4. Conclusions

4.1 Contributions

The contributions of the thesis are: (i) it provides economic evaluation of screening for scoliosis; and estimates the cost of shoulder pain in Sweden, the cost of colorectal cancer, and the potential gain to the Norwegian society by reducing the recurrence rate for CRC (e.g. by increasing the quality of CRC surgery); (ii) it provides development of the first general (multi applicative) simulation model for colorectal cancer, which estimates both treatment cost and survival time; (iii) it provides an example of face-, internal-, cross-over- and external validation (few validations are so far conducted within health economics); and (iv) it provides a discussion of the strengths and weaknesses of using Markov models, COI and cost-minimization analyses.

4.2 Key conclusions

In Paper I, a cost-of-illness (COI) study on shoulder pain in Sweden, showed that the mean health care cost per patient was €326 during 6 months, and physiotherapy treatments accounted for 60% of this cost. The mean annual total cost was €4 139 per patient. Of this, sick leave accounted for 84% of the cost, but different methods for estimating sick leave cost can provide very different results.

In Paper II a semi-Markov model with 70 health states was presented and validated. We tracked age and time since specific health states using tunnels and a three-dimensional data matrix. The structure and complexity of the model and the variety of data sources implied that we faced parameter and methodological uncertainty, as well as modelling uncertainty. Therefore, the model was validated using face, internal, cross and external validation. The main result from Paper II was the validation, and this revealed a satisfactory match with other models and empirical estimates of both the cost of colorectal cancer treatment and survival time, which are the two main outcomes of the model. We performed no preceding calibration of the model.

In Paper III, we found that altered decisions about palliative treatment can increase the average CRC cost substantially. Reducing the recurrence rate by better surgery and implementing preventive efforts like screening of asymptomatic persons could have a considerable cost-effectiveness potential. Further, we saw that expectations about the future are important for cost and survival estimates. Because many evaluations have time horizons of 20-40 years, PSA that is based on parameter probability distributions estimated from "yesterday's data" can be misleading.
In Paper IV, we compare costs in screening and non-screening scenarios using a cost-minimization analysis. Many relevant factors can be assumed to differ from country to country. We found that the cost-effectiveness of screening is heavily dependent on (i) the percent of the non-screened that receive some

kind of treatment (surgery or bracing) for their scoliosis, and (ii) the share of surgery versus bracing, in both screened and non-screened children. We also found that it is more cost-effective to screen girls only rather than screening all children.

4.3 Policy implications

From Paper I, we saw that production loss, via sick leave, accounted for most of the total health care cost caused by shoulder pain. Then, it can be a problem for the decision makers that sick leave is often a cost that is excluded from COI analyses and FEEs, and when included, different relevant methods for estimating production loss provide very different results. Further we found that it is a relatively small group of patients that contribute to most of the cost to society, and this is particularly related to production losses.

Two policy implications from paper II and III could be mentioned: (i) society could gain from more research on how to reduce the CRC recurrence rate, and (ii) to attain credible cost-effectiveness analysis of CRC treatment interventions, there is often a need for flexible general models with the ability to include expectations about future prices, resource use, recurrence rates, background mortality etc., and to compare different categories of CRC interventions, e.g. chemotherapy versus screening or new surgery techniques. The health sector often simultaneously assesses different types of treatments for the same disease in order to identify the right mix of treatments. Then, the evaluations depend on the use of general (i.e., multi-applicative) models that can estimate the aggregate effects of many different treatments. Nevertheless, to work with specialized models have been the practice so far.

Paper IV: To policy makers it could be important that it is far more cost-effective to screen only girls than all children, screening will hardly be cost-effective without a rather effective and inexpensive screening procedure, and the result from the CMA was heavily depending on the type of treatment (surgery or bracing) received by the screened and non-screened.

For all tree evaluations presented in this thesis we saw that modelling or methodological uncertainly was considerable. In Paper I the choice between using friction or human capital methods gave alternative predictions outside the confidence interval for the base case alternative. In paper II future decisions about palliative treatment were important for the total CRC treatment cost, and for Paper IV questionable fundamental assumptions were important for the whole model. This kind of uncertainty is important to consider when decision makers assess health economic evaluations which often use PSA based solely on parameter uncertainty.

References

1. Gray A, Clarke PM, Wolstenholme JL, Wordsworth S. Applied methods of cost-effectineness analysis in health care. Oxford: Oxford University Press; 2011. 313 p.
2. Drummond M, Sculpher M, Torrance G, O'Brien B, Stoddart G, editors. Methods for the Economic Evaluation of Health Care Programmes. Third ed: Oxford University Press; 2005.
3. Gold MR, Siegel JE, Russel LB, Weinstain MC. Cost-Effectiveness in Health and Medicine. Oxford: Oxford University Press; 1996.
4. Mcintosh E, Clarke P, Frew E, Louviere J. Applied methods of cot-benefit analysis in health care. Oxford: Oxford University Press; 2010.
5. Mishan E, Quah E. Cost-benefit analysis. New York: Routledge; 1976.
6. Briggs AH, O'Brien BJ. The death of cost-minimization analysis? Health Economics. 2001;10(2):179-84.
7. Nord E. Cost-value analysis in health care. Cambridge: Cambridge University Press; 1999.
8. Hunink M, Glasziou P, Siegel J, Weeks J, Pliskin J, Elstein A, et al. Decision making in health and medicine - integrating evidence and values. New York: Cambridge University Press; 2001.
9. Glick HA, Doshi JA, Sonnad SS, Polsky D. Economic Evaluation in Clinical Trials. Oxford: Oxfors University Press; 2007.
10. Bowling A. Measuring disease. second ed. Philadelphia: Open University Press; 2001.
11. Rice DP. Estimating the cost of illness. American Journal of Public Health and the Nations Health. 1967;57(3):424-40.
12. Tarricone R. Cost-of-illness analysis. Health Policy.77(1):51-63.
13. Hodgson TA, Meiners MR. Cost-of-Illness Methodology: A Guide to Current Practices and Procedures. The Milbank Memorial Fund Quarterly/Health and Society. 1982;60(3):34.
14. Brouwer W, Rutten F, Koopmanschap M. Costing in economic evaluations. In: Drummond M, McGuire A, editors. Economic evaluation in health care. Oxford: Oxford University Press; 2001. p. 286.
15. Wordsworth S, Ludbrook A, Caskey F, Macleod A. Collecting unit cost data in multicentre studies. The European Journal of Health Economics. 2005;6(1):38-44.
16. Brouwers W, Rutten F, Koopmanschap M. Costing in economic evaluations. In: Drummond M, Alistair G, editors. Economic evauation in health care - merging theory with practice. Oxford: Oxford University Press; 2001. p. 68-93.
17. Koopmanschap MA, Rutten FF, van Ineveld BM, van Roijen L. The friction cost method for measuring indirect costs of disease. J Health Econ. 1995;14(2):171-89.
18. Sculpher M. The role and estimation of productivity costs in economic evaluation. In: Drummond M, McGuire A, editors. Economic evaluation in health care - merging theory with practice. Oxford: Oxford University Press; 2001. p. 94-112.
19. Shiell A, Gerard K, Donaldson C. Cost of illness studies: An aid to decision-making? Health Policy.8(3):317-23.
20. Helsedirektoratet. Economic evaluation of health intervention - a guide. Oslo: The Norwegian Directorate of Health; 2012.
21. Meltzer D. Accounting for future costs in medical cost-effectiveness analysis. Journal of Health Economics. 1997;16(1):33-64.

22. Kuntz K, Weinstein M. Modelling in economic evaluation. In: Drummond M, McGuire A, editors. Economic evauation in health care - merging theory with practice: Oxford University Press; 2001. p. 31.
23. Sculpher MJ, Claxton K, Drummond M, McCabe C. Whither trial-based economic evaluation for health care decision making? Health Economics. 2006;15(7):677-87.
24. Drummond M, Manca A, Sculpher M. Increasing the generalizability of economic evaluations: Recommendations for the design, analysis, and reporting of studies. International Journal of Technology Assessment in Health Care. 2005;21(02):165-71.
25. Briggs AH, Claxton K, Sculpher M. Decision modelling for health economic evaluation. Oxford: Oxford University Press; 2006.
26. Bland JM, Altman DG. Bayesians and frequentists. BMJ. 1998;317(7166):1151-60.
27. Briggs AH. A BAYESIAN APPROACH TO STOCHASTIC COST-EFFECTIVENESS ANALYSIS. International Journal of Technology Assessment in Health Care. 2001;17(01):69-82.
28. Karnon J. Alternative decision modelling techniques for the evaluation of health care technologies: Markov processes versus discrete event simulation. Health Economics. 2003;12(10):837-48.
29. Barton P, Bryan S, Robinson S. Modelling in the economic evaluation of health care: selecting the appropriate approach. J Health Serv Res Policy. 2004;9(2):110-8.
30. Cooper K, Brailsford SC, Davies R. Choice of modelling technique for evaluating health care interventions. J Oper Res Soc. 2007(2):168 - 76.
31. Griffin S, Claxton K, Hawkins N, Sculpher M. Probabilistic Analysis and Computationally Expensive Models: Necessary and Required? Value in Health. 2006;9(4):244-52.
32. Weinstein MC. Recent Developments in Decision-Analytic Modelling for Economic Evaluation. PharmacoEconomics. 2006;24:1043-53.
33. Drummond MF, Schwartz JS, Jönsson B, Luce BR, Neumann PJ, Siebert U, et al. Key principles for the improved conduct of health technology assessments for resource allocation decisions. International Journal of Technology Assessment in Health Care. 2008;24(03):244-58.
34. Sonnenberg FA, Beck JR. Markov Models in Medical Decision Making: A Practical Guide. Medical Decision Making. 1993;13(4):322-38.
35. Chiang AC, Wainwright K. Fundamental methods of mathematical economics. Fourth ed: Mc Graw Hill; 2005.
36. Joranger P, Nesbakken A, Hoff G, Sorbye H, Oshaug A, Aas E. Modeling and validating the cost and clinical pathway of colorectal cancer. Medical Decision Making. 2014.
37. Naimark DMJ, Bott M, Krahn M. The Half-Cycle Correction Explained: Two Alternative Pedagogical Approaches. Medical Decision Making. 2008.
38. Naimark DMJ, Kabboul NN, Krahn MD. The Half-Cycle Correction Revisited: Redemption of a Kludge. Medical Decision Making. 2013;33(7):961-70.
39. Krupnick A, Morgenstern R, Batz M, Nelson P, Burtraw D, Shih J-S, et al. Not a sure thing: making regulatory choices under uncertainty. Washington: Resources For the Future; 2006.
40. Claxton K. Exploring Uncertainty in Cost-Effectiveness Analysis. PharmacoEconomics. 2008;26(9):781-98.
41. Drummond M, McGuire A, editors. Economic Evaluation in Health Care - merging theory with practice. Oxford: Oxford University Press; 2001.
42. Briggs A. Handling Uncertainty in Cost-Effectiveness Models. PharmacoEconomics. 2000;17(5):479-500.

43. Briggs A. Handling uncertainty in economic evaluation and presenting the results. In: Drummond M, McGuire A, editors. Economic evaluation in health care. Oxford: Oxford University Press; 2001. p. 172-214.
44. Briggs A, Sculpher M, Buxton M. Uncertainty in the economic evaluation of health care technologies: The role of sensitivity analysis. Health Economics. 1994;3(2):95-104.
45. StatensLegemiddelverk. Guidelines for pharmacoeconomic analyzes. Oslo: 2012.
46. Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the Panel on Cost-Effectiveness in Health and Medicine. JAMA: The Journal of the American Medical Association. 1996;276(15):1253-8.
47. Bojke L, Claxton K, Sculpher M, Palmer S. Characterizing Structural Uncertainty in Decision Analytic Models: A Review and Application of Methods. Value in Health. 2009;12(5):739-49.
48. Jackson CH, Bojke L, Thompson SG, Claxton K, Sharples LD. A Framework for Addressing Structural Uncertainty in Decision Models. Medical Decision Making. 2011;31(4):662-74.
49. Bloom B, Bruno D, Maman D, Jayadevappa R. Usefulness of US Cost-of-Illness Studies in Healthcare Decision Making. PharmacoEconomics. 2001;19(2):207-13.
50. Rice DP. Cost of illness studies: what is good about them? Injury Prevention. 2000;6(3):177-9.
51. Byford S, Torgerson DJ, Raftery J. Cost of illness studies. BMJ : British Medical Journal. 2000;320(7245):1335-.
52. Currie G, Kerfoot KD, Donaldson C, Macarthur C. Are cost of injury studies useful? Injury Prevention. 2000;6(3):175-6.
53. Berto P, D'Ilario D, Ruffo P, Virgilio RD, Rizzo F. Depression: cost-of-illness studies in the international literature, a review. The Journal of Mental Health Policy and Economics. 2000;3(1):3-10.
54. Segel JE. Cost-of-illness - A primer. RTI-UNC Center of Excellence in Health Promotion Economics: RTI International 2006.
55. Sharp L, Tilson L, Whyte S, O'Ceilleachair A, Walsh C, Usher C, et al. Cost-effectiveness of population-based screening for colorectal cancer: a comparison of guaiac-based faecal occult blood testing, faecal immunochemical testing and flexible sigmoidoscopy. Br J Cancer. 2012;106(5):805-16.
56. Tappenden P, Chilcott J, Eggington S, Sakai H, Karnon J, Patnick J. Option appraisal of population-based colorectal cancer screening programmes in England. Gut. 2007;56(5):677-84.
57. Frazier AL, Colditz GA, Fuchs CS, Kuntz KM. Cost-effectiveness of screening for colorectal cancer in the general population. JAMA: The Journal of the American Medical Association. 2000;284(15):1954-61.
58. Tilson L, Sharp L, Usher C, Walsh C, S W, O'Ceilleachair A, et al. Cost of care for colorectal cancer in Ireland: a health care payer perspective. The European Journal of Health Economics. 2012;13(4):511-24.
59. Taplin SH, Barlow W, Urban N, Mandelson MT, Timlin DJ, Ichikawa L, et al. Stage, Age, Comorbidity, and Direct Costs of Colon, Prostate, and Breast Cancer Care. Journal of the National Cancer Institute. 1995;87(6):417-26.
60. Altman DG, Bland JM. Statistics notes: Absence of evidence is not evidence of absence1995 1995-08-19 07:00:00. 485 p.
61. Drummond M, O'Brien B, Stoddart G, Torrance G. Methods for the economic evaluation of health care programmes. Second ed1997.
62. Ross J. The use of economic evaluation in health care: Australian decision makers' perceptions. Health Policy. 1995;31(2):103-10.

63. Claxton K, Sculpher M, Drummond M. A rational framework for decision making by the National Institute For Clinical Excellence (NICE). The Lancet. 2002;360(9334):711-5.
64. NICE. Guide to the methods of technology appraisal 2013. London: National Institute for health and Care Excellence, 2013.
65. Claxton K, Sculpher M, Palmer S, Culyer AJ. CAUSES FOR CONCERN: IS NICE FAILING TO UPHOLD ITS RESPONSIBILITIES TO ALL NHS PATIENTS? Health Economics. 2015;24(1):1-7.
66. Claxton K, Martin S, Soares M, Rice N, al. e. Methods for the Estimation of the NICE Cost Effectiveness Threshold. York: University of York, 2013.
67. Norheim O, Allgott B, Aschim B, Førde R, al. e. Open and fair - priorities in health. Oslo: Ministry of Health and Care Services, 2014.
68. Bateman I, Carson R, Hanemann M, Hanley N, al e. Economic valuation with stated preference techniques - a manual. Transport Df, editor. Cheltenham, UK: Edward Elgar; 2002.
69. Drummond M. Health Technology Assessment and Its Interface with Regulation, Policy and Management. In: del Llano-Señarís JE, Campillo-Artero C, editors. Health Technology Assessment and Health Policy Today: A Multifaceted View of their Unstable Crossroads: Springer International Publishing; 2015. p. 3-14.
70. NICE. Guide to the methods of technology appraisal. London: National Institute for Clinical Excellence, 2004.
71. Simoens S. Use of Economic Evaluation in Decision Making. Drugs. 2010;70(15):1917-26.
72. Williams I, Bryan S. Understanding the limited impact of economic evaluation in health care resource allocation: A conceptual framework. Health Policy. 2007;80(1):135-43.
73. Luce BR, Drummond M, JÖNsson B, Neumann PJ, Schwartz JS, Siebert UWE, et al. EBM, HTA, and CER: Clearing the Confusion. Milbank Quarterly. 2010;88(2):256-76.
74. Squire L, Van DerTak H. Economic analysis of projects. Bank W, editor. London: The Johns Hopkins University Press; 1989.
75. Tsuchiya A, Williams A. Walfare economics and economic evaluation. In: Drummond M, McGuire A, editors. Economic evaluation in health care - merging theory with practice. Oxford: Oxford University Press; 2001. p. 21.
76. Sugden R, Williams A. The principles of practical cost-benefit analysis. Oxford: Oxford University Press; 1978.
77. Hoffmann C. The influence of economic evaluation studies on decision making. Health Policy.52(3):179-92.
78. Bryan S, Williams I, McIver S. Seeing the NICE side of cost-effectiveness analysis: a qualitative investigation of the use of CEA in NICE technology appraisals. Health Economics. 2007;16(2):179-93.
79. Husereau D, Culyer A, Neumann P, Jacobs P. How do Economic Evaluations Inform Health Policy Decisions for Treatment and Prevention in Canada and the United States? Appl Health Econ Health Policy. 2015;13(3):273-9.
80. Dakin H, Devlin N, Feng Y, Rice N, O'Neill P, Parkin D. THE INFLUENCE OF COST-EFFECTIVENESS AND OTHER FACTORS ON NICE DECISIONS. Health Economics. 2014:n/a-n/a.
81. Grepstad M, Kanavos P. A comparative analysis of coverage decisions for outpatient pharmaceuticals: Evidence from Denmark, Norway and Sweden. Health Policy.119(2):203-11.
82. Tina Shih Y-C, Mullins Daniel C, Drummond M. The Role of Economic Evaluation in Meeting IOM's Recommendations on Delivering High-Quality Cancer Care. Value in Health. 2014;17(5):497-500.

83. Chambers JD, Neumann PJ, Buxton MJ. Does Medicare Have an Implicit Cost-Effectiveness Threshold? Medical Decision Making. 2010;30(4):E14-E27.
84. Mason A, Drummond M, Ramsey S, Campbell J, Raisch D. Comparison of Anticancer Drug Coverage Decisions in the United States and United Kingdom: Does the Evidence Support the Rhetoric? Journal of Clinical Oncology. 2010;28(20):3234-8.
85. Innvær S, Vist G, Trommald M, Oxman A. Health policy-makers' perceptions of their use of evidence: a systematic review. Journal of Health Services Research & Policy. 2002;7(4):239-44.
86. Lessard C, Contandriopoulos A-P, Beaulieu M-D. The role (or not) of economic evaluation at the micro level: Can Bourdieu's theory provide a way forward for clinical decision-making? Social Science & Medicine. 2010;70(12):1948-56.
87. Eddama O, Coast J. Use of economic evaluation in local health care decision-making in England: A qualitative investigation. Health Policy.89(3):261-70.
88. Brousselle A, Lessard C. Economic evaluation to inform health care decision-making: Promise, pitfalls and a proposal for an alternative path. Social Science & Medicine. 2011;72(6):832-9.
89. Hoffmann C, Stoykova BA, Nixon J, Glanville JM, Misso K, Drummond MF. Do Health-Care Decision Makers Find Economic Evaluations Useful? The Findings of Focus Group Research in UK Health Authorities. Value in Health. 2002;5(2):71-8.
90. Bell CM, Urbach DR, Ray JG, Bayoumi A, Rosen AB, Greenberg D, et al. Bias in published cost effectiveness studies: systematic review2006 2006-03-23 22:46:51. 699-703 p.
91. Miners AH, Garau M, Fidan D, Fischer AJ. Comparing estimates of cost effectiveness submitted to the National Institute for Clinical Excellence (NICE) by different organisations: retrospective study2005 2005-01-06 22:55:15. 65 p.
92. Kleinbaum D, Klein M. Survival analysis. Second ed. New York: Springer; 2005.
93. Yabroff KR, Borowski L, Lipscomb J. Economic studies in colorectal cancer: Challenges in measuring and comparing costs. JNCI Monographs. 2013;2013(46):62-78.
94. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB. Model Transparency and Validation: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force–7. Medical Decision Making. 2012;32(5):733-43.
95. Buxton MJ, Drummond MF, Van Hout BA, Prince RL, Sheldon TA, Szucs T, et al. Modelling in Ecomomic Evaluation: An Unavoidable Fact of Life. Health Economics. 1997;6(3):217-27.
96. Drummond M, Pang F. Transferability of economic evaluation results. In: Drummon M, MCcGuire A, editors. Economic evaluation in health care - merging theory with practice. Oxford: Oxford University Press; 2001. p. 256-76.
97. Just PM, Riella MC, Tschosik EA, Noe LL, Bhattacharyya SK, de Charro F. Economic evaluations of dialysis treatment modalities. Health Policy. 2008;86(2–3):163-80.
98. Sculpher M, Pang F, Manca A, Drummond M, Golder S, Urdahl H. Generalisability in economic evaluation studies in healthcare: a review and case studies. Health Technology Assessment. 2004;8(49):206.
99. Lee CF, Fong DYT, Cheung KMC, Cheng JCY, Ng BKW, Lam TP, et al. Costs of School Scoliosis Screening: A Large, Population-Based Study. Spine. 2010;35(26):2266-72 10.1097/BRS.0b013e3181cbcc10.
100. Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C, et al. Principles of Good Practice for Decision Analytic Modeling in Health-Care Evaluation: Report of the ISPOR Task Force on Good Research Practices—Modeling Studies. Value in Health. 2003;6(1):9-17.

Part II: The papers

Paper I

Virta et al. BMC Musculoskeletal Disorders 2012, 13:17
http://www.biomedcentral.com/1471-2474/13/17

BMC Musculoskeletal Disorders

RESEARCH ARTICLE Open Access

Costs of shoulder pain and resource use in primary health care: a cost-of-illness study in Sweden

Lena Virta[1*], Pål Joranger[2], Jens Ivar Brox[3] and Rikard Eriksson[2]

Abstract

Background: Painful shoulders pose a substantial socioeconomic burden. A prospective cost-of-illness study was performed to assess the costs associated with healthcare use and loss of productivity in patients with shoulder pain in primary health care in Sweden.

Methods: The study was performed in western Sweden, in a region with 24 000 inhabitants. Data were collected during six months from electronic patient records at three primary healthcare centres in two municipalities. All patients between 20 and 64 years of age who presented with shoulder pain to a general practitioner or a physiotherapist were included. Diagnostic codes were used for selection, and the cases were manually controlled. The cost for sick leave was calculated according to the human capital approach. Sensitivity analysis was used to explore uncertainty in various factors used in the model.

Results: 204 (103 women) patients, mean age 48 (SD 11) years, were registered. Half of the cases were closed within six weeks, whereas 32 patients (16%) remained in the system for more than six months. A fifth of the patients were responsible for 91% of the total costs, and for 44% of the healthcare costs. The mean healthcare cost per patient was €326 (SD 389) during six months. Physiotherapy treatments accounted for 60%. The costs for sick leave contributed to 84% of the total costs. The mean annual total cost was €4139 per patient. Estimated costs for secondary care increased the total costs by one third.

Conclusions: The model applied in this study provides valuable information that can be used in cost evaluations. Costs for secondary care and particularly for sick leave have a major influence on total costs and interventions that can reduce long periods of sick leave are warranted.

Background

Shoulder pain is a common cause of lost work days and disability. A majority of the patients are treated in primary health care [1-3]. In Sweden, health and medical care are organised in three levels: regional medical care, county medical care, and primary care which is organised by the county councils. Primary care is intended to meet the needs of most patients for medical treatment, care, preventive measures and rehabilitation. When more specialised care is necessary, patients are referred to the county hospitals. The regional hospitals treat rare and complicated cases. There were very few private care providers in the county at the time of this study. Resources are scarce, and the Swedish Health and Medical Services Act states that priority should be given to those who are in the greatest need of health and medical care. Quality of care can be defined as a combination of structure, process, and outcome [4]. Cost-of-illness studies can provide information about healthcare resources and costs allocated to different groups of patients.

Net costs to healthcare authorities for health and medical care in Sweden in 2005 were 16% for primary care and 52% for specialised physical care [5], most of which is financed from tax revenues. There is a government-imposed patient's cost ceiling for health care, meaning that no patient needs to pay more than €100 during a 12-month period, and no patient needs to pay more than €200 for prescription drugs covered by the benefits.

About 6,500 shoulders were operatively treated in Sweden in 2004 [6], and since 1998 the number of shoulder surgeries has increased by about 10% annually. A recent study reported a four-fold increase in the number of acromioplasties for rotator cuff disorders in New York

* Correspondence: lena.virta@vgregion.se
[1]Medical Faculty, University of Oslo, Oslo, Norway
Full list of author information is available at the end of the article

State from 1996 to 2006 [7]. Multifactorial reasons were suggested for this increase, with patient-based, surgeon-based, and systems-based factors all playing a role. The differential diagnoses for shoulder pain are based on the history, acute or chronic nature of the pain, physical examination, and, if needed, completed with imaging. Tests for diagnostic accuracy [8] as well as surgical indications, are being discussed [1,9,10]. Although evidence from case series supports the effectiveness of surgical interventions for shoulder pain when used appropriately [1], the increase in shoulder surgery cannot be explained by the practice of evidence-based medicine. Three randomised clinical trials [11-13] comparing supervised exercises for subacromial pain with surgery, have concluded that supervised exercises are equally effective as surgery - and less expensive. One additional study found that only 10% of the patients awaiting surgery were finally operated on after being treated with physiotherapist-supervised exercises in a hospital setting [14]. This indicates a need for economic evaluations of current treatment strategies in primary health care.

The initial steps taken to diagnose and treat the patient in primary care may be essential for effective treatment, and may contribute to fewer patients being referred to surgery as well as lower costs to society. Kuijpers et al [15] reported costs of shoulder pain in primary care patients who presented with shoulder pain to their general practitioner (GP) in the Netherlands in 2006. Patients were followed for six months and their shoulder pain related costs were calculated by using patients' cost diaries. The patients reported all expenses relevant to their shoulder complaints; direct costs, such as visits to healthcare centres, and indirect costs, such as sick leave, and paid and unpaid help. In their study, 70% had persistent symptoms after six weeks and 46% after six months. They found that 12% of the patients with shoulder pain were responsible for 74% of the total costs, mostly a result of sick leave from paid work. Our study was performed to investigate the situation in Swedish primary health care, using an alternative design.

In Sweden, electronic patient records (EPR) based on diagnostic codes are used mainly in the clinical care of patients and rarely to evaluate healthcare programmes or cost-effectiveness aspects. Completeness and accuracy of diagnostic codes have been found acceptable [16,17], in spite of a coding system poorly adapted to primary health care. Attempts have been made, using EPR, to monitor the burden of illness for patients with low back pain [18], diabetes [19], and groups of patients according to their health status [20]. Linking costs and consequences based on already collected patient data may be useful to monitor the cost of illness in selected groups of patients.

The aim of this study was to assess the costs associated with healthcare use and loss of productivity caused by shoulder pain in Sweden, by auditing data from the EPR.

Questions asked in the study:

- What are the shoulder pain related treatment costs in primary care consulters in Sweden (direct costs)?
- What are the costs of shoulder pain in defined subgroups of the selected population (highest costs)?
- What are the costs for sick leave (indirect costs)?
- What are the total costs?

Methods

Setting

The study was performed in 2009 in two municipalities, comprising 24 000 inhabitants, in a prosperous region on the Swedish west coast. The labour market in this region is based on trade and tourism, as well as many small and medium-sized enterprises. Three primary healthcare centres with three adjacent physiotherapy units were responsible for almost all primary health care in the area. There were few private alternatives to physiotherapy and no private physicians, making it possible to capture almost all patients who presented with shoulder pain in primary health care. In western Sweden, patients do not need a referral for physiotherapy. Sick leave for more than eight days must be prescribed by a doctor, although some employers require this from day one. The inclusion of patients was based on EPR in primary health care. We included all patients that presented with shoulder pain to any of these six units during the measurement period of six months, regardless of trauma or other diseases. Patients being permanent residents in either of the two municipalities and between 20 and 64 years of age were included, if any of the diagnostic codes given at the visit qualified them.

Costing

A prospective cost-of-illness study was performed to explore the most important cost components of treating shoulder pain in primary health care. Healthcare costs and total costs, including cost for sick leave, were assessed. Costing involves identifying, measuring and valuing all resource changes that occur as certain healthcare interventions are carried out. In a bottom-up approach, individual elements are specified in detail. The three steps of the costing procedure in this study were:

1. Identification of relevant cost-items
2. Quantification of the use of the identified cost-items
3. Valuing the identified items

Electronic patient records (EPR)

With very few exceptions, all units in primary health care in Sweden are computerised, and several EPR systems are in use. The data collected from the EPR were organized in a data matrix containing patients' personal

identity number, age, sex, dates of encounter and diagnostic codes for every visit, number of admissions and referrals to specialist care, x-rays, number of drug prescriptions and sick leave periods prescribed by a GP. Our first step was to retrieve all visits to general practitioner or physiotherapist (PT) caused by shoulder pain during the measurement period. All data were anonymised before analysis.

At all participating units, notice boards were used to inform the patients. Receptionists were also asked to leave information sheets to patients who sought treatment for shoulder pain. All inhabitants in the area had been told that information from their EPR could be accessed and processed without consent for planning and quality assurance. The procedures of this study were approved by the Regional Ethical Review Board of Gothenburg.

Management of shoulder pain

The Swedish guidelines for the management of shoulder pain [6] are similar to those for GPs in other countries [15,21-23]. Conservative (non-operative) care is recommended, including information on the prognosis of shoulder pain and advice regarding physical activities. In addition, the guidelines recommend a step-by-step treatment progression, consisting of physiotherapy treatment, pain relief and glucocorticoid injections (administered with or without local anaesthetic). If conservative treatment fails to reduce the symptoms, the patient is referred to an orthopaedic surgeon. In the present study physiotherapy treatments were adapted to each patient's condition and supervised exercises were emphasised.

The local hospital has a radiology department providing ultrasound evaluation of suspected tendon ruptures. MRI is regarded as a tool for orthopaedic surgeons and is seldom used in primary care in this part of Sweden.

The diagnostic coding system International Classification of Disease, version 10 (ICD-10), was used. Initially, a pilot study was performed at all six participating units to find out which diagnostic codes that were used for patients who consulted for shoulder pain. Fractures and dislocations of the shoulder were included. All visits with known and potential codes for shoulder pain were retrieved from the EPR system. Each individual with a potential code was scrutinized by comparing data within the EPR to verify the cause of visit. In the last step, 29 codes were classified in four categories, presented in Table 1: subacromial pain (including nonspecific shoulder pain), stiffness (adhesive capsulitis, arthritis), dislocations, and fractures.

Procedure

The cost-of-illness calculation was based on all registered actions related to shoulder pain during the measured period. Patients referred to orthopaedic surgeon for evaluation were followed up to monitor whether they were selected for surgery or not.

Total treatment time and sick leave at inclusion were retrieved from the EPR. The period between first and last dates of visit to a GP or a PT with the qualifying code was defined as the total treatment time. At least one visit per month had to be registered, except during the holiday period.

Half of the patients started and ended treatment within six months. Some started before and some ended after the measured period. We believe that this would be the case at any chosen period during the year. Costs for all patients passing through during six months can be multiplied by 2 to estimate the annual cost for this group of patients. This estimate can then be used to compare with annual costs in other regions. This method is also suitable to investigate the relative size of the different treatment components.

Calculation of treatment costs per patient requires complete registration of all activities during the whole treatment period. Patients must be monitored from their first encounter for shoulder pain, although onset may be difficult to define. They should preferably be monitored for a long time period, ideally for the rest of their lives.

Table 1 Diagnostic codes.

Subgroups	Diagnostic codes	Patients N (%)	Age (years) Mean (SD) Median	Sex: Male N (%)	Had Surgery N (%)
Subacromial pain	M751-9, M759P, M709, M779, M791, M799, M255, M255B, M629, M795, M796B	181(89)	48 (11) 51	89(49)	9 (48)
Stiffness	M750, M190B, M192B	10 (5)	52 (10) 54	7 (70)	2 (10)
Fractures	S420, S4200, S429	7 (3)	48 (14) 52	3 (42)	6 (32)
Dislocations	S430, S431, S435, S460	6 (3)	51 (13) 55	2 (33)	2 (10)

According to International Classification of Diseases, version 10 (ICD-10), used for shoulder pain and merged into four categories. Code names are presented below. All patients, N = 204. Patients who had surgery or other orthopaedic intervention, N = 19

M75. Shoulder lesions; M750 Adhesiv capsulitis

M70-M79 (B) Other soft tissue disorders (shoulder); M629 Disorder of muscle

M255 (B) Pain in joint (shoulder); M19.(B) Arthrosis (shoulder)

Valuing healthcare costs

Costs used in the economic evaluation are presented in Table 2. Healthcare costs per visit to GP were in our study set at €107. This figure was based on reports from the National Board of Health and Welfare, in which the cost was calculated to €92 in 2004, costs for medication and medical services excluded. To this we added an annual increase in costs of 3%. We compared this with the local inter-county price list in Sweden for 2009 [24], where a visit to GP, including x-ray, medication and laboratory services, was charged with €124. From these figures we found our estimate per visit to be appropriate. We used the cost for physiotherapy treatment, €50, from the same inter-county price list, since no other figures were available. Charges to primary care for x-ray and ultrasound evaluations were retrieved from the hospital's radiology department.

Medication prescribed during the registered visits was retrieved from the EPR. Medication purchased without prescription was not registered. Costs for analgesics and nonsteroidal anti-inflammatory drugs were calculated as if every prescription was filled once and as if the patient had free medication, meaning that the costs were paid by the primary care unit. Costs were retrieved from the hospital pharmacy.

Patients who were referred to an orthopaedic specialist and to surgery generated additional costs. From the local inter-county price list [24] and the hospital administration we retrieved the costs for visits in outpatient care and a mean cost for ambulatory surgery in 2009, based on actual costs per patient. We estimated that ten MRI investigations would be performed in the patients evaluated in the present study. These figures illustrate the higher costs for secondary care (Table 2).

Table 2 Costs used in the economic evaluation

Costs	(Euros*)
Direct healthcare costs (per visit)	
General practitioner (25 min)	107
Physiotherapist (60 min)	50
x-ray, shoulder	65
Ultrasonography, shoulder	124
Medicine	Prices July 2009
Orthopaedic specialist	335
MRI	308
Shoulder surgery, uncomplicated, ambulatory care	2420
Indirect costs	
Sick leave from paid work (human cost method) per day	205

*1 Euro = 10.62 SEK. Average values in 2009, http://www.riksbank.se (Swedish National Bank)

Valuing productivity costs

The costs for sick leave were for the baseline value calculated according to the human capital approach [25,26]. This method places monetary weights on healthy time using market wage rates. It is an estimation of changes in productivity, based on the opportunity cost of the production that people would have contributed to, had they been at work. We assumed that the production costs were reflected by the salary. In this study, we only had data on sick leave periods (graded from 25 to 100% of full working time) prescribed by GPs. Partial sick leave was converted to 100% sick leave for each patient. The cost per day was calculated from the mean income in the region in 2008, provided by the Swedish Bureau of Statistics. The costs for productivity loss due to sick leave were calculated after this model presented by the Swedish Ministry of Industry in 2001 [27]:

Costs for productivity loss = Mean income + social fares + indirect taxes.

We assumed that social fares were 40% of the main cost and indirect taxes were 28%.

This equation shows what the worker must produce to cover his own income, payroll taxes and fees by law and agreement.

Human capital versus friction cost method

An alternative approach to the human capital method is the friction cost method [25,28]. In that case we assume that when a person has a period of sick leave, there is a pool of unemployed people that can replace the sick person. Hence, there will only be a productivity loss in a "friction" period until the new employee is recruited and trained to do the job. It is frequently argued that evaluations using the human capital approach overestimate the true costs to society [25]. Koopmanschap et al [29] found that cost of absence from work in 1988 when using the friction cost method was 38.7% of what they found by using the human capital approach. The cost for disability was 0.3% and for mortality 1.9% if the friction cost method was used. As part of the sensitivity analysis we displayed the effect of using the friction cost instead of the human capital method.

Data analysis

Costs were calculated for a six-month period. The arithmetic mean, standard deviations (SD), and median value were used to provide information about the total cost of treatment for all patients, and to illustrate the skewness in the distribution of costs and resource use. The total costs during six months, were multiplied by 2 in order to get the total annual costs for patients with shoulder pain in primary health care.

One-way and two-way sensitivity analyses were performed to explore the uncertainty [30], to demonstrate the impact of one parameter varying in the model, and to examine the relationship of two or more different parameters changing simultaneously.

We used a multivariable linear regression analysis to explore how gender, age and municipality, as independent variables, predicted costs.

Results

Patients

During six months 204 patients were registered; 103 women and 101 men. Mean age was 48 (SD 11) years. Eighty-nine per cent presented with subacromial or non-specific shoulder pain (Table 1). Nineteen patients (9%) came for postoperative rehabilitation. Twenty-nine patients (14%) were referred to an orthopaedic surgeon, and four of these (2%) went on to have surgery within a year. Fifty per cent of the cases were closed within six weeks, whereas 32 patients (16%) remained in the system for more than six months. Seven of these patients had been operated on. Baseline characteristics of the group are presented in Table 3.

Use of healthcare resources and sick leave

Consumption of healthcare resources and sick leave from work during six months are presented in Table 4. Forty patients (20%) had a period of sick leave prescribed by GP, mean 9.0 days (SD 29.2). Three patients (1.5%) were on sick leave due to their shoulder pain for more than six months; two of them with concomitant back pain and one with concomitant diabetes. Partial sick leave amounted to 11% (202 days) during the measured period.

Fifty-five patients (27%) consulted both a GP and a PT within 4 weeks. Nineteen patients (9%) had more than 10 physiotherapy treatments, 68 patients (33%) had none. The whole group had a mean of 6.7 (SD 7.0) physiotherapy treatments and 24.0 (SD 50.2) days of sick leave.

The consumption of medication, x-ray and ultrasound evaluations was low.

Costs

Costs for healthcare use and sick leave are presented in Table 4. The mean healthcare cost per patient was €326 (SD 389). Physiotherapy treatments accounted for 60%. This cost was twice as high as for visits to GP. The group of 73 patients that used the direct access to PT incurred a higher mean total cost for physiotherapy but lower healthcare and total costs.

The healthcare costs for the group with persistent symptoms were one fourth of all healthcare costs during six months. Median healthcare costs were €200 (Inter Quartile Range 113-397) for the whole group, whereas the median total costs were €249 (IQR 119-661). Eighty-four per cent of the total costs were due to sick leave prescribed by GP, for the whole group and for those who had surgery.

Total costs for the 45 patients (22%) with costs > €1000 during six months are presented in Table 5. Sick leave in this group amounted to 91% of the total costs, and for 44% of the healthcare costs (Figures 1 and 2). Seven patients in this group had no registered sick leave. Eighteen patients had symptoms for more than 6 months; five of them had no registered sick leave. The three patients with sick leave > 6 months contributed to 25% of the total costs.

Table 3 Baseline characteristics of patients with shoulder pain.

Characteristics	n = 204	n = 45*	n = 19**
Age (years); mean (SD)	48 (11)	48 (11)	48 (13)
Sex: male; n (%)	101 (49)	20 (44)	10 (51)
Treatment duration of current shoulder complaints***			
0-6 weeks	101 (50)	15 (33)	4 (21)
7-12 weeks	28 (14)	4 (9)	2 (10)
12-26 weeks	42 (21)	12 (29)	6 (32)
> 6 months	33 (16)	14 (29)	7 (37)
Duration of sick leave in the 8 weeks preceding inclusion****			
0 weeks	193 (95)	37 (82)	17 (89)
0-1 weeks	4 (2)	3 (7)	1 (5)
1-8 weeks	7 (3)	5 (11)	1 (5)

Numbers (percentages) are presented unless stated otherwise

* patients generating costs of > €1000 in 6 months

** had shoulder surgery

*** total treatment time for all patients. Costs calculated for 6 months

**** sick leave due to shoulder pain

Table 4 Costs (€) and consumption of healthcare resources and sick-leave during 6 months.

Direct costs	Mean number of visits	Total number	Cost per patient	Total costs
General Practitioner	0.89 (0.97)	181	95 (105)	19429
Physiotherapy	3.91 (7.40)	798	195 (369)	39825
X-ray*	0.28 (0.45)	57	18 (29)	3719
Ultrasound*	0.11 (0.31)	23	14 (39)	2857
Medicine*	0.28	58	4 (6)	718
Total healthcare costs			**326 (389)**	**66548**
Indirect costs				
Sick-leave**	9.04 (29.17)	1844	1743 (5626)	**355610**
Total costs			**2069 (5730)**	**422158**

N = 204. Means (SD) or total numbers are presented

* Number of patients given prescriptions for medicine, x-rays or ultrasound

** Days

The mean annual total cost for patients with shoulder pain in primary health care was €4139 per patient. Additional healthcare costs were generated by 29 patients (14%), MRI investigations, and from four cases of surgery in ambulatory care. The costs for secondary care for this group were estimated at €22475, corresponding to one third of the total costs for primary care.

Uncertainty

To show the uncertainty of the results we have reported the 95% confidence interval (CI) for the base case scenario for total costs and for healthcare costs in Table 6. The CI is €1 283- 2856 and €273-380, respectively. These intervals reflect the uncertainty caused by the fact that different patients use services such as x-ray and PT consultations with different frequencies. Additional uncertainty is related to the cost per unit of health services and the cost of sick leave per day. To show the importance of this uncertainty we performed a sensitivity analysis. For x-ray cost per examination we chose as an example +30% as a maximum average value and -30% as a minimum value. For each tested parameter value we computed the new expected costs and 95% CI based on the sample variation related to the (unchanged) frequencies of health service use and the new cost level per unit. The sensitivity analysis showed that the total cost was most sensitive to the choice of method for estimating the sick leave cost. Compared to the base case scenario where we used the human capital method, the friction cost method gave a reduction of the total cost per patient of 51.6% to €1001. Because of the dominance of the sick leave cost, the reasonable change of healthcare cost has just a minor influence on the total cost. A 30% change in physiotherapy cost or a 50% change in physician cost contributes just to a 2.8 and 2.3% change in the total cost, when changing these parameters one by one (one-way sensitivity analysis). When changing all the parameters of health service cost in the same direction (multi-way sensitivity analysis) as shown in Table 6, the total cost only changes by 5.7%.

The sensitivity analysis showed that the physiotherapy unit cost makes the biggest contribution to uncertainty in the health service cost. A 50% change in the

Table 5 Costs (€) and consumption of healthcare resources and sick-leave during 6 months for the group that cost > €1000.

Direct costs	Mean number of visits	Total	Cost per patient	Total costs
GP	1.71 (1.42)	77	184 (153)	8266
PT (cost per visit)	8.20 (13.62)	369	409 (680)	18415
X-ray*	0.33	15	22 (31)	979
Ultrasound*	0.27	12	33 (56)	1490
Medicine*	0.47	21	6 (7)	262
Total healthcare costs			**654 (671)**	**29412**
Indirect healthcare costs	Days			
Sick-leave**	40.82 (50.98)	1837	7875 (9833)	**354356**
Total cost			**8528 (9829)**	**383768**

N = 45. Means (SD) or total numbers are presented

* Number of patients given prescriptions for medicine, x-rays or ultrasound

** Days

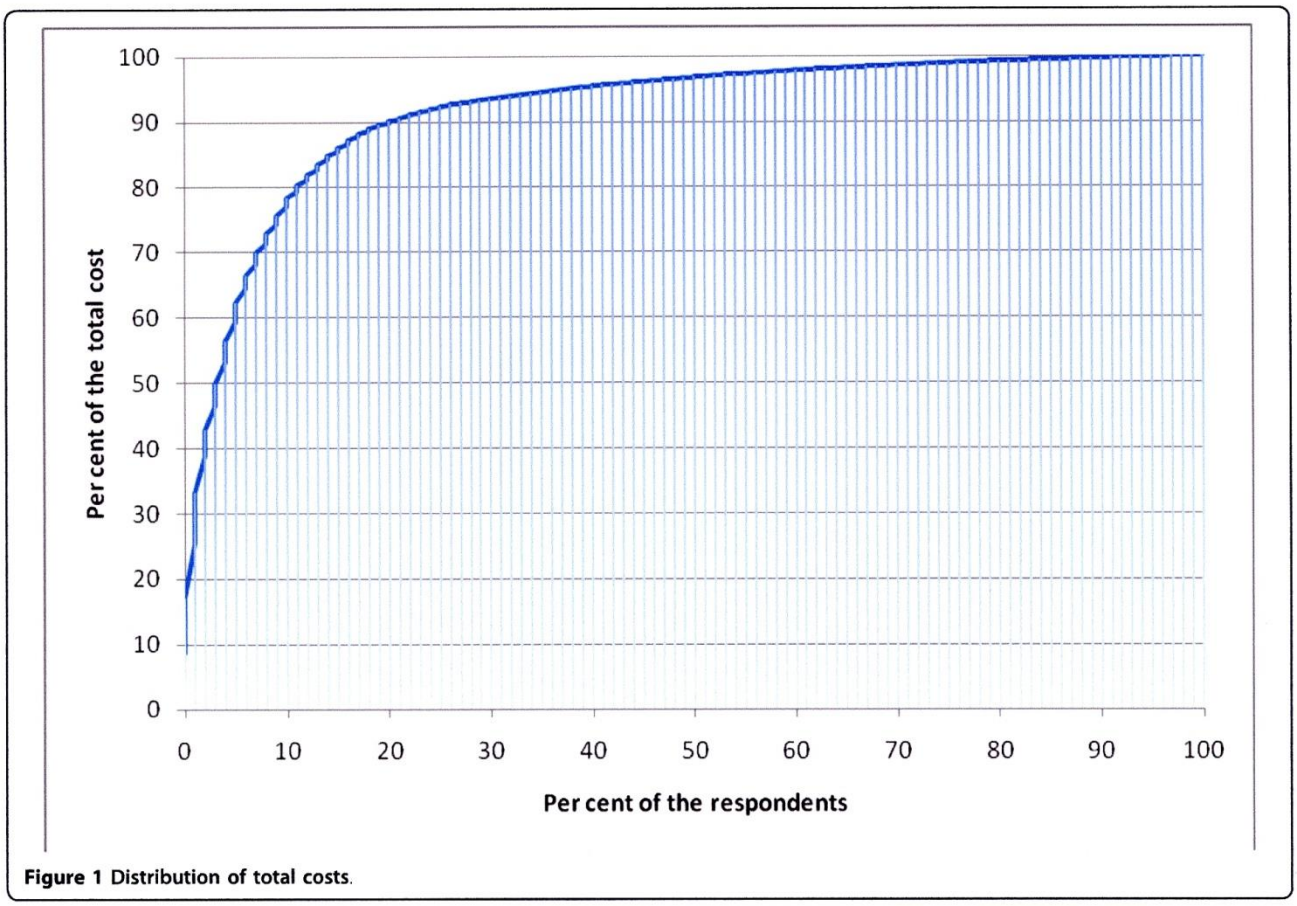

Figure 1 Distribution of total costs.

physician unit cost changes the health service cost by 14.6%, and a 30% change of physiotherapy costs gives a change of 18.0%. Relevant changes in the costs of x-ray, ultrasound and medicine only have a minor influence.

Gender, age and place of treatment did not influence total costs or health service costs. A sensitivity analysis using the logarithm of total cost and health service cost did not change this conclusion (Table 7).

Discussion

The main finding in the present study is that the mean healthcare costs amounted to less than 20% of mean total costs for patients with shoulder pain. Contrary to this, median healthcare costs contributed to 80% of median total costs, reflecting a minority of patients incurring high costs from long lasting sickness absence. Our findings are in keeping with previously published results on patients with shoulder and back pain [15,18,31].

Treatment strategies

The majority of patients were managed in primary care. Fifty per cent were treated within six weeks, and only two per cent were selected for surgery. This is in line with the intentions in guidelines and literature. Surgery should be considered if it represents an evidence-based approach when conservative measures fail. A treatment strategy for patients with subacromial pain is currently evaluated [32]. The observed increase in shoulder surgery does not correspond with a similar increase in prevalence of shoulder pain [33]. Vitale et al [7] discussed the increasing utilization of surgical procedures overall in recent years, and Hofmann [34] argued that there is a technological imperative in health care.

The inter quartile range of total costs varied from 119 to 661, illustrating the impact of long periods of sick leave. A fifth (22%) of the population generated costs of more than €1000 and accounted for 91% of the total costs. In the Dutch study [15], 12% of the patients cost more than €1000 and contributed to 74% of the total costs. The three patients with sick leave > 6 months contributed to 25% of the total costs. Efforts have been made to reduce long periods of sick leave, often combined with programmes for pain management [35-37]. Multidisciplinary rehabilitation programmes for patients with chronic low back, neck or shoulder pain are reported to be superior to treatment as usual for return to work [38,39]. However, a Cochrane review [40] on the subject did not find evidence to recommend multidisciplinary

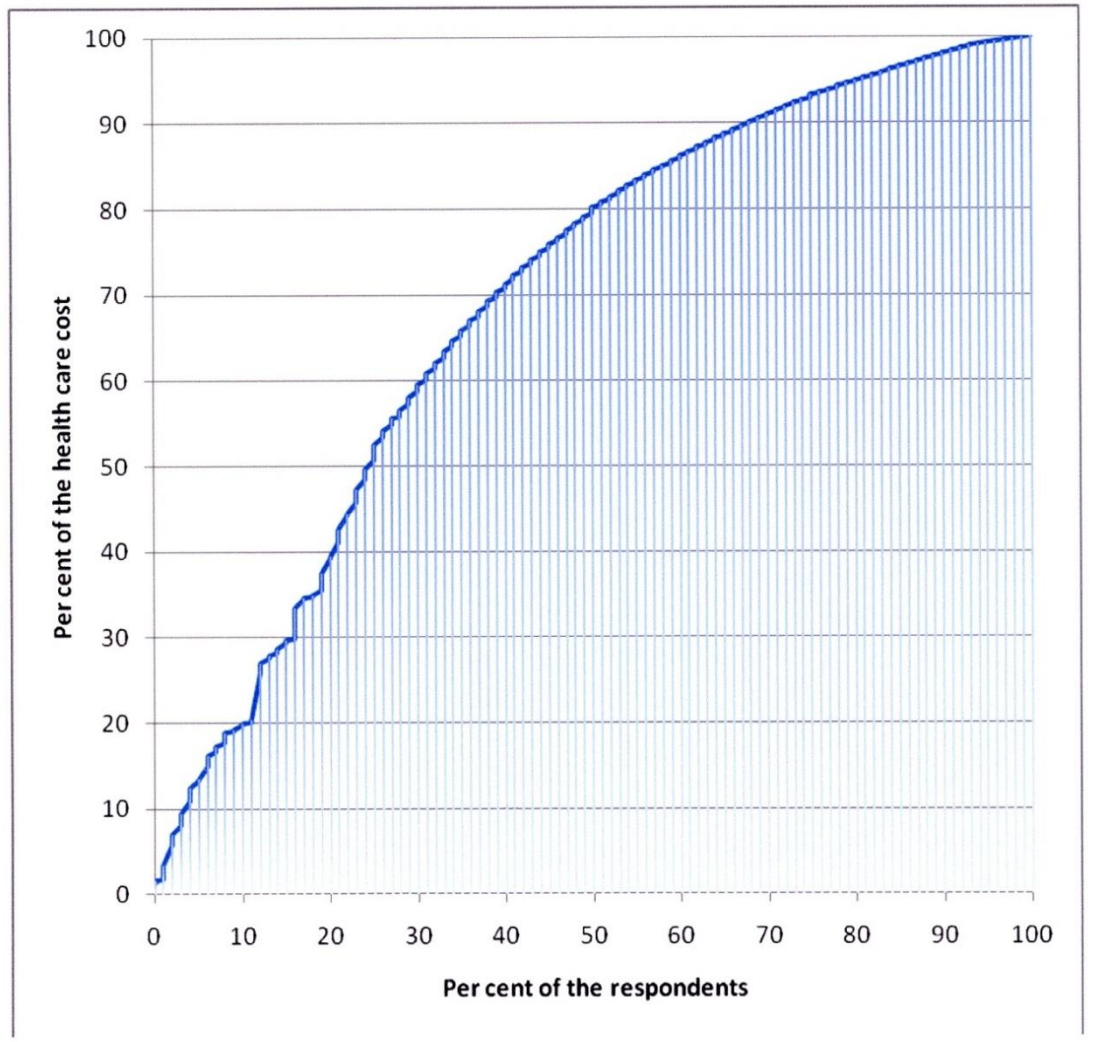

Figure 2 Distribution of total healthcare costs.

rehabilitation for patients with neck and shoulder pain. In the present study, physiotherapy treatments accounted for 60% of the healthcare costs and two thirds of the patients consulted a PT 3-4 times on average.

Whether an intervention programme is cost-effective or not depends on the relevance of the clinical outcomes and the costs needed to achieve this [41-43].

In the present study, 89% were diagnosed with subacromial or nonspecific shoulder pain. Feleus et al [44] found that 41% of the patients with non-traumatic neck, arm, or shoulder pain were given an unspecific diagnostic code at the first consultation in primary health care, and no differences were found in severity, complaints or functional limitations compared to patients with a specific diagnostic code. A specific diagnosis was given in 59% of the cases, mostly subacromial impingement syndrome. Distinction between diagnostic groups is important if these groups have different prognoses or require different management. Patients with specific diagnoses were more frequently referred for specialist treatment, while patients with non-specific diagnoses were more frequently referred for physiotherapy in the Dutch study [44]. Non-specific shoulder pain - the presence of pain without specific physical signs and pathology - is common, and Miranda et al [45] found that subjective complaints without clinical findings may indicate adverse psychological and psychosocial factors rather than an underlying pathologic condition. Several studies have reported that long-term sickness absence was associated with work conditions rather than with individual characteristics [46].

Future studies should include cost-effectiveness evaluation of various physiotherapy regimens or comparisons of physiotherapy with other treatments for shoulder pain. Functional limitations and duration of sick leave should be included as outcome measures. Such studies will be

Table 6 Uncertainty

Changed parameter or method	Percentage change in parameter	Percentage change in total cost	Percentage change in HC cost	Total cost (95% confidence interval)	Healthcare cost (95% confidence interval)
Base case scenario	0.0	0.0	0.0	2069 (1283-2856)	326 (273-380)
Parameters, one-way sensit. analysis					
PT cost per consultation	+30	2.8	18.0	2128 (1339-2917)	385 (317-453)
	-30	-2.8	-18.0	2011 (1227-2795)	268 (228-307)
GP cost per consultation	+50	2.3	14.6	2117 (1328-2906)	374 (318-430)
	-50	-2.3	-14.6	2022 (1238-2805)	279 (227-330)
Sick leave cost per day	+30	25.3	0.0	2592 (1575-3610)	326 (273-380)
	-30	-25.3	0.0	1546 (991-2102)	326 (273-380)
X-ray cost per consultation	+30	0.3	1.7	2075 (1288-2861)	332 (278-385)
	-30	-0.3	-1.7	2064 (1278-2850)	321 (268-374)
Ultrasound cost per consultation	+30	0.2	1.3	2074 (1287-2860)	330 (277-384)
	-30	-0.2	-1.3	2065 (1279-2851)	322 (269-375)
Medicine, unit used	+100	0.2	1.1	2073 (1286-2859)	330 (276-383)
	-50	-0.1	-0.5	2068 (1281-2854)	324 (271-378)
Parameters, multi-way sensit. analysis					
PT, GP per consultation	+30. +50	5.1	32.6	2176 (1384-2967)	432 (364-502)
	-30. -50	-5.1	-32.6	1963 (1182-2744)	220 (183-257)
PT, GP, x-ray, ultras. per consultation	As for one- way sensit. a.	5.7	35.8	2186 (1394-2979)	443 (373-513)
		-5.7	-35.8	1952 (1172-2733)	209 (173-246)
Method, one-way sensitivity analysis					
Sick leave cost based on friction method	-61.3	-51.6	0.0	1001 (685-1316)	326 (273-380)

Calculation of percentage change and new levels of total costs and health service costs by changing the unit costs of the different cost components

extensive and time-consuming, but study protocols have been presented [47,48].

Strengths and limitations of the study

A limitation of the present study is that we do not know whether the patients were relieved from their symptoms when the treatment period was ended, or if they disappeared out of the system for other reasons. The costs were limited to the primary diagnoses for the visit, and ignored costs associated with comorbidity. This is often the case in cost-of-illness studies and a simplification of real life, as has been pointed out by Koopmanschap [49]. When we looked closer into the three cases with sick leave more than six months, we found that they all had additional diagnoses. We could not gather such information for the rest of the group with the method applied.

Table 7 Multivariable linear regression analysis

Independent variables	Log (total cost)	Log (health service cost)
Place of treatment	-0,043	-0,119
	(0,118)	(0,224)
Gender	0,160	0,202
	(0,117)	(0,222)
Age	0,003	-0,002
	(0,005)	(0,010)
Constant	5,207	6,007
	(0,275)	(0,523)
Observations	203	203

The numbers in parentheses below the estimates are the standard errors

We had information about sick leave periods prescribed by GP, but we do not know if patients were actually absent from work all that time. We had no information about short-term sick leave, nor whether patients had sick leave prescribed by orthopaedic surgeon post-operatively. To fully estimate the cost for productivity loss additional data would have been required, for instance self-reported data from cost diaries or logbooks [15,42], or questionnaires [32]. However, a recent study suggests that self-reported data are less valid than register-based data to measure the number of days on sick leave [50].

The cost for medication is probably underestimated in this study. We had no information on the consumption of drugs, nor of the medication paid out of pocket. However, medication had a minor contribution to the total cost, and we do not expect that costs for medication would have an important impact on the results.

Generalization to other settings might be difficult, and will depend on how diagnostic codes are used, how reliable the registration is, and how costs are determined. The reliability of the cost estimates and varying research methodology have been under debate [51]. Charges for hospital services, like radiographic imaging, do not always reflect the actual unit cost of a production, but is merely a vehicle for transferring money between healthcare service units. However, these costs are easily available and most often the only costs available and therefore used in the present study. The measurement of productivity loss due to illness is highly dependent of the choice of approach, and this calls for standardisation on a national level. In the Netherlands a "Standardisation of costs; a manual for costing in economic evaluations" [52] was issued to eliminate some of the price differences between studies and to give guidelines for a uniform costing methodology.

The strength of the present study is that we were able to capture almost all patients consulting with all types of shoulder pain during a six-month period. There were few alternatives to medical care and data were manually controlled. We can double the total cost to illustrate the annual cost to society and to the health care system for shoulder pain in the chosen area. Our study provides direct and meaningful information about the size of the problem and can be an essential component in further cost-effectiveness analyses of different treatment strategies in primary health care.

Conclusions

Costs for sick leave for shoulder pain contributed to more than 80% of the total costs for society for this patient category. These results are in line with other studies on neck, shoulder and back pain. Health care interventions should focus on getting people back into the workforce, with special attention towards the small group that generates the highest costs. The model applied in the current study may be applied in future studies to analyse changes over time in terms of illness patterns in medical and health economic perspectives. A societal perspective is needed for the inclusion of all consequences of the interventions.

Acknowledgements

We wish to thank Dr Suzanne Palmblad-Olow, Department of Radiology, NU hospital, and Anders Lindgren, Secretary, County Council Västra Götalandsregionen, for helpful contributions in the initial stages of this study.

Author details

[1]Medical Faculty, University of Oslo, Oslo, Norway. [2]Department of Faculty of Health, Nutrition and Management, Faculty of Health Sciences, Oslo and Akershus University College of Applied Sciences, Postbox 4, St. Olavsplass, 0130 Oslo, Norway. [3]Department of Orthopaedics, Oslo University Hospital, 0027 Oslo, Norway.

Authors' contributions

LV, PJ and RE designed the study. LV collected the data. All authors participated in the study, and drafted the manuscript. PJ performed the statistical analysis. All authors read, critically revised, and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 15 May 2011 Accepted: 10 February 2012
Published: 10 February 2012

References

1. Chaudhury S, Gwilym SE, Moser J, Carr AJ: **Surgical options for patients with shoulder pain.** **Nat Rev Rheumatol** 2010, **6(4)**:217-226.
2. Mitchell C, Adebajo A, Hay E, Carr A: **Shoulder pain: diagnosis and management in primary care.** **BMJ** 2005, **331(7525)**:1124-1128.
3. Boykin RE, Heuer HJ, Vaishnav S, Millett PJ: **Rotator cuff disease - basics of diagnosis and treatment.** **Rheumatology Rep (Online)** 2010, **2(1)**:1-12.
4. Larson JS, Muller A: **Managing the quality of health care.** **J Health Hum Serv Adm** 2002, **25(3)**:261-280.
5. Regeringskansliet, Government offices of Sweden: **Hälso- och sjukvården i Sverige (Health and medical care in Sweden).** Socialdepartementet. Stockholm; 2007.
6. Nordqvist A, Rahme H, Hovelius L, Etzner M: **Axelns sjukdomar. (Shoulder diseases; in Swedish).** **Lakartidning** 2007, **104(19)**:1492-1496.
7. Vitale MA, Arons RR, Hurwitz S, Ahmad CS, Levine WN: **The rising incidence of acromioplasty.** **J Bone Joint Surg Am** 2010, **92(9)**:1842-1850.
8. Park HB, Yokota A, Gill HS, El Rassi G, McFarland EG: **Diagnostic accuracy of clinical tests for the different degrees of subacromial impingement syndrome.** **J Bone Joint Surg Am** 2005, **87(7)**:1446-1455.
9. Henkus HE, de Witte PB, Nelissen RG, Brand R, van Arkel ER: **Bursectomy compared with acromioplasty in the management of subacromial impingement syndrome: a prospective randomised study.** **J Bone Joint Surg Br** 2009, **91(4)**:504-510.
10. Moosmayer S, Lund G, Seljom U, Svege I, Hennig T, Tariq R, Smith HJ: **Comparison between surgery and physiotherapy in the treatment of small and medium-sized tears of the rotator cuff: a randomised controlled study of 103 patients with one-year follow-up.** **J Bone Joint Surg Br** 2010, **92(1)**:83-91.
11. Brox JI, Gjengedal E, Uppheim G, Bohmer AS, Brevik JI, Ljunggren AE, Staff PH: **Arthroscopic surgery versus supervised exercises in patients with rotator cuff disease (stage II impingement syndrome): a prospective, randomized, controlled study in 125 patients with a 2 1/2-year follow-up.** **J Shoulder Elbow Surg** 1999, **8(2)**:102-111.
12. Haahr JP, Andersen JH: **Exercises may be as efficient as subacromial decompression in patients with subacromial stage II impingement: 4-8-years' follow-up in a prospective, randomized study.** **Scand J Rheumatol** 2006, **35(3)**:224-228.
13. Ketola S, Lehtinen J, Arnala I, Nissinen M, Westenius H, Sintonen H, Aronen P, Konttinen YT, Malmivaara A, Rousi T: **Does arthroscopic acromioplasty provide any additional value in the treatment of shoulder impingement syndrome?: a two-year randomised controlled trial.** **J Bone Joint Surg Br** 2009, **91(10)**:1326-1334.
14. Virta L, Mortensen M, Eriksson R, Möller M: **How many patients with subacromial impingement syndrome recover with physiotherapy? A follow-up study of a supervised exercise programme.** **Advances in Physiotherapy** 2009, **11(3)**:166-173.
15. Kuijpers T, van Tulder MW, van der Heijden GJ, Bouter LM, van der Windt DA: **Costs of shoulder pain in primary care consulters: a**

prospective cohort study in The Netherlands. **BMC Musculoskelet Disord** 2006, **7**:83.
16. Nilsson G, Ahlfeldt H, Strender LE: **Computerisation, coding, data retrieval and related attitudes among Swedish general practitioners-a survey of necessary conditions for a database of diseases and health problems. Int J Med Inform** 2002, **65(2)**:135-143.
17. Nilsson G, Ahlfeldt H, Strender LE: **Textual content, health problems and diagnostic codes in electronic patient records in general practice. Scand J Prim Health Care** 2003, **21(1)**:33-36.
18. Ekman M, Jönhagen S, Hunsche E, Jönsson L: **Burden of illness of chronic low back pain in Sweden. SPINE** 2005, **30(15)**:1777-1785.
19. Wirehn AB, Andersson A, Ostgren CJ, Carstensen J: **Age-specific direct healthcare costs attributable to diabetes in a Swedish population: a register-based analysis. Diabet Med** 2008, **25(6)**:732-737.
20. Carlsson L, Borjesson U, Edgren L: **Patient based 'burden-of-illness' in Swedish primary health care. Applying the Johns Hopkins ACG case-mix system in a retrospective study of electronic patient records. Int J Health Plann Manage** 2002, **17(3)**:269-282.
21. Murphy R, Carr A: **Management of shoulder pain in general practice. InnovAiT** 2009, **2(7)**:402-407.
22. Robb G, Arroll B, Duncan R, Goodyear-Smith F: **Summary of an evidence-based guideline on soft tissue shoulder injuries and related disorders - part 2: management. J Primary Health Care** 2009, **1(1)**:42-49.
23. Geraets JJ, de Jongh AC, Boeke AJ, Buis PA, Spinnewijn WE, Geijer RM, Goudswaard AN: **Summary of the practice guideline for shoulder complaints from the Dutch College of General Practitioners. Ned Tijdschr Geneeskd** 2009, **153**:A164.
24. **Price list for the Western Health Care Region in Sweden 2009 (Prislista för Västra sjukvårdsregionen, Utomlänspriser 2009. För vårdtjänster enligt samverkansavtal om hälso- och sjukvård inom Västra Sjukvårdsregionen).** Edited by: Gotaland CoV. Skovde, Sweden; 2009:.
25. Drummond M, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddard GL: **Methods for the economic evaluation of health care programmes.** 3 edition. Oxford: Oxford University Press; 2005.
26. van Tulder MW, Koes BW, Bouter LM: **A cost-of-illness study of back pain in The Netherlands. Pain** 1995, **62**:233-240.
27. Malmquist C: **Kostnader i samband med arbetsbetingad ohälsa och stress. (publication on costs for productivity loss in relation to work-induced illness and stress).** Regeringskansliet, (Sweden.gov.se). Stockholm; 2001.
28. Schulper M: **The role of estimation of productivity cost in economic evaluation.** In **Economic evaluation in health care - merging theory with practice.** Edited by: Drummond M, McGuire A. Oxford: Oxford University Press; 2001:93-111.
29. Koopmanschap MA, Rutten FF, van Ineveld BM, van Roijen L: **The friction cost method for measuring indirect costs of disease. J Health Econ** 1995, **14(2)**:171-189.
30. Taylor M: **What is sensitivity analysis.** Consortium YHE: University of York; 2009, 1-8.
31. Seferlis T, Lindholm L, Nemeth G: **Cost-minimisation analysis of three conservative treatment programmes in 180 patients sick-listed for acute low-back pain. Scand J Prim Health Care** 2000, **18(1)**:53-57.
32. Dorrestijn O, Stevens M, Diercks RL, van der Meer K, Winters JC: **A new interdisciplinary treatment strategy versus usual medical care for the treatment of subacromial impingement syndrome: a randomized controlled trial. BMC Musculoskelet Disord** 2007, **8**:15.
33. Murphy RJ, Carr AJ: **Shoulder pain. Clin Evid (Online)** 2010, Jul 22; 2010.pii: 1107.
34. Hofmann B: **Is there a technological imperative in health care. Int J Technol Assess Health Care** 2002, **18(3)**:675-689.
35. Geraets JJ, Goossens ME, de Bruijn CP, Koke AJ, de Bie RA, Pelt RA, van den Heuvel WJ, van der Heijden GJ: **A behavioural treatment for chronic shoulder complaints: concepts, development, and study design. Aust J Physiother** 2004, **50(1)**:33-38.
36. Geraets JJ, Goossens ME, de Bruijn CP, de Groot IJ, Koke AJ, Pelt RA, Van der Heijden G, Dinant GJ, van den Heuvel WJ: **Cost-effectiveness of a graded exercise therapy program for patients with chronic shoulder complaints. Int J Technol Assess Health Care** 2006, **22(1)**:76-83.
37. Faber E, Kuiper JI, Burdorf A, Miedema HS, Verhaar JA: **Treatment of impingement syndrome: a systematic review of the effects on functional limitations and return to work. J Occup Rehabil** 2006, **16(1)**:7-25.
38. Storro S, Moen J, Svebak S: **Effects on sick-leave of a multidisciplinary rehabilitation programme for chronic low back, neck or shoulder pain: comparison with usual treatment. J Rehabil Med** 2004, **36(1)**:12-16.
39. Westman A, Linton SJ, Theorell T, Ohrvik J, Wahlen P, Leppert J: **Quality of life and maintenance of improvements after early multimodal rehabilitation: a 5-year follow-up. Disabil Rehabil** 2006, **28(7)**:437-446.
40. Karjalainen K, Malmivaara A, van Tulder M, Roine R, Jauhiainen M, Hurri H, Koes B: **Multidisciplinary biopsychosocial rehabilitation for subacute low back pain among working age adults. Cochrane Database Syst Rev** 2003, **2, Art. No: CD002194. DOI: 10.1002/14651858.CD002194.**
41. Österås H, Torstensen TA, Arntzen G, Österås BS: **A comparison of work absence periods and the associated costs for two different modes of exercise therapies for patients with longstanding subacromial pain. J Med Econ** 2008, **11(3)**:371-181.
42. De Bruijn C, Goossens M, de Bie R, Ament A, Geraets J, Dinant GJ: **Cost-effectiveness of an education and activation program for patients with acute and subacute shoulder complaints compared to usual care. Int J Technol Assess Health Care** 2007, **23(1)**:80-88.
43. Bergman GJ, Winter JC, van Tulder MW, Meyboom-de Jong B, Postema K, van der Heijden GJ: **Manipulative therapy in addition to usual medical care accelerates recovery of shoulder complaints at higher costs: economic outcomes of a randomized trial. BMC Musculoskeletal Disorders** 2010, **11**:200.
44. Feleus A, Bierma-Zeinstra SM, Miedema HS, Verhaar JA, Koes BW: **Management in non-traumatic arm, neck and shoulder complaints: differences between diagnostic groups. Eur Spine J** 2008, **17(9)**:1218-1229.
45. Miranda H, Viikari-Juntura E, Heistaro S, Heliovaara M, Riihimaki H: **A population study on differences in the determinants of a specific shoulder disorder versus nonspecific shoulder pain without clinical findings. Am J Epidemiol** 2005, **161(9)**:847-855.
46. Brox JI: **Regional musculoskeletal conditions: shoulder pain. Best Pract Res Clin Rheumatol** 2003, **17(1)**:33-56.
47. Bennell K, Coburn S, Wee E, Green S, Harris A, Forbes A, Buchbinder R: **Efficacy and cost-effectiveness of a physiotherapy program for chronic rotator cuff pathology: a protocol for a randomised, double-blind, placebo-controlled trial. BMC Musculoskelet Disord** 2007, **8**:86.
48. Kromer TO, de Bie RA, Bastiaenen CH: **Effectiveness of individualized physiotherapy on pain and functioning compared to a standard exercise protocol in patients presenting with clinical signs of subacromial impingement syndrome. A randomized controlled trial. BMC Musculoskelet Disord** 2010, **11**:114.
49. Koopmanschap MA: **Cost-of-illness studies. Useful for health policy. Pharmacoecon** 1998, **14(2)**:143-148.
50. Grovle L, Haugen AJ, Keller A, Natvig B, Brox JI, Grotle M: **Poor agreement found between self-report and a public registry on duration of sickness absence. J Clin Epidemiol** 2012, **65(2)**:212-218, Epub 2011 Aug 17.
51. Rice DP: **Cost of illness studies: what is good about them. Inj Prev** 2000, **6(3)**:177-179.
52. Oostenbrink JB, Koopmanschap MA, Rutten FF: **Standardisation of costs: the Dutch manual for costing in economic evaluations. Pharmacoecon** 2002, **20(7)**:443-454.

Pre-publication history
The pre-publication history for this paper can be accessed here:
http://www.biomedcentral.com/1471-2474/13/17/prepub

doi:10.1186/1471-2474-13-17
Cite this article as: Virta et al.: Costs of shoulder pain and resource use in primary health care: a cost-of-illness study in Sweden. BMC Musculoskeletal Disorders 2012 13:17.

Paper II

Modeling and Validating the Cost and Clinical Pathway of Colorectal Cancer

Paal Joranger, MS, Arild Nesbakken, PhD, Geir Hoff, PhD, Halfdan Sorbye, PhD, Arne Oshaug, PhD, Eline Aas, PhD

Background. *Cancer is a major cause of morbidity and mortality, and colorectal cancer (CRC) is the third most common cancer in the world. The estimated costs of CRC treatment vary considerably, and if CRC costs in a model are based on empirically estimated total costs of stage I, II, III, or IV treatments, then they lack some flexibility to capture future changes in CRC treatment. The purpose was 1) to describe how to model CRC costs and survival and 2) to validate the model in a transparent and reproducible way.* ***Methods.*** *We applied a semi-Markov model with 70 health states and tracked age and time since specific health states (using tunnels and 3-dimensional data matrix). The model parameters are based on an observational study at Oslo University Hospital (2049 CRC patients), the National Patient Register, literature, and expert opinion. The target population was patients diagnosed with CRC. The model followed the patients diagnosed with CRC from the age of 70 until death or 100 years. The study focused on the perspective of health care payers.* ***Results.*** *The model was validated for face validity, internal and external validity, and cross-validity. The validation showed a satisfactory match with other models and empirical estimates for both cost and survival time, without any preceding calibration of the model.* ***Conclusions.*** *The model can be used to 1) address a range of CRC-related themes (general model) like survival and evaluation of the cost of treatment and prevention measures; 2) make predictions from intermediate to final outcomes; 3) estimate changes in resource use and costs due to changing guidelines; and 4) adjust for future changes in treatment and trends over time. The model is adaptable to other populations.* ***Key words:*** *Markov model; validation; colorectal cancer; treatment cost; survival.* ***(Med Decis Making XXXX;XX:xx–xx)***

Cancer is a major cause of morbidity and mortality in the Western world, and colorectal cancer (CRC) is the second most common cancer in women and third in men.[1] The 5-year relative survival rates is 47% in Europe and 60% in the US.[2] The economic burden of cancer is expected to increase in the future, partly due to changing demographics and the introduction of new and resource-demanding treatments and screening methods. Thus, it is important to monitor the clinical course of cancer in patient cohorts to estimate cancer costs and develop sound methods to evaluate different treatment and screening regimens.

During the last decades, several models have been developed with an emphasis on describing and modeling the preclinical course of cancer.[3–8] A workshop among leading academic teams concluded that there

Received 19 August 2013 from Norwegian University of Life Science, Ås, Norway/Oslo and Akershus University College of Applied Sciences, Oslo, Norway (PJ); Oslo and Akershus University College of Applied Sciences, Oslo, Norway (AO); Oslo University Hospital, Oslo, Norway/K.G. Jebsen Colorectal Cancer Research Centre, Oslo, Norway/University of Oslo, Oslo, Norway (AN); the Cancer Registry of Norway/University of Oslo/Telemark Hospital, Skien, Norway (GH); Haukeland University Hospital, Bergen, Norway (HS); and University of Oslo, Oslo, Norway (EA). Financial support for this study was provided entirely by the authors' employers. The funding agreement ensured the authors' independence in designing the study, interpreting the data, and writing and publishing the report. The structure of the model and preliminary results were presented at the yearly Norwegian Conference Focusing Health Economists 2012, a workshop organized by the South-Eastern Norway Regional Health Authority. A presentation was also given at the Department of Health Management and Health Economics, University of Oslo. Revision accepted for publication 20 June 2014.

DOI: 10.1177/0272989X14544749

Supplementary material for this article is available on the *Medical Decision Making* Web site at http://mdm.sagepub.com/supplemental.

Address correspondence to Paal Joranger, Faculty of Health Sciences, Oslo and Akershus University College of Applied Sciences, Postboks 4 St. Olavs plass, NO-0130 Oslo, Norway; e-mail: Paal.joranger@hioa.no.

was considerable variation in cost estimates used in the various models and that future research should address modeling costs both more precisely and more transparently.[9] This limitation was confirmed in a review of economic evaluations of laparoscopic surgery.[10]

Compared with estimating the cost of CRC treatment empirically, using model-based estimates has several advantages.[11–14] Within a model framework, it is easier to adjust for changes like mortality rates, recurrence rates, and new treatments. Further, a model facilitates extrapolations of both costs and outcomes, allowing predictions 10–30 years into the future.

The model presented in this paper is similar to the model of Tilson and others[15] but has some extensions. First, time is defined explicitly by using a Markov framework instead of decision trees. A 3-dimensional data matrix captures time dependence according to the age of the patient and how long he or she has remained in a defined health state. Several other extensions were included in relation to 1) independent costs according to exclusively local recurrence, distant recurrence only, or a combination of local and distant recurrence; 2) survival and cost of treatment for re-recurrence, which is defined as a new recurrence after an apparently curative treatment of the first recurrence; 3) a separate decision tree for palliative chemotherapy; 4) side effects from surgery; and 5) (neo) adjuvant and palliative therapy.

To improve confidence in models, attention has been paid to transparency and validation.[16–18]

The purpose of this paper was to contribute to modeling CRC cost and survival by presenting a transparent model and validating it. The structure of the Markov model was specified in detail, explaining how different data sources were used to estimate costs and transition probabilities. We validated the model according to standard methods in order to show the precision of the model.[18]

The model was intended to be "multiply applicable" (general) so it could address a range of problems related to CRC treatment (surgery, chemotherapy, radiation, screening, lifestyle changes, etc.) and to be transferable to other countries that have access to a similar type of data.

THE MODEL

Model Structure and Flow of CRC Patients

The main outcomes of the model were recurrence, survival, and costs of CRC, which were estimated by means of a semi-Markov model.[16–20] The model structure was based on literature about CRC treatment and the natural history of CRC, national guidelines on CRC treatment, and expert opinion (oncologist, colorectal surgeon, and gastro physician). Figure 1 illustrates the Markov model with mutually exclusive health states (squares) and how patients could move between the health states. The model simulates the flow of a 70-year-old cohort of CRC patients from the year of diagnosis through periods of treatment and health states without CRC symptoms, until they died from CRC or other causes or were 100 years old (red lines). The length of 1 cycle was defined as 1 year. Each arrow was represented by a transition probability. Loop arrows illustrate health states where the patient can stay for more than 1 year.

The TNM system, which classifies disease as stages I–IV, was used to stage the disease at the time of diagnosis (see definition in Appendix 1).

In the model, standard half-cycle corrections were applied to adjust for mortality.[21,22] For costs, half-cycle corrections were not explicitly modeled but were performed indirectly, as the empirical data used to estimate CRC treatment costs consider compliance and mortality.

Algorithms and Modeling of Treatment and Disease Course

Primary treatment. The cohort entered the model in one of the TNM stages of year 0 when diagnostic and supplementary examinations were performed to establish disease stage, comorbidity, and the patient's general condition. Based on this workup, it was decided whether the treatment intention was curative or palliative. Curative treatment always implied resection of the primary tumor and regional lymph nodes, with or without pre- or postoperative radiotherapy and/or chemotherapy. After the histopathological report was finished, TNM staging and R (Residual tumor) classification were done. The treatment was defined as curative (called R0 resection) if the entire tumor was resected, there was no microscopic invasion of the resection margins, and there was no radiological evidence of residual tumor in other organs (no distant metastasis).

In the model, according to the primary treatment (year 0), the patients could go to "disease free" (DF) after a R0 resection or receive palliative treatment (or no treatment) if curative treatment was not possible. Palliative treatment included R1–2 resections and other palliative surgical procedures and/or (radio-) chemotherapy. Subsequently, the patients

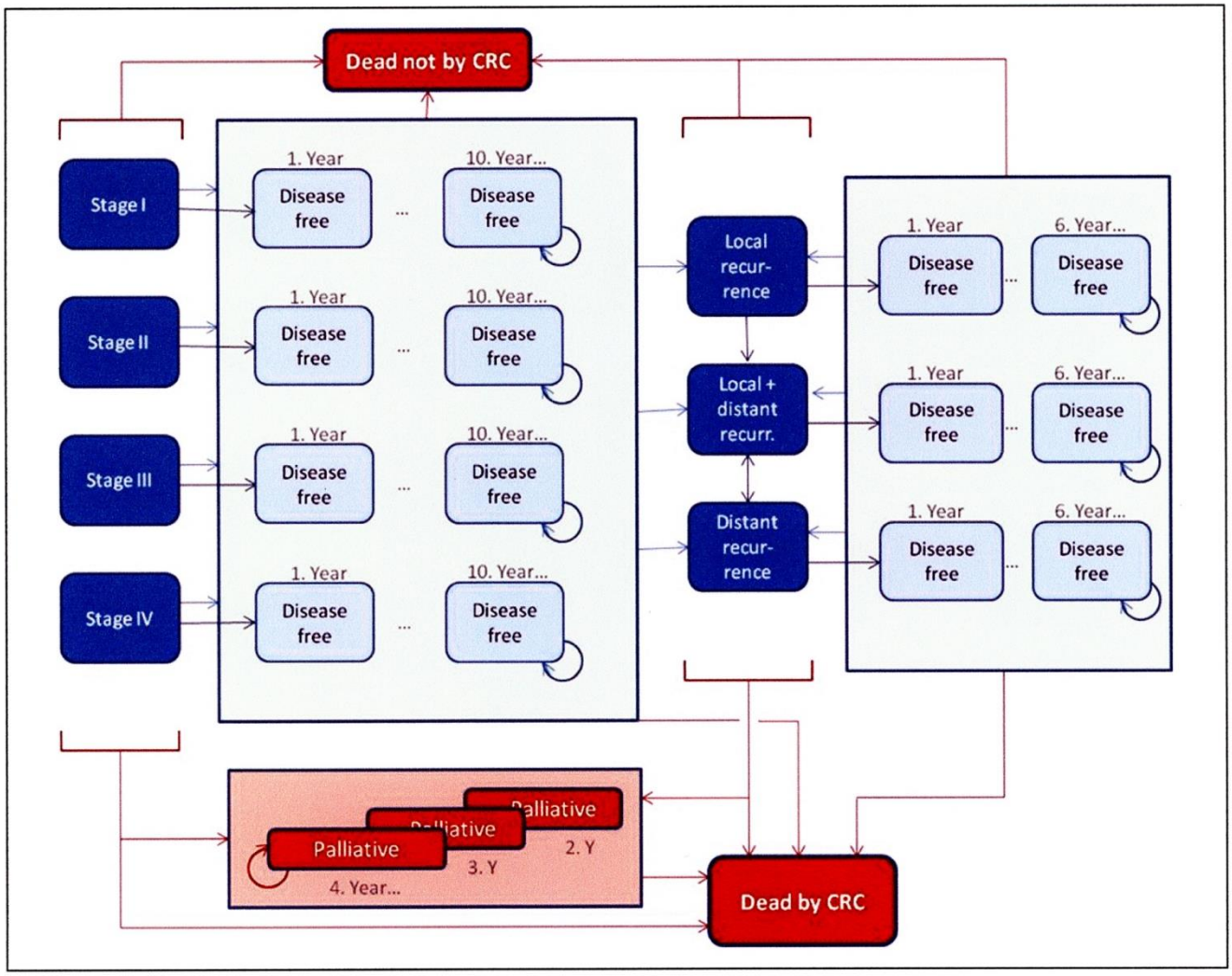

Figure 1 *The structure of the semi-Markov model. CRC = colorectal cancer.*

could die within 30 days after the operation (often due to treatment complications that resulted in the classification of CRC death), die later of CRC or causes other than CRC, or develop a recurrence for which some would receive treatment. A proportion of the patients started palliative chemotherapy during primary treatment and moved to further palliative chemotherapy the following years (years 2, 3, and 4 in palliation).

Follow-up and treatment of recurrence. From years 1 to 4 in the DF state after primary resection, the patients entered follow-up programs that varied according to the estimated risk of relapse and the national guidelines. From the point of DF, the patient could move to the next year in the DF state, could die (of CRC or other causes), or could have a recurrence, a local recurrence (LR), metastasis (distant recurrence [DR]), or a combination of LR and DR (local and distant recurrence, LDR).

The model principles behind the recurrence stages were similar to the treatment pattern based on the primary diagnosis. The patient would stay in the recurrence state for 1 cycle and receive treatment independent of whether the intention was curative. In the following cycle, the patient moved to DF after LR, DR, or LDR; died (of CRC or other causes); or moved on to palliative treatment. It was assumed that after a diagnosis of recurrence, the course of the individual patient would be identical to that of all other individuals with the same type of recurrence

and independent of the TNM stage at the time of primary diagnosis.

Survival and Mortality

In the model, mortality was dichotomized according to cause of death—mortality caused by CRC (disease-specific mortality) or any other cause (all-cause mortality exclusive of CRC mortality). Patients might die after surgery (<30 days); during palliative treatment followed by a terminal phase, whether due to CRC or the palliative treatment; or from an unrelated cause in any of the health states defined, including a DF state. Dying from CRC within the DF state is a correction for the patients who die from CRC within a month after diagnosis or for cases where the CRC is detected after the time of death (autopsy). Therefore, no CRC treatment costs created by recurrence are included for these patients. This correction comprises from 0.07% (stage I) to 0.7% (stage IV) of the patients every year.

Tunnel States

Probabilities for recurrence and death often depend on the duration of clinical disease, that is, the time from diagnosis. This time dependency (memory) was captured in the model by using "tunnel states."[19,20] By using tunnels, we can incorporate heterogeneity and simultaneously estimate survival and costs according to age groups. Including age is important when one is evaluating interventions like screening or primary prevention where the individual can be diagnosed with cancer and enter the model at any age.

All the DF after primary resection (light blue in Figure 1), DF after recurrence, and the 3 "palliative" states were parts of tunnels. The DF tunnels begin in year 1 after primary treatment and are continuous throughout year 10 (cycle 11).

Perspective and Cost

The model has the perspective of the health care payer. The CRC treatment costs included in-hospital CRC treatment (including diagnostics), treatment for complications, treatment of recurrence, radiation and chemotherapies, follow-up, and patient visits to a general practitioner.

To estimate the cost in each health state, submodels were used to reflect the treatment pathways. The unit cost of the different treatments was mainly based on the reimbursement systems in Norway. The Hospitals in Norway were reimbursed partly by block grants and partly through fees for service. The fee-for-service component was directly linked with diagnosis-related groups (DRGs).* The unit cost for the chemotherapy drug is based on Oncolex† estimates.

To determine cost, frequency, and compliance with follow-up and surveillance, we applied market prices, data from Kørner and others,[23] and the national guidelines for CRC treatment.[24]

For palliative chemotherapy treatment, a decision tree was used to estimate costs according to treatment paths and was then distributed according to the treatment years in the Markov model. For each treatment in the decision tree, a separate cost model was developed, which took into account the cost of the drug, computed tomography scanning, complications, and the time spent on therapy by the nurse, pharmacist, and medical practitioner. The model corrected for noncompliance and discontinuation of chemotherapy.

Data Source

Inputs were based on Norwegian data as far as possible. An important source was an observational study in the period 1993–2010 of 2049 patients diagnosed with CRC at Oslo University Hospital–Aker (referred to as the OUS data).[25,26] The OUS data included a wide range of variables related to CRC treatment (mostly surgical procedures) and time to recurrence and death.

Information from the National Patient Register (referred to as the NPR data), based on data related to an analysis by Aas,[27] was used to quantify some types of treatments from the years 2003 and 2004. The NPR data were collected for 2 counties in Norway and should be representative of the general population. There have been differences in CRC risk between regions in Norway. Around the year 2000, the relative risk of CRC for inhabitants in these 2 counties was close to 1.0 compared with the overall national risk.

Other data sources were national life tables, international scientific publications (e.g., overall survival of patients receiving palliative chemotherapy), and

*In 2010, approximately 900 DRGs were used to reflect the hospital case mix. One DRG received a value reflecting the average cost of treating 1 patient. Each DRG received a weight reflecting the intensity of the treatment compared with treating the average patient.

†Oncolex is a Norwegian encyclopedia for oncology health personnel (www.oncolex.org). It includes background information and updated procedures for treatment of CRC.

expert opinions in parameters considered not to have essential effects on outcome. Data for treatment after recurrence were limited and based on literature and expert opinion.

A Scandinavian prospective, population-based, observational study was an important data source concerning palliative chemotherapy.[28]

Calibration is normally used as a complement to data sources, but it was concluded that calibration was not needed due to a good fit of the model.

VALIDATION OF THE MODEL

The model was validated according to face validity, internal and external validity, and cross-validity.[29] We concentrated the validation on survival and the cost of CRC treatment, being the 2 main outcomes of the model. They are endpoint estimates of numerous intermediate calculations in the model and therefore indirectly also represent a rough test of the subparts of the model.

Face Validity

To determine face validity, we assess whether the results make sense and can be explained at an intuitive level.[29] The model structure, including health states, patient flow, and the data used, was closely evaluated by medical experts and could therefore easily be recognized. The estimates used for important cost components during the first year of treatment and during palliative treatment were based on established assumptions and classification of treatment options and their costs (see Appendix 1).

Internal Validity

Internal validity implies that the mathematical calculations were correct and consistent with the specification of the model.[29] Algorithms were used to check that the row of the data matrices for the annual transition probabilities summed to 1 and that the number of patients in the Markov model was constant for all cycles.

Validation of the economic model was more complicated. Extensive use of checking calculations was performed to test whether the results based on the model were replicable. Approximately 150 one-way sensitivity simulations were run to test whether the model behaved as expected (i.e., according to size, direction and symmetry). No anomalies were found.

Cross-Validity

For cross-validation or between-model validation, we compared models (or methods) that were independently developed, but aimed at estimating the same outcomes, to investigate whether they achieved similar results.

Overall survival. A 10-year overall survival estimated by the model was compared with statistical estimations (Weibull distribution, STATA) based on the OUS data (Figure 2). Overall survival was used as an endpoint because it reflected the sum of all moves, was not used directly as an input in the model, and was normally more reliable than the other relevant output measures.

In cross-validation, the degree of model and data independence is important. The higher the degree of independence, the more valuable is the validation. Independence is obtained if the models compared use completely different data sources and apply different types of methods. In the cross-validation, different methods were used but the data were partially dependent. In the model, OUS data were used to estimate disease-free survival and time to recurrence after primary treatment (for the patients who underwent R0 resection), while in the statistical analyses, data for all 2049 patients were used to estimate overall survival. In addition to the OUS data mentioned, the literature and expert opinions were used to find the R0 resection rate in patients with recurrent disease and to estimate the survival for patients in palliative treatment. A simplified model (based on a portion of the patients in the OUS data) was then developed to estimate disease-free survival after R0 resection for recurrence and time to re-recurrence.

The curves in Figure 2 indicate that the structure and assumptions of the model correspond well with the Weibull regression. Even without a preceding calibration of the model, the differences in survival between the 2 methods of estimation in the fifth year after diagnosis were –0.002, –0.002, 0.014, and –0.004 for stages I, II, III, and IV, respectively, and in the tenth year were –0.001, –0.018, –0.004, and 0.001, respectively. The areas between the curves (based on the model and the Weibull estimation) showed the difference in survival between the 2 methods and were 0.27, –0.22, 1.62, and –0.04 months, respectively. Over 10 years, the weighted difference was on average 11.5 days (0.38 months) for all stages.

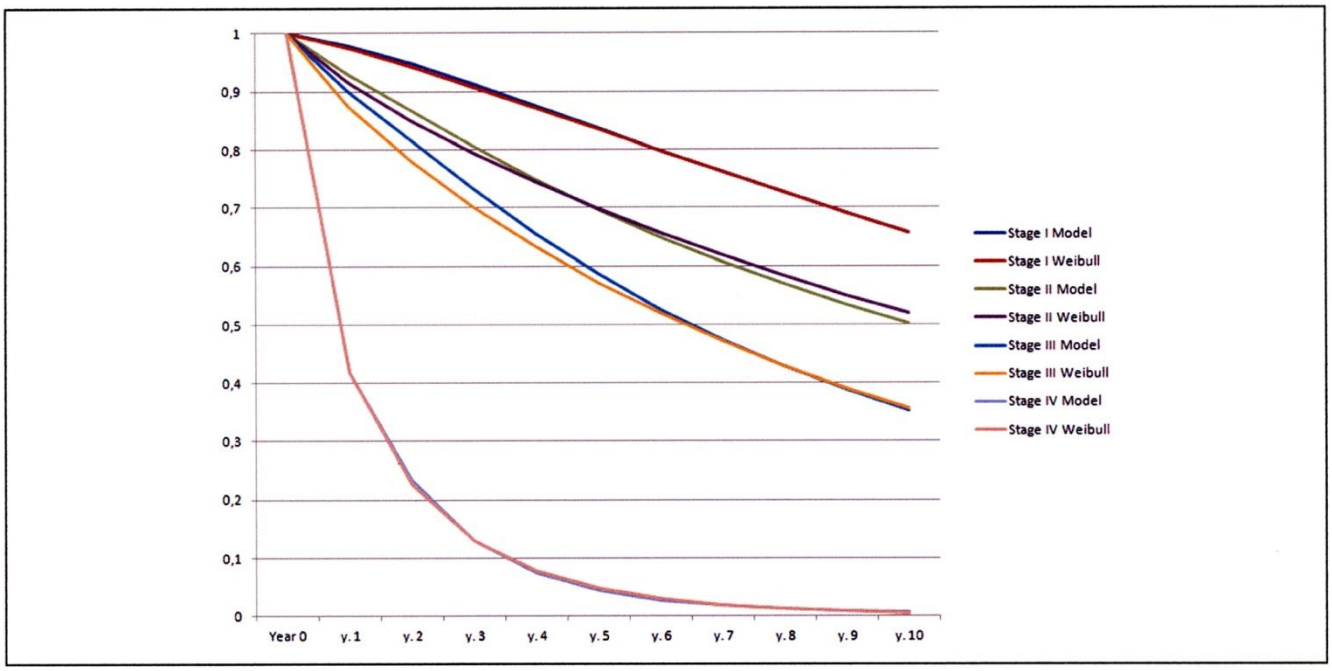

Figure 2 Model *means the overall survival curve simulated by the model, and* Weibull *means the overall survival curve estimated with Weibull distribution using our data from Oslo University Hospital.*

The curves based on the model for stages I, II, and III overestimated the survival during the first years. This seemed to be caused by structural elements in the model due to a mortality lag. A DF patient with a recurrence death within 12 months was not defined as dead during that cycle according to the model (except for a small fraction who would die before knowing about the recurrence) but was rather moved to one of the health states of treatment for recurrence. This could also explain why stage I patients have the smallest deviation (lowest recurrence) and stage III the greatest, with stage II in between. We could reduce this problem by shortening the cycles from 12 to 6 months or even 1 month.

Furthermore, the model underestimated the survival curve for stage II in the last part of the 10-year period (Figure 2). The reasons seemed partly to be that patients in stage II were relatively more often struck by isolated local recurrence than the average CRC patient, and patients with local recurrence survived for a longer time than patients who experienced DR or LDR, while the model assumed the same survival time for all 3 types of recurrence.

The good fit between the statistical estimations and the simulation model of the 40 estimates of comparison indicated that the structure seemed to be close to reality and captured the true treatment pathways.

From cycle 11, after completing the DF tunnel, the patient was assumed to have no recurrence. For this part of the model, the overall mortality was based mainly on data from the Norwegian Life Table. To verify consistency with natural survival for the Norwegian population, estimated overall survival from the model was validated against the life table for patients aged 70–100 years.

As expected, none of the overall survival curves of the 4 stages cross, none of these cross the overall survival curve based on the National Life Table, and the curves show a gradual change (Figure 3). The data used were partly dependent in this validation, because the life table was also used in the model as part of the background mortality from cycle 11 or when the cohort was 81 years of age (year 11 in Figure 3). Another weakness in this validation was the expected difference between the natural survival of the general population used in the model and the expected background survival for the CRC group. The latter group seemed to have a lifestyle that increased the risk of death apart from CRC.[30–33]

CRC treatment costs. Comparing our CRC costs with a non-Norwegian study is difficult as a result of often major differences related to time horizon of costs, treatment regimens, unit costs, general health conditions, and whether the cost of recurrence and

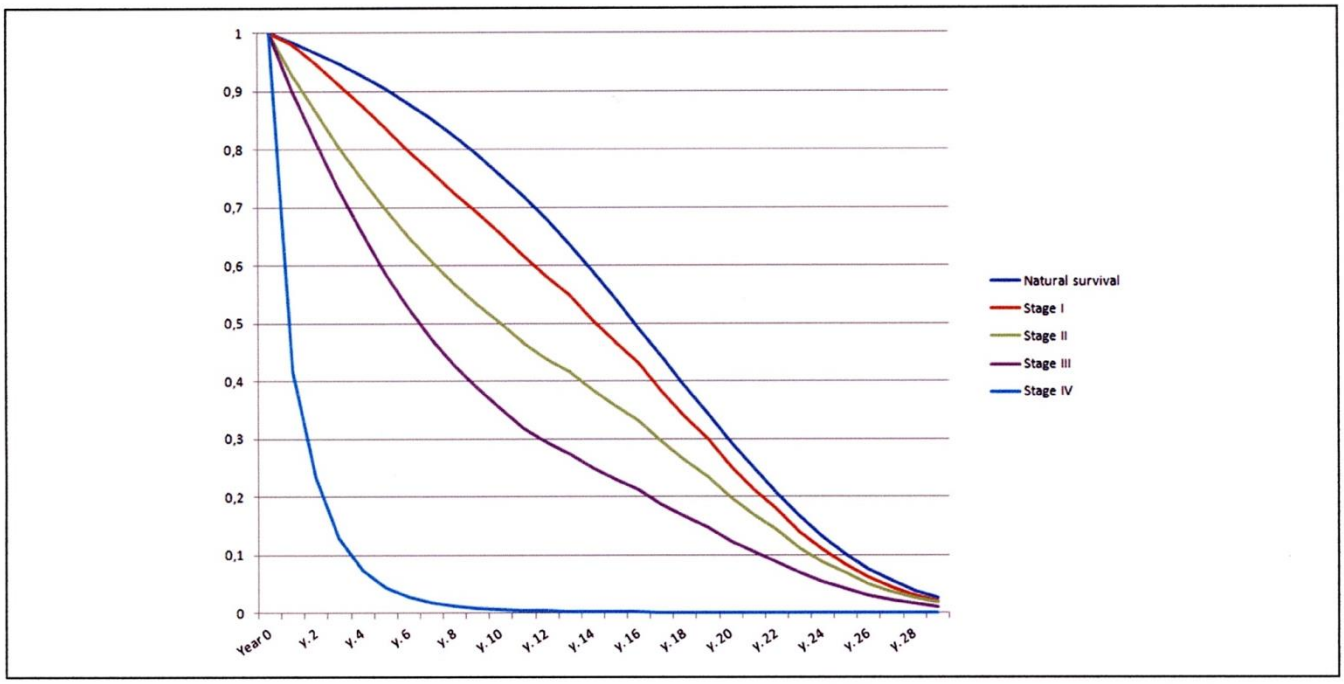

Figure 3 Overall survival for the Norwegian population (without colorectal cancer [CRC]) and for CRC patients according to the disease stage.

palliative treatment was included. In addition, diagnostics, treatment regimens, and cost can change significantly over time. Nevertheless, we compared our results with a recent Irish study that thoroughly described the treatment regimen and other important conditions so that we could correct for the differences in the assumptions of the Irish study and our study.[15] The study published by Tilson and others[15] was a model study (decision trees) based on 4268 CRC patients (National Cancer Registry Ireland, 2004–2005), local hospitals' databases and protocols, literature, and expert clinical opinions. The Irish model was developed independently from ours with regard to both data and structure.

Our model estimated the total lifetime CRC costs to be as follows: for stage IV patients, €61,396; for stage III patients, €49,894; for stage II patients, €33,501; and for stage I patients, €23,386 (average 2011: 1 Euro = 7.79 NOK and 1 Euro = 1.39 USD). When corrected for the exchange rate (8.19%, average 2008–2011) and annual inflation (3.4%, average Irish Consumer Price Index for health 2008–2011, see www.cso.ie), the costs for stages I, II, and III in Ireland were 17.7%, 26.7%, and 13.8% higher, respectively, than in our model, while stage IV was 30.7% lower. There were some important differences in prices and treatment regimens between the studies.[‡] After we adjusted for these factors in our model, the cost differences between Tilson and others'[15] model and our model (Tilson's model minus our model) were –3.0%, –1.3%, 3.6%, and –1.2% for stages I, II, III, and IV, respectively. These 4 deviations were all within the estimated confidence interval in the study of Tilson and others, which varied between ±12% and 29% of the stage cost estimates.

External Validation

External validation compares actual event data with the result from a model simulating the same scenario. For a multi-application model like ours, validation could be general or could be specific to each application of the model. "External validation and predictive validation are critical as they most closely correspond to the model's purpose—to help decision makers anticipate what will occur if they take certain actions."[18(p738)]

[‡]Tilson and others[15] assumed higher prices for resections for both the colon and rectum; less use of palliative chemotherapy; less use of adjuvant chemotherapy for stage III; no recurrence for stage I; a cost for recurrence equal to the cost of stage IV; and no category given for 'nonsurgical supportive treatment and care' except best supportive care.

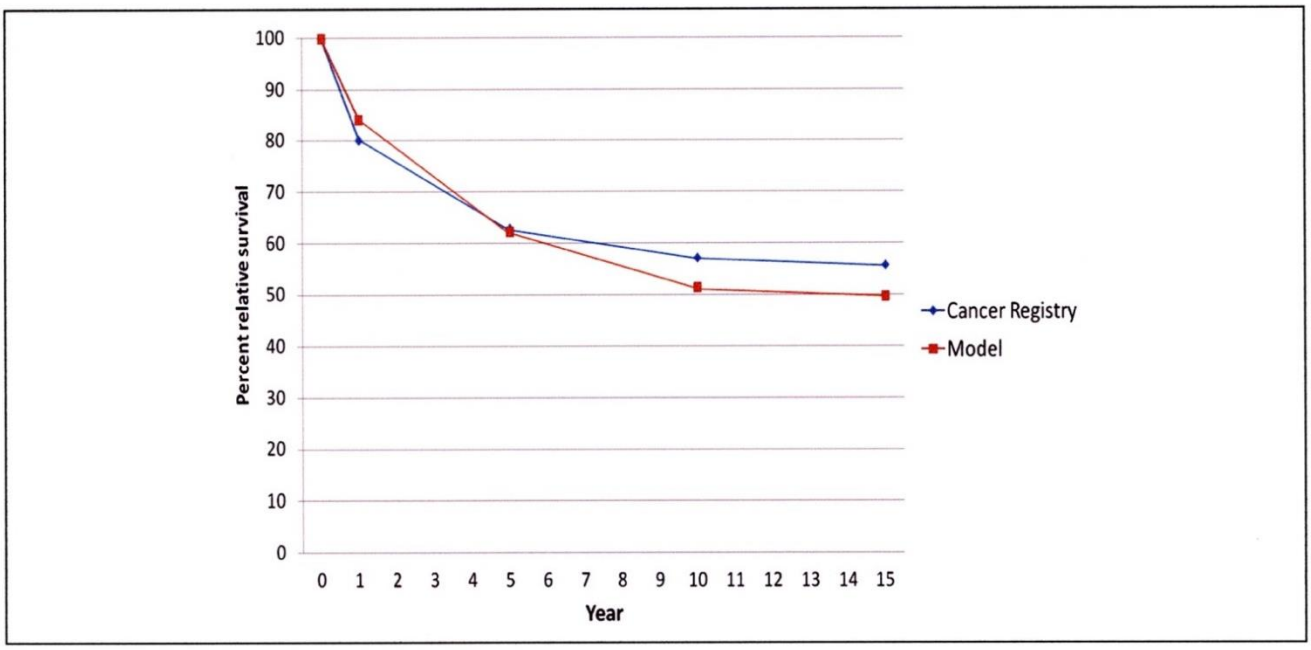

Figure 4 External validation between the model and data from the Cancer Registry of Norway.

Relative survival. The first external validation was done by comparing relative survival estimated by the model with patients monitored by the Cancer Registry of Norway, which monitors the whole Norwegian population (about 5 million), while our OUS data were based on a catchment area of about 4.2% of the total population. Further, as the Cancer Registry of Norway organizes the data differently from the OUS data, the 2 data sources were highly independent.

In Figure 4, the comparison revealed a 3.9% higher relative survival for the model during the first year and 0.9%, 5.6%, and 5.6% lower relative survival 5, 10, and 15 years after diagnosis. As observed for the cross-validation, this external validation showed that the model predictions were too high immediately after diagnosis, rather accurate after 5 years, and slightly lower at year 10. Fifteen years after diagnosis, the validation indicated that the difference had stabilized between 5% and 6%.

The overestimation of survival up to 5 years after diagnosis can partly be explained by 1-year cycles, as argued above. The general picture beyond 5 years was a higher mortality in our model. One possible explanation might be that our statistical analyses were based on older data (1993–2010), while the analyses from the Cancer Registry of Norway displayed relative survival estimates for the follow-up period 2008–2010. According to the Cancer Registry of Norway, relative survival has increased gradually during the last decades.

CRC treatment costs. In an external validation with a relevant population, total costs were compared with empirically estimated ("model-free") total costs based on a Norwegian population study by Aas.[27] The methods used for estimation in the 2 studies were therefore highly independent, although for 16% of the treatments, the 2 studies used the same data source (NPR data), but these data were collected from different time periods.

Adjusted for the annual price change of the Norwegian DRGs, the estimated CRC costs reported by Aas[27] were €29,890 for all patients (2011 Euro) in the control group with no screening, while the model estimate was €41,548 (39.0% higher). This difference could partly be explained by the increased intensity of palliative treatment in the period between Aas's study (1999–2001) and our study (2010–2011). In 1999–2001, biological agents were rarely used in Norway. In addition, the use of radiation therapy and surgery for metastases was not included in Aas's study. When adjusting for these two, as well, the model result was 9.1% higher than the estimate by Aas. Furthermore, taking into account Aas's estimated costs for a 5-year period, the model estimate was 3.1% higher than Aas's model-free estimate. Additionally, corrections due to different follow-up schemes for the

2 periods must be made. Our model was based on the Norwegian guidelines from 2010, while Aas's study was based on the actual follow-up years 1999–2001.[24] If we instead use the guidelines from that period, the model estimate for the average CRC cost was 1.3% higher than the model-free estimate.[34] Both with and without this last correction, the model seems to fit well when taking into account that the lower and upper confidence interval for Aas's CRC cost estimate was 11%–12%.

DISCUSSION, APPLICATION, AND FURTHER IMPROVEMENTS

This study demonstrates a multi-application (general) model for estimating survival and costs for CRC patients. The validation of the model revealed a good match with reality both in survival and in costs. The model is suitable for addressing a wide range of CRC-related themes, the most important being the estimation of cost and survival associated with different treatments and prevention measures. Such information is essential for future revisions of guidelines and health care providers.

Application of the Model

The following applications and advantages of the model should be emphasized. The model 1) estimates the costs and survival time of an average CRC patient according to different disease stages; 2) estimates final outcomes from changes in intermediate outcomes such as decline in both recurrence and mortality rate due to improvement in preoperative diagnostics; 3) can be used in economic evaluations by applying modest adjustments and developments needed to perform economic evaluations of different types of screening, prevention, introduction of new treatment, and follow-up alternatives; 4) estimates resource use; 5) can adjust for changed parameters over time (time-dependency) and simultaneously account for the time since CRC treatment, consequences of CRC patient age, alterations in treatment, and changes in cost and resource use over time (i.e., by using the 8 tunnels and the 3-dimensional data matrix); and 6) is transferable to other countries with access to the same types of data. Since calibration has not been used in this model, applying data from another country and building the model with the recommendations and assumptions provided in the present article and appendixes should, in principle, effect a similar goodness of fit. See more on applications in Appendix 2.

Weaknesses and Further Developments of the Model

The cycles in the model were set to 1 year, which to some extent restricts the precision of the model. Linearity is especially likely to be an unsatisfying approximation during the first year after a diagnosis of stage IV, the first year after a diagnosis of recurrence, and the first year of palliative treatment. The problem would be reduced if the cycles were reduced. One-month cycles could be incorporated into the model by changing the cycle lengths for all health states or by building separate sub-Markov models for selected health states, making the cycle length shorter (such as a cycle length of a week or a month) for the selected states, and retaining the 1-year cycle for the other health states. Shorter cycles would make the model more complex and would require more detailed data, which accentuates the tradeoff between model complexity and accuracy. Our plan for the next generation of this model is to include shorter cycles for some health states.

The OUS data range from 1993 to 2010. Since some of these data are relatively old, survival in the model is lower, which can be explained by the older and less effective treatments. These deviations are 0.9%, 5.6%, and 5.6% lower relative survival at 5, 10, and 15 years, respectively, after diagnosis when the model estimates are compared with data from the Cancer Registry of Norway (see the section "Relative Survival").

In the model, we used a cohort of patients diagnosed at the age of 70 years. This may have resulted in a higher survival rate than if we had used the average age in the OUS sample. The average age during stages I–IV at the time of diagnosis was 69.9, 72.3, 70.4, and 70.5 years, respectively. When we compared these average ages with our 70-year-old patients (based on Weibull regressions), we found that the differences in overall survival 10 years after diagnosis were –0.2%, 4.2%, 0.7%, and 0.03%, respectively, for the 4 stages. These differences are quite small and could only to some extent affect the external validation between the model and data from the Cancer Registry of Norway.

The palliative submodel was suitable for exploring treatment paths and costs, but there was no explicit, built-in time dimension. An approximation was therefore used to disperse the costs over time. A better solution could be to build a separate sub-Markov

model with weekly or monthly cycles for palliation into the main model.

The model has 70 health states and 8 "tunnels" and uses a 3-dimensional data matrix to handle the changing rates of recurrence and mortality by age and year since primary CRC or recurrence. The complexity could be a drawback for decision makers to fully understand all the mechanisms of the model. Still, hardly any part of the model could be further simplified; quite the contrary seems to be the issue.

In the model, the OUS data were used as the basis for the survival analysis. Because the OUS data were collected over a long time period, it could imply that subgroups of patients are treated differently from current guidelines. Basing the inputs on newer data could adjust for these differences.

Many articles analyze the costs during the last year that patients are alive or for other time periods, but no relevant articles analyzing costs of the activity related to "best supportive care" were found. We assume, however, that this cost is partly included in "digestive malignancy" (Table 7, Appendix 1).

There seems to be a lack of data on the resource use related to treatment for local and distant recurrence, separately or combined, mainly because the relevant registers are not organized to estimate this. Solutions could be to conduct observational or retrospective studies on resource use after a recurrence or expand the registries to include such data.

In such general models as ours, external validations can be applied to some components of the model or to the model as a whole.[18] Our external validation was applied to the model as a whole by validating for the main outcomes: survival and cost. One problem with these could be that errors in different parts of the model may cancel out each other and in sum give a result for the model consistent with the external data. Future validations of the model should therefore also focus on different components of the model. Some of these could be certain categories of cost (e.g., palliative chemotherapy and treatment for metastasis), time to recurrence, and time to re-recurrence.

Our validations for survival were comparisons with observed data for survival rates at whole-year points. For future validations of the model, a suitable alternative to this approach could be life years or life expectancy.

Further development of the model should include quality of life and should elaborate on the palliative and re-recurrence part of the model. Including quality of life during primary treatment, treatment for recurrence, and palliative care would, to a greater extent, capture the severity according to the TNM stage.

The effect of CRC diagnosis on quality of life in the disease-free stages should also be considered.

ACKNOWLEDGMENTS

The authors thank 2 anonymous referees and the editors for their valuable comments. We acknowledge the Department of Health, Nutrition and Management, the Faculty of Health Sciences, Oslo, and Akershus University College of Applied Sciences for funding Paal Joranger's doctorate.

REFERENCES

1. Jemal A, Center MM, DeSantis C, Ward EM. Global patterns of cancer incidence and mortality rates and trends. Cancer Epidemiol Biomarkers Prev. 2010;19(8):1893–907.

2. Coleman MP, Quaresma M, Berrino F, et al. Cancer survival in five continents: a worldwide population-based study (CONCORD). Lancet Oncol. 2008;9(8):730–56.

3. Frazier AL, Colditz GA, Fuchs CS, Kuntz KM. Cost-effectiveness of screening for colorectal cancer in the general population. JAMA. 2000;284(15):1954–61.

4. Ness RM, Holmes AM, Klein R, Dittus R. Cost-utility of one-time colonoscopic screening for colorectal cancer at various ages. Am J Gastroenterol. 2000;95(7):1800–11.

5. Tappenden P, Chilcott J, Eggington S, Sakai H, Karnon J, Patnick J. Option appraisal of population-based colorectal cancer screening programmes in England. Gut. 2007;56(5):677–84.

6. Lansdorp-Vogelaar I, van Ballegooijen M, Zauber AG, Habbema JDF, Kuipers EJ. Effect of rising chemotherapy costs on the cost savings of colorectal cancer screening. J Natl Cancer Inst. 2009; 101(20):1412–22.

7. Ramsey SD, Wilschut J, Boer R, van Ballegooijen M. A Decision-analytic evaluation of the cost-effectiveness of family history–based colorectal cancer screening programs. Am J Gastroenterol. 2010;105(8):1861–9.

8. Sharp L, Tilson L, Whyte S, et al. Cost-effectiveness of population-based screening for colorectal cancer: a comparison of guaiac-based faecal occult blood testing, faecal immunochemical testing and flexible sigmoidoscopy. Br J Cancer. 2012;106(5):805–16.

9. Pignone M, Russel L, Wagner J. Economic Models of Colorectal Cancer Screening in Average-Risk Adults—Workshop Summary. Washington, DC: Institute of Medicine and National Research Council; 2005.

10. Murray A, Lourenco T, de Verteuil R, et al. Clinical effectiveness and cost-effectiveness of laparoscopic surgery for colorectal cancer: systematic reviews and economic evaluation. Health Technol Assess. 2006;10(45):1–141.

11. Etzioni R, Ramsey SD, Berry K, Brown M. The impact of including future medical care costs when estimating the costs attributable to a disease: a colorectal cancer case study. Health Econ. 2001;10(3):245–56.

12. Clerc L, Jooste V, Lejeune C, et al. Cost of care of colorectal cancers according to health care patterns and stage at diagnosis in France. Eur J Health Econ. 2008;9(4):361–7.

13. Brown C, Fenech D, McLeod R. Reconstructive techniques after rectal resection for rectal cancer. Cochrane Database Syst Rev. 2008;(4):CD006040.

14. Ramsey SD, Berry K, Etzioni R. Lifetime cancer-attributable cost of care for long term survivors of colorectal cancer. Am J Gastroenterol. 2002;97(2):440–5.

15. Tilson L, Sharp L, Usher C, et al. Cost of care for colorectal cancer in Ireland: a health care payer perspective. Eur J Health Econ. 2012;13(4):511–24.

16. Smith-Spangler CM. Transparency and reproducible research in modeling: why we need it and how to get there. Med Decis Making. 2012;32(5):663–6.

17. Caro JJ, Briggs AH, Siebert U, Kuntz KM. Modeling good research practices—overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force–1. Med Decis Making. 2012;32(5):667–77.

18. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force–7. Med Decis Making. 2012;32(5):733–43.

19. Briggs AH, Claxton K, Sculpher M. Decision Modelling for Health Economic Evaluation. Oxford (UK): Oxford University Press; 2006.

20. Kuntz KM, Weinstein M. Modelling in economic evaluation. In: Drummond M, McGuire A, eds. Economic Evaluation in Health Care. Oxford (UK): Oxford University Press; 2001. p 286.

21. Naimark DMJ, Bott M, Krahn M. The half-cycle correction explained: two alternative pedagogical approaches. Med Decis Making. 2008;28(5):706–12.

22. Naimark DMJ, Kabboul NN, Krahn MD. The half-cycle correction revisited: redemption of a kludge. Med Decis Making. 2013; 33(7):961–70.

23. Körner H, Söreide K, Stokkeland P, Söreide J. Systematic follow-up after curative surgery for colorectal cancer in Norway: a population-based audit of effectiveness, costs, and compliance. J Gastrointest Surg. 2005;9(3):320–8.

24. Helsedirektoratet. Nasjonalt handlingsprogram med retningslinjer for diagnostikk, behandling og oppfølging av kreft i tykktarm og endetarm [Contract No. IS–1792]. Oslo (Norway): Helsedirektoratet; 2010.

25. Sjo OH, Lunde OC, Nygaard K, Sandvik L, Nesbakken A. Tumour location is a prognostic factor for survival in colonic cancer patients. Colorectal Dis. 2008;10(1):33–40.

26. Nesbakken A, Nygaard K, Westerheim O, Mala T, Lunde OC. Local recurrence after mesorectal excision for rectal cancer. Eur J Surg Oncol. 2002;28(2):126–34.

27. Aas E. Cost-Effectiveness of Screening for Colorectal Cancer with Once-Only Flexible Sigmoidoscopy and Faecal Occult Blood Test. Oslo (Norway): Oslo University, Health Economics Research Programme; 2009.

28. Sorbye H, Pfeiffer P, Cavalli-Björkman N, et al. Clinical trial enrollment, patient characteristics, and survival differences in prospectively registered metastatic colorectal cancer patients. Cancer. 2009;115(20):4679–87.

29. Weinstein MC, O'Brien B, Hornberger J, et al. Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices—Modeling Studies. Value Health. 2003;6(1):9–17.

30. Vrieling A, Kampman E. The role of body mass index, physical activity, and diet in colorectal cancer recurrence and survival: a review of the literature. Am J Clin Nutr. 2010;92(3):471–90.

31. Marchand LL, Wilkens LR, Kolonel LN, Hankin JH, Lyu L-C. Associations of sedentary lifestyle, obesity, smoking, alcohol use, and diabetes with the risk of colorectal cancer. Cancer Res. 1997; 57(21):4787–94.

32. Giovannucci E. Modifiable risk factors for colon cancer. Gastroenterol Clin North Am. 2002;31(4):925–43.

33. Huxley RR, Ansary-Moghaddam A, Clifton P, Czernichow S, Parr CL, Woodward M. The impact of dietary and lifestyle risk factors on risk of colorectal cancer: a quantitative overview of the epidemiological evidence. Intl J Cancer. 2009;125(1):171–80.

34. (Norwegian Gastro Intestinal Cancer Group) Norwegian Gastro Cancer Group. Kolorektalcancer og Analcancer. En Veiledning for Leger. Aurskog (Norway): 1999.

Appendix 1 (paper II): Data and statistical analyses

(App. 1: Published in a Web-only format)

1. Data source

In this study, as far as possible, inputs were based on Norwegian data. An important source of input data was an observational study at Oslo University Hospital – Aker (OUS), in the following referred to as OUS data, which included a wide range of variables related to CRC treatment. Most of the variables described different surgical procedures, the times to recurrence and the time of death. The study consisted of 2,049 patients diagnosed with CRC in the period 1993-2010, including all the CRC patients diagnosed at OUS. The hospital treated all patients from a defined catchment area of approximately 210,000 inhabitants.

Textbox 1.

The Markov model was based on transition matrixes with the following notation:

$$tp_{t,a}^{f,s} = tp_{time\ in\ tunnel,age}^{from,to}$$

f = the health state from which the patient was moving

s = the health state to which the patient was moving

t = number of years (time) the patient has been in the tunnel

t = 1, 2…10

t = 0, the patient had not entered a tunnel, but was in one of the treatment states

a = the age of the patient leaving a health state

The abbreviations for *health states*:

TNM (I, II, III or IV): TNM stages defining primary treatment the first year after CRC-diagnosis

REC (LR, DR or LDR): Treatment states for recurrence; local (LR), distant (DR) and both local and distant recurrence (LDR)

R-REC (R-LR, R-DR or R-LDR): Treatment states for re-recurrence or later recurrence

c: Referring to "*disease free"* after primary resection *or "disease free"* after REC (a supplement, like IIIc, DRc or R-DRc)

D: Death by all causes other than CRC

30d: Death within 30 days after surgery

CD: Death by CRC more than 30 days after surgery

Pa: Palliation

The TNM classification system (AJCC/UICC) was used to classify the disease stage at the time of diagnosis, where T-stage reflects the depth of tumor invasion into and through the bowel wall, N-stage reveals whether or not there are metastases in the regional lymph nodes, and M-stage shows the presence

of distant metastases. According to TNM, the disease is classified in stages I-IV. Stage I means the tumor is confined to the intestinal wall; stage II means the tumor is invading through the intestinal wall (and might invade adjacent organs or perforate the visceral peritoneum); stage III has lymph node(s) metastases; stage IV has distant metastases.

Information from the National Patient Register (in the following, referred to as the NPR data), based on data related to an analysis by Aas [2] was used to quantify some types of treatment from the years 2003 and 2004. The data were collected for two counties in Norway and should be representative for the general population.

National life tables (Statistics Norway) and four international published papers estimating the overall survival for patients receiving palliative chemotherapy, were used. Two of the studies were based on European populations [3 4], one on North Americans [5], and the last on Scandinavian countries [6].

When information from public sources was not available, *expert opinion* (oncologist, colorectal surgeon, and a gastro physician) was considered a legitimate method for assessing parameters [7]. Generally, expert opinions were used if the parameters were considered (based on literature, model simulation or expert opinion) not to have essential effects on output. If they were considered to affect the output significantly, sensitivity analyses were carried out. Expert opinion has been used in computing parts of the treatment model for palliative chemotherapy, partly for the use of radiation and for certain parts of the sub-model for recurrence and re-recurrences. Calibration is normally used as a complement to data sources [8]. Calibration would imply a systematic adjustment of model parameters by letting the model output govern the model input. After comparing the result of the model with the data from the same population (see Chapter 4), it was concluded that calibration was not needed due to a good fit of the model.

Input data presented here are mostly estimates from calculations and statistical analysis, and presented with a precision that does not always correspond with the quality of the underlying data source. This high "precision" is used in this appendix to make it easier for the readers to test the model by doing their own simulations with in-data close to the data calculated in our Excel-based model.

Endpoints have often been defined differently in studies of CRC, leading to a lack of comparability, so our CRC survival analyses were performed with endpoint definitions agreed upon in a recent consensus conference [9] and shown in Table 1.

Table 1*. Definitions of events according to three main endpoints*

Event	Endpoint		
	Disease-free survival	**Time to recurrence**	**Overall survival**
Local or regional recurrence	F	F	I
Distant metastasis (DR)	F	F	I
Second primary, CRC	F	I	I
Second primary, other cancer	F	I	I
Death from CRC	F	F	F
Death from other cancer	F	C	F
Non-cancer death	F	C	F
Treatment-related death	F	C	F
Loss to follow-up	C	C	C

Failures is F, censoring is C and ignoring is I.

2. Incorporation of data

Important factors in the model, such as survival curves, transition probabilities and frequencies were derived from data in the literature, from the primary CRC data or from official registers. Some of the data could directly be found in published papers, such as the probability of a patient getting a certain treatment, e.g., the probability of prescribing adjuvant chemotherapy to a patient with a stage-III disease. Often, available data could not be directly incorporated into the model. Important sources of data for modeling the course of CRC were different kinds of survival curves presented in literature.

These often presented the cumulative survival for a certain period and indicated the probability for an average patient to survive at least to time t. given by

$$S(t) = P(T > t) = 1- F(t)$$

where *t* was years and F(t) the cumulative density function. Let S (t-u) be the cumulative survival for the last period in time, where the *u* is the length of a Markov cycle. Then the probability of surviving through one cycle was defined as

$$s(t) = S(t) / S(t-u). \qquad (0)$$

Based on Equation (0), the probability of failure (recurrence or death) during a cycle was defined by

$$tp(t_u) = 1- s(t) = 1 - [S(t) / S(t-u)]. \qquad (1)$$

If two- and three-year survival was 0.9 and 0.8, respectively, then the probability of staying alive from year two to three was S(3) / S(3-1) = 0.8/0.9 = 0.89, and the transition probability of dying between years two and three would be 1 – [S(t)/S(t-u)] = 1- (0.8/0.9) = 0.11.
Based on data from four studies [3-6], Equation (1) was used to estimate the survival function for patients going through palliative treatment. The survival curves from each study were scanned and visually extracted. The four datasets were merged by weighting each study equally, and the probability of surviving years one through four were computed to be 0.675, 0.350, 0.175 and 0.087, respectively. Finally, equation (0) was used to estimate the transition probabilities $tp_{1,a}^{Pa,Pa}$ (see textbox 1) of staying alive between each cycle (or year), such as between year one and year two.

2.1 Estimating survival curves and transition probabilities based on individual data

Several statistical models can be applied to estimate the transition probabilities from our individual level data. As pointed out in Briggs et al. [10], the parametric survival function, Weibull, is preferable, as it allows the transition probabilities to change as a function of duration (such as time since diagnosis). Based on the Weibull model, separate hazard rates and transition probabilities could be estimated according to TNM stages for each year. The following equation and parameter was used:
Let S(t) = exp [-H(t)] and H(t) = λt^p , and inserted in Equation (1) the probability of failure during one cycle was given by

$$tp(t_u)_{fail} = 1 - \exp[\lambda(t-u)^p - \lambda t^p] \qquad (1^*)$$

where λ was the "scale" factor and p the "shape" factor. If p<1, then (and statistical significant) there was evidence for a decreasing hazard over time. Further, let $\lambda = \exp(a_0 + \sum a_i X_i)$ were *i* goes from 1 to *n*, then the nomenclature was given by

$$tp(t_u)_{fail} = 1 - \exp([\exp(a_0 + \sum a_i X_i)](t-u)^p - [\exp(a_0 + \sum a_i X_i)]\, t^p) \quad (2)$$

where a_0 is the constant, a_1 referred to age at diagnosis and a_2 to gender. The estimated coefficients are reported in Table 2. Based on Equation (2), the transition probability of no failure during one cycle was given by

$$tp(t_u)_{no\text{-}fail} = 1 - tp(t_u)_{fail} = tp_{t,a}^{TNM,TNM} \quad (3)$$

The estimated coefficients from Equation (2) were used to estimate the transition probability of a failure during one specific year conditioned on surviving until the start of that specific year, reported in Table 2. For instance, will the transition probability for a cohort of 70-year-old individuals, diagnosed with TNM, stage II and with a 50/50 mix of men and women be 0.0757. Thus, according to Equation (3), the transition probability $tp_{2,72}^{IIc,IIc}$ of staying in *"disease free after II"* from year two to three would be 0.9253 (1-0.0757).

$$\begin{aligned} tp(3_u)_{fail} &= 1 - \exp([\exp(-6.91471 + 0.0389469 * 70 + 0.266219 * 0.5)](36-12)^{0.7850102} \\ &\quad - [\exp(-6.91471 + 0.0389469 * 70 + 0.266219 * 0.5)] * 36^{0.7850102}) \\ &= 0.0757 \end{aligned}$$

where the parameter *u* (length of Markov cycle) was set to 12, because months were used in the analysis, while the length of a cycle was measured in years. Thus, after being treated according to the R0 resection and surviving without recurrence during the first and second year, 0.0757 of the patients would get a recurrence (local, distant, or both local and distant recurrence) or die of causes other than CRC during the third year. Then, according to Equation (3), the transition probability $tp_{2,72}^{IIc,IIc}$ of staying in *"disease free after II"* from year two to three would be 1 – 0.0757 = 0.9253.

Table 2*. The parameters for estimating the transformation probability related to stages I–IV with R0 resection.*

Variables	Parameter for DFS curve				Parameter for TTR curve				OS for recurr.
	I	II	III	IV	I	II	III	IV	
ϒ	1.113	.785	.776	.837	.977	.655	.675	.714	.797
a_0	-11.087	-6.915	-6.086	-2.362	-11.609	-5.776	-4.876	-1.695	-5.299
a_1 (age)	.068	.039	.034	-.007	.071	.026	.020	-.011	.039
a_2 (gender)	.408	.266	.251	.006	.422	.182	.150	-.268	.138

OS is overall survival, DFS is disease-free survival, TTR is time to recurrence and ϒ is p in Stata. Source: OUS data.

To estimate the transition probabilities from the primary year of treatment (year 0) to *"disease free after TNM,"* the above method needed to be adjusted. Only a proportion of the patients received R0 surgery, e.g., 0.943 in stage III (Table 4). Thus, if 0.824 of the R0 patients were estimated to be eligible for the "*disease free after TNM*" state (disease free survival) after primary treatment, then 0.824 * 0.943 = 0.777 of all the patients diagnosed with stage III would move from the clinical stage III to the first year of "disease free after stage III", defined by $tp_{0,70}^{III,IIIc}$. The R0 correction was particularly important for stage IV, where only 0.059 got a R0 operation. In Table 3, row four, all the adjusted transition probabilities were reported. Adjustments were also needed for the transition probabilities connecting the treatment year after recurrence and the first year of being "*disease free after recurrence*". For the rest of the years of "*disease free after TNM*", Equation (3) was used to estimate the transition probabilities directly from the Weibull regressions, as argued above.

Transition probabilities from primary treatment to *recurrence* were more complex to estimate. First, the parameters for time to recurrence (TTR) were estimated using the OUS data similarly to disease-free survival (DFS), and then the transition probabilities for recurrence were estimated by using Equation (2), i.e., the proportion of CRC-patients suffering a recurrence during primary treatment (year 0). The transition probability from primary treatment to recurrence also had to be adjusted by categorizing recurrences into local recurrence, distant recurrence and both local and distant recurrence.

Table 3*. Transition probabilities from primary treatment (according to TNM stages) to other health states. For the abbreviation in the first column, see also text box 1.* *Source: OUS.*

The probability to:	Abbrevi-ation	TNM stages			
		Stage I	Stage II	Stage III	Stage IV
Die within 30 days after surgery	$tp_{0,70}^{TNM,30d}$	.0030	.0310	.0290	.1120
Die of CRC the first year after surgery	$tp_{0,70}^{TNM,CD}$	.0016	.0108	.0270	.4185
Receive palliative treatment the first and second year after treatment	$tp_{0}^{TNM,Pa}$	.0000	.0000	.0000	.3718
Be considered disease-free the first year after treatment	$tp_{0,70}^{TNM,TNMc}$	.9646	.8479	.7766	.0375
Get a local recurrence during the first year after treatment	$tp_{0,a}^{TNM,LR}$	.0029	.0122	.0100	.0007
Get both local and distant recurrence during the first year after treatment	$tp_{0,a}^{TNM,LDR}$	.0007	.0061	.0149	.0010
Get a distant recurrence during the first year after treatment	$tp_{0,a}^{TNM,DR}$	.0117	.0626	.0985	.0074

The estimated transition probability from primary treatment in clinical stage II to local recurrence for a patient 70 years old was, as an example, defined by

$$tp_{0,70}^{II,LR} = \text{REC}_{\text{II}} * \text{R0}_{\text{II}} * \text{LR}of\text{R}_{\text{II}} * (1 - \text{CD}_{\text{REC30d}} - \text{CD}_{\text{TNM+REC}}) \quad (4)$$

where R0_{II} was the portion of the stage II patients having R0 surgery and categorized as disease-free (Table 4) - only a disease-free person could get a recurrence. REC_{II} was the probability of getting a recurrence during the first year of primary treatment for a stage-II patient, given that the patient had R0 surgery for the primary CRC. Further, $\text{LR}of\text{R}_{\text{II}}$ was the portion of the REC_{II} getting a local recurrence (LR). $\text{CD}_{\text{REC30d}}$ was the probability of dying during the first month after recurrence and was estimated to be 0.0422. These patients were excluded, because it was assumed that they died within 30 days after diagnosis of recurrence, or the recurrence was diagnosed post-mortem (autopsy). Further, it was assumed that these patients did not receive any treatment. $\text{CD}_{\text{TNM+REC}}$ was the probability of dying from CRC in the period of 2-12 months after the primary treatment, given that the patient had a recurrence that year. Hence, a double-counting of cost for patients dying within the first year was avoided (for both primary treatment and the treatment cost of recurrence). Inserting the coefficients from Table 4 for a 70-year-old patient in Equation (4), the transition probability was

Table 4*. Conditional probabilities for estimating transition probabilities related to recurrence, given that the patients have received a R0 surgery.*

The probability of:	Abbrevi-ation	Stage I	Stage II	Stage III	Stage IV
Getting a recurrence (during the year of primary treatment) according to TNM stages, given R0 surgery for the primary CRC	REC_{TNM}	.0177	.0994	.1665	.3668
Getting a R0 surgery according to TNM stages	$R0_{TNM}$	1.000	.957	.943	.059
Having a local recurrence (LR) (during the year of primary treatment) according to TNM stages, given a recurrence (R) and R0 surgery for the primary CRC	$LRofR_{TNM}$	.190	.151	.081	.081
Having a local and distant recurrence (LDR) (during the year of primary treatment) according to TNM stages, given a recurrence (R) and R0 surgery for the primary CRC	$LDRofR_{TNM}$	.048	.075	.121	.108
Having a distant recurrence (DR) (during the year of primary treatment) according to TNM stages, given a recurrence (R) and R0 surgery for the primary CRC	$DRofR_{TNM}$	.762	.774	.798	.811
Dying during the first month after being diagnosed with the first recurrence	$CD_{TNM+REC}$	.088	.107	.172	.536
Dying of CRC in the period two to twelve months after the primary treatment, given that the patient got a recurrence this year	CD_{REC30d}	.0422	.0422	.0422	.0422

$$tp_{0,70}^{II,LR} = 0.0994 * 0.957 * 0.151 * (1-0.0422-0.107) = 0.0122$$

To estimate the transition probabilities for the subsequent years of recurrence from the health state of, e.g., "disease free after II" (see Figure 1), another formula was applied. For instance, the transition probability of a local recurrence (LR) for a stage-II patient in year three in the "disease free after II" was given by

$$tp_3^{IIc,LR} = (REC_{II,3} - (REC_{II,3} * CD_{REC30d})) * LRofR_{II} \quad (5)$$

$REC_{II,3}$ was the probability of a patient getting a recurrence in stage II, given R0 surgery for the primary CRC and no recurrence until the end of the third year. By estimating $REC_{II,3} * CD_{REC30d}$,

the probability for the stage II patients with recurrence to die within 30 days was found. $REC_{II,3}$ was estimated by Equation (2) to be 0.04635. Inserting this result together with the parameters in Table 4 in Equation (5), the transition probability used in the model for stage-II patients moving from "*Disease free after II*" in year three to treatment of local recurrence during year four was given by

$$tp_3^{IIc,LR} = (0.04365 - (0.04365 * 0.0422)) * 0.151 = 0.00631$$

Table 5*. Transition probabilities from the first year of recurrence to other health states the next year. For the abbreviation in the second column, recall Box 1. Source: OUS.*

The probability of:	Abbrevi-ation	Type of recurrence the patient are leaving (REC)		
		Local recurr.	Local & distant	Distent recurr.
Dying of CRC the first year after recurrence	$tp_0^{REC,CD}$	.2234	.5934	.4334
Receiving palliative treatment the first and second years after recurrence	$tp_0^{REC,Pa}$	.6035	.3600	.3853
Being considered disease-free the first year after recurrence	$p_0^{REC,RECc}$	.1030	.0	.1097
Getting a local recurrence the first year after being treated for recurrence	$p_0^{REC,R-LR}$	.0019	.0	.0
Getting both a local and distant recurrence the first year after being treated for recurrence	$p_0^{REC,R-LDR}$	.0025	.0	.0
Getting a distant recurrence the first year after being treated for recurrence	$p_0^{REC,R-DR}$	.0191	.0	.0251

To estimate survival after recurrence, some simplifications were carried out because of the scarcity of data. Overall survival after recurrence (Table 2) was estimated, but due to lack of data, estimating re-recurrence and disease-free survival was impossible. As an approximation, time to recurrence and disease-free survival for Stage IV was used (Table 2) and adjusted by the difference between overall survival for recurrence and Stage IV.

The adjusted time to recurrence and disease-free survival was shown in Table 6.

When considering the changes in both disease-free survival and time to recurrence and re-recurrence in the states *"disease-free after TNM" and "disease-free after REC"*, tunnels of ten and six years were built respectively. An essential part of building the model was to use precise clinical endpoints, and the definitions of events (failures) and censoring of data were defined in Table 1.A 10-year time frame was chosen for the tunnel states after primary treatment; consequently, no recurrence was assumed to occur 11 years after diagnosis (year of primary treatment and 10 years into *"disease free after TNM"*). Further, time in the tunnel *"disease free after REC"* was limited to six years, as only a small fraction was left in the tunnel and the re-recurrence rate seemed to stabilize.

***Table 6**: Transition probabilities for patients moving through the model after being "disease free" for the first recurrence or later recurrences. "Disease free" means that there is still no sign of CRC after a R0 resection after recurrence (or later recurrences).*

The number of years "disease free" after the year of R0 resection for recurrence or later recurrence		Probability of moving to state of new recurrence (local = R-LR, distant = R-DR, both = R-LDR) or dying					From one year to the next in the "disease free tunnel" †
		R-LR	R-DR	R-LDR	CRC Mortality*	No CRC Mortality	
Move from	Disease free year 1	.0166	.1662	.0221	.0090	.0391	.7468
	Disease free year 2	.0162	.1618	.0215	.0088	.0444	.7473
	Disease free year 3	.0157	.1570	.0209	.0085	.0479	.7500
	Disease free year 4	.0152	.1521	.0203	.0083	.0510	.7532
	Disease free year 5	.0147	.1473	.0196	.0080	.0538	.7565
	Disease free years 6, 7, etc.	.0143	.1427	.0190	.0078	.0434	.7729

** : The estimates show the probability of getting a recurrence and dying within one month (see also the text).*

†: The estimates show the transition probabilities moving from one year of being disease-free to the next year of being disease-free (the probability of staying in the tunnel from one year to the next), that means the probability of staying "cured" another year.

2.2. Mortality

2.2.1 CRC death

CRC-death means death caused by colorectal cancer disease and death caused by CRC treatment. In the model, CRC deaths mainly occur in the year of treatment for primary diagnosed CRC or for recurrence and during the year with palliation. This means, if a patient got a recurrence in month four in year three after the primary treatment and died six months after CRC, then the model defines this as a recurrence, and the patient would move to the recurrence state for the next year, receive treatment, and die in that state of health (except those who die within 30 days after recurrence, as mentioned in the chapter above). The effect of this simplification is discussed in the article.

To estimate the overall survival curve for the patient receiving palliative chemotherapy, the average of four studies was used [3-6]. The percentages of patients surviving the first four years were 0.675, 0.350, 0.175, and 0.087, respectively. Equation (1) was used to estimate the transition probability of staying alive.

To estimate the transition probabilities from the primary treatment of recurrence (first year of palliative treatment) to the second year of palliative treatment, we used Kaplan-Meier on OUS data and estimated the overall survival for the three groups of recurrence patients. The parameters used for mortality the first year after recurrence are 0.211, 0.593 and 0.427 for patients with local (LR), local and distant (LDR) and distant recurrences (DR), respectively.

Based on OUS data, the probability of dying of CRC during the primary treatment year (according to stage) was estimated to be 0.0045, 0.0418, 0.0560 and 0.5305 for stage I, II, III and IV, respectively.

2.2.2 Non-CRC mortality

In the model, a distinction was made between mortality caused by CRC and all-cause mortality other than CRC. For the first 10 years after primary CRC-treatment, for a patient considered to be disease-free, the non-CRC mortality rate was calculated based on the OUS data according to stages I, II and III. The calculated mortality rate for the first years after primary CRC treatment was higher for 70-year-old patients considered to be disease-free than for cohorts of the same age in Norway. This could be attributable to the side effects of the treatment or co-morbidity by

other lifestyle-related diseases than CRC. At the end of the 10-year period, however, the mortality was 1.7 – 2.4% less for stages I-III than the normal rate, which could be attributed to a situation where the frailest persons died at the beginning of the period. After the 10-year period, this difference was subtracted from the relevant age-specific mortality rate collected from the Norwegian Life Table, and the result was used as non-CRC mortality for patients of ages 81-99. This age span was split into the age groups 81-83, 84-86, 87-89, 90-92, 93-95, 96-97 and 98-99. For stage IV, the non-CRC mortality was estimated to be higher than the normal population in the whole 10-year period. As an example, the non-CRC mortality for a patient of 70 years at the time of diagnosis who entered stage II and was disease-free for three years was given by

$$tp_{3,73}^{IIc,D} = 1 - tp_{3,73}^{IIc,REC} - tp_{3,73}^{IIc,IIc} - tp_{3,73}^{IIc,CD}$$

Where $tp_{3,73}^{IIc,CD} = REC_{II,3} * CD_{REC30d}$

The same approach was used to estimate the non-CRC mortality for those disease-free after recurrence, except that the tunnel-state period lasted for six years. The same mortality probability was then used for all three types of recurrence (Table 6).

The yearly probability of non-CRC death for patients in palliative chemotherapy treatment was assumed to be 4.66% the first 10 years and thereafter followed the same change in mortality risk as for those disease-free.

2.3. The data for the economic models

Major cost components in this model were diagnostics; primary treatment (the first year after diagnosis), including surgery, chemotherapy, radiotherapy and side-effect treatment; follow-up; treatment related to recurrence (first year after recurrence); and palliative treatment.

The health care cost per person per cycle depended on health states. By multiplying the cost for one patient staying one year in a health state by the number of patients staying in that specific health state for the same year, the total cost for all patients in each health state per year was estimated. The expected total CRC cost per patient was estimated by aggregating the cost over the total lifespans of the patients for all health states.
Discussions about treatment often concern colon or rectum cancer, not colorectal cancer. The model merges these two together. To analyze colon and rectum cancer separately, the present

model structure could be used for both, but different data for treatment cost, type of treatment and recurrence rate have to be applied.

2.3.1 Diagnostics

When colorectal cancer is suspected, investigations to confirm the diagnosis and staging of the disease are carried out. The costs related to these examinations are listed in Table 8. In the model, every patient was assumed to receive one unit of each type of examination, except for rectal ultrasound and MRI, which was only received by patients with rectal cancer.

2.3.2 Primary treatment cost

In the model, the first year of treatment included cost of preoperative examinations, cancer treatment, palliative treatment of patients with non-resectable synchronous metastatic disease and the initial part of the follow-up. The quantity per patient and the cost per unit for the different components in the treatment are listed in Table 7.

Generally, several of the frequencies were estimated by using a decision tree, where the distribution of CRC diagnosis between colon and rectum cancer was an important parameter. Based on OUS data, the percentage diagnosed with colon cancer at stages I, II, III and IV were 51.0, 68.1, 65.9 and 70.6, respectively.

The probability for a patient in stage II of receiving "colon resection with no complications (DRG 149)" was 0.401, while it was 0.023 for a patient in stage IV. The major treatment category for a patient diagnosed with a stage IV was "digestive malignancy with complications", DRG 172,[2] where the proportion of patients that received such treatment during the first year after diagnosis was 1.525 – i.e., patients in this category had more than 50% probability of receiving the treatment (DRG 172) more than once in the first year.

The unit cost for DRG rows no. 1-19 (Table 7) was based on DRG weights, while the unit cost for no. 20 and 21 was estimated on the basis of drug cost, time use, CT-scanning, and side effects.

The source of the frequency estimate for stages I-IV, rows five to six and eleven to thirteen (Table 7) were based on NPR data [2]. The other frequencies in rows 1-14 were based on OUS data.

The parameter for stage IV concerning metastasis (rows 15-18) was based on literature [11 12], Norwegian guidelines(24) and expert opinions. It was assumed that 0.5 of the stage-IV patients had metastasis in the liver, and 0.25 of these were eligible for resection. The equivalent parameters for metastasis in the lungs were 0.25 and 0.075, respectively. Further, resections for metastasis in any other organs were not included. Rows 17 and 18 indicate no-surgical supportive treatment and care and were adjusted upward to cover the all costs for treatment and care for metastasis.

The frequencies in rows 20 and 21 were based on data from literature and expert opinion [13-18].

2.3.3 Follow-up

After CRC diagnosis and primary treatment, the patients were allocated to regular *follow up,* which could be given during the year of primary treatment (year 0). The frequency for the different types of *follow-up* and the cost per-unit of follow-up for stages II, III and IV were shown in Table 8. For a patient treated for stage I, one unit outpatient consultation, CEA-test and colonoscopy during the primary treatment year after surgery, and one outpatient consultation and CEA-test annually during the ensuing five years was assumed. The follow-up frequency was based on national guidelines (24), and the rate of compliance was based on literature [19].

2.3.4 Recurrence

The data for treatment after recurrence was limited and not included in Table 7, but based on literature, similarities with the primary treatment shown in Table 7 were assumed. For a local recurrence, the frequencies of treatments described in rows one to four (Table 7) were estimated by assuming that 30% of all local recurrence underwent an attempt of curative resection of the recurrent tumor, and this was split between the colon and the rectum, as for stage III. The frequency of treatment related to rows 11-13 was assumed to be identical with stage III. For radiotherapy (row 19), the frequencies were assumed to be 0.215, based on the Norwegian Rectal Cancer Registry and literature [15]. For palliative chemotherapy (row 20), adjustments were made, as many patients do not receive palliative chemotherapy due to old age, co-morbidity and poor performance status [20]. The first period after local recurrence, the frequency for receiving palliative chemotherapy was in fact marginal, but some period after diagnosis, a proportion of these patients would receive palliative treatment. This proportion was assumed to be 0.49 and was used in the estimation of LR treatment cost.

For distant recurrence (DR), the frequencies were assumed to be identical with stage IV for rows 13 and 15-21. For patients with both local and distant recurrence (LDR), the same parameters as for stage IV were used, except that the following was assumed: there was (i) no major resection, (ii) the probability of getting palliative treatment was increased to 0.754, and (iii) frequencies in rows 19 and 21 were assumed to be zero.

2.3.5 Radiotherapy

To find the parameter for radiotherapy in Table 7, we used decision trees and split the CRC patients into those with colon cancer and those with rectum cancer and used literature and expert opinion to get estimates of radiotherapy use [13 16-18 20]. Further, for some of the parameters, we also had to consider that only patients with *resection* in stages I–III receive radiotherapy. The cost per fraction of radiotherapy is based on the DRG score, and patients were assumed to have 25 fractions each.

2.3.6 Adjuvant and perioperative chemotherapy

The parameter for adjuvant chemotherapy stages II and III was estimated the same way as for radiotherapy by dividing colon and rectum cancers. Literature and expert opinion was used to obtain estimates for chemotherapy use [21]. For rectum stage II, we assumed no adjuvant chemotherapy, according to Norwegian guidelines (24). A problem with these parameter values is the changes over time. The administration of adjuvant chemotherapy in stage-III patients older than 75 years increased from 19% in the years 1989–1993 to 79% in 2004–2006, and from 1% to 19% in these periods for stage-III patients 75 years or older [13].

Table 7*. Frequency per-patient and values per-unit for primary treatments, used within the base case model analysis. The frequencies show how many times the average patient with a certain diagnosis receives the treatment stated (see also the text). Treatment for recurrence is not included.*

Treatment first year after primary diagnosis (DRG, medical: M, surgical: S)	row no.	Primary treatment stage				Unit cost, (€)	Source
		I	II	III	IV		
Resection of primary tumor							
Colon resection, w (148, S)	1	.210	.280	.458	.443	23,913	OUS
Colon resection, n (149, S)	2	.300	.401	.192	.023	11,688	OUS
Rectal resection, w (146, S)	3	.267	.174	.218	.120	18,546	OUS
Rectal resection, n (147, S)	4	.221	.145	.119	.0	12,486	OUS
Non-resectional surgery							
Endoscopic therapy colon; closure stoma, w (152, S)	5	.0	.0	.045	.026	9,539	NPR
Endoscopic therapy colon; closure stoma, n (153, S)	6	.036	.036	.090	.026	6,758	NPR
Endoscopic therapy rectum; TEM, w (157, S)	7	.0	.0	.0	.101	5,519	OUS
Endoscopic therapy rectum; TEM, n (158, S)	8	.0	.0	.0	.034	2,748	OUS
GI obstruction, w (180, S)	9	.0	.0	.0	.044	3,939	OUS
GI obstruction, n (181, S)	10	.0	.0	.0	.015	2,140	OUS
Endoscopic/other treatment							
Digestive malignancy, w (172, M)	11	.0	.107	.493	1.526	7,526	NPR
Digestive malignancy, n (173, M)	12	.0	.0	.164	.184	4,409	NPR
Aftercare and rehabilitation (465)	13	.0	.0	.030	.553	6,207	NPR
Endoscopic insertion of stent to gastro. tract, short therapy (703O)	14	.0	.0	.0	.008	1,310	OUS
Treatment for metastasis							
Resection							
Liver metastasis resection, w (191B, S)	15	.0	.0	.0	.125	26,528	[19,20], [11,12]
Lung metastasis resection (75, S)	16	.0	.0	.0	.019	18,968	[19,12]
No-surgical supportive treatment and care							
Liver metastasis (203, M)	17	.0	.0	.0	.188	6,468	NPR, exp
Lung metastasis (82, M)	18	.0	.0	.0	.075	7,664	NPR, exp

(Continuing)

Treatment first year after primary diagnosis (DRG, medical: M, surgical: S)	row no.	Primary treatment stage				Unit cost, (€)	Source
		I	II	III	IV		
Chemo- and radiotherapy							
Radiotherapy (409E, M)	19	.033	.075	.147	.056	645 *	[15 20], exp
Palliative chemotherapy (M)	20	.0	.0	.0	.610	20,183†	[13 15-18]
Adjuvant chemotherapy (M)	21	.0	.054	.535	.05	8,677/ 7,494	[13 15-18]

w: with complications or co-morbidities
n: without complications or co-morbidities
exp: Expert opinion
OUS: Observational study at Oslo University Hospital – Aker
NPR: National Patient Register based on data organized by Aas [2]
*: Cost per visit at hospital for radiotherapy
†: The cost first year of palliative treatment

Perioperative chemotherapy was only assumed for stage IV, and based on expert opinions, 10% of the stage IV patients are assumed to receive this therapy [22].

For estimating the cost per therapy for stage III in Table 7, row 21, we assume nine rounds of oxaliplatin therapy (development of neurotoxicity) and 12 cycles of 5FU. We also assume that 50% receive 5FU and the other 50% receive the other therapy. For stages II and IV, we assume 12 rounds of therapy for both.

The cost of the drug from pharmacy is based on oncolex.org. Further, we took into account the cost of CT-scanning, complications, and time that the nurse, pharmacist and medical practitioner use when giving the therapy.

Table 8. *Examinations before surgery (all stages) and follow-up during primary treatment and the following years, compliance, cost per unit and data source.*

Examination	Before surgery	Follow up, stage II and III (stage IV)							Cost /unit (€)	Source unit cost
		Prim. treat.	Y.1	Y 2	Y 3	Y 4	Y 5	Comp		
Outpatient consultation	1	3 (3)	2(2)	2(2)	1(2)	1(2)	1(2)	1	308	DRG
CT abdomen/liver/ pelvis ("bekken")	1	1 (3)	(2)	(2)	(2)	(2)	1(2)	0.85	411	Marked
Colonoscopy	1	1 (1)					1	0.57	289	DRG
CEA-test	1	2 (3)	2(2)	2(2)	1(2)	1(2)	1(2)	0.63	16	Marked
Ultrasound, rectum	ReCa								128	Marked
MRI, rectum	ReCa								250	Marked
Biopsy	1								494	DRG
Proctoscopy		ReCa			ReCa	ReCa	ReCa	0.57	251	DRG
CT scan lungs			1	1	1	1		0.85	141	Marked
CT scan liver			2	2	1	1		0.85	193	Marked

"ReCa": only rectal cancer

"Marked": price per unit was based on the private market of health service in Norway

"Comp.": compliance for following up

2.3.7 Palliative chemotherapy

The structure of the data for palliative chemotherapy treatment required that a decision tree (Figure 5) be used to estimate costs according to treatment paths before being distributed to the treatment years in the Markov model. In Figure 5, number at each branch indicated the conditional probability, and the number in the brackets was the joint (total) probability for obtaining a certain type of treatment for a patient starting with palliative chemotherapy treatment. As an example, it was assumed that 71% of the patients have good health (high PS) and obtained first-line palliative chemotherapy treatment. Of these patients, 40% receive bevacizumab, together with FLIRI or FLOX, which constitutes 28.4% (0.71 * 0.4 * 100%) of all patients receiving palliative chemotherapy. Of the others, 30% received FLIRI and 70% FLOX [23]. Population-based studies shows that approximately one third of mCRC patients do not receive palliative chemotherapy at all [13 16-18 20]. In the model, 61% of stage-IV patients received palliative chemotherapy[13 16-18 20]. Of the patients receiving first-line treatment, 60% received second-line palliative chemotherapy treatment [4]. Of these, 60% will be KRAS wild types and suitable for EGFR (epidermal growth factor receptor) inhibitor treatment.

For each treatment in the decision tree, separate cost models took into account the cost of the drug, CT-scanning, complications, and time the nurse, pharmacist and medical practitioner use when giving the therapy. Table 9 shows the cost of components for the different palliative chemotherapy treatments. The medicine costs include all costs at the pharmacy (drug, time use, equipment, etc.) and were derived from oncolex. The cost of 5-FU/FA (5-fluororuracil/folinic acid) was based on the Nordic 5FU/FA schedule (Nordic Flv) [6]. The costs related to CT, time usage and treatment intensity were derived from literature [24]. Unit costs for side effects (excl. nausea) were derived from the DRG system 2011. The model is corrected for non-compliance and withdrawal from the chemotherapy treatment.

The cost model for palliative treatment above has no timeline, which complicated the discounting of the costs. As a simplification, we have distributed the total costs for the three lines of treatment over a three-year period and then summarized the total costs for all three lines. The total palliative treatment costs were €35,880 and were distributed to each treatment year (one, two and three) with the weights of 56.3, 34.9 and 8.9%, respectively.

Table 9*. The costs of different components of the palliative chemotherapies (euro).*

Components in the treatment	5-FU/FA	Bevacizumab + FLIRI	Bevacizumab + FLOX	FLIRI		FLOX		EGFR + irinotecan
				1. line	2. line	1. line	2. line	
Medicine (from pharmacy)	3,081	32,734	31,772	5,789	4,211	5,083	3,697	30,873
Administered in hospital	479	1,197	1,197	439	319	878	638	2,154
CT-scanning	1,029	1,179	1,179	943	686	943	686	2,831
Out-patient consultation	1,497	1,834	1,834	1,384	1,047	1,384	1,047	2,171
Side effects*	1,121	1,699	1,699	1,072	925	1,072	925	1,548
Sum cost	7,206	38,643	37,681	9,627	7,188	9,360	6,993	39,577

* Side effects include sepsis, intestine perforation, arterial thromboembolism and medicine for nausea. Diarrhea is included in another part of the model.

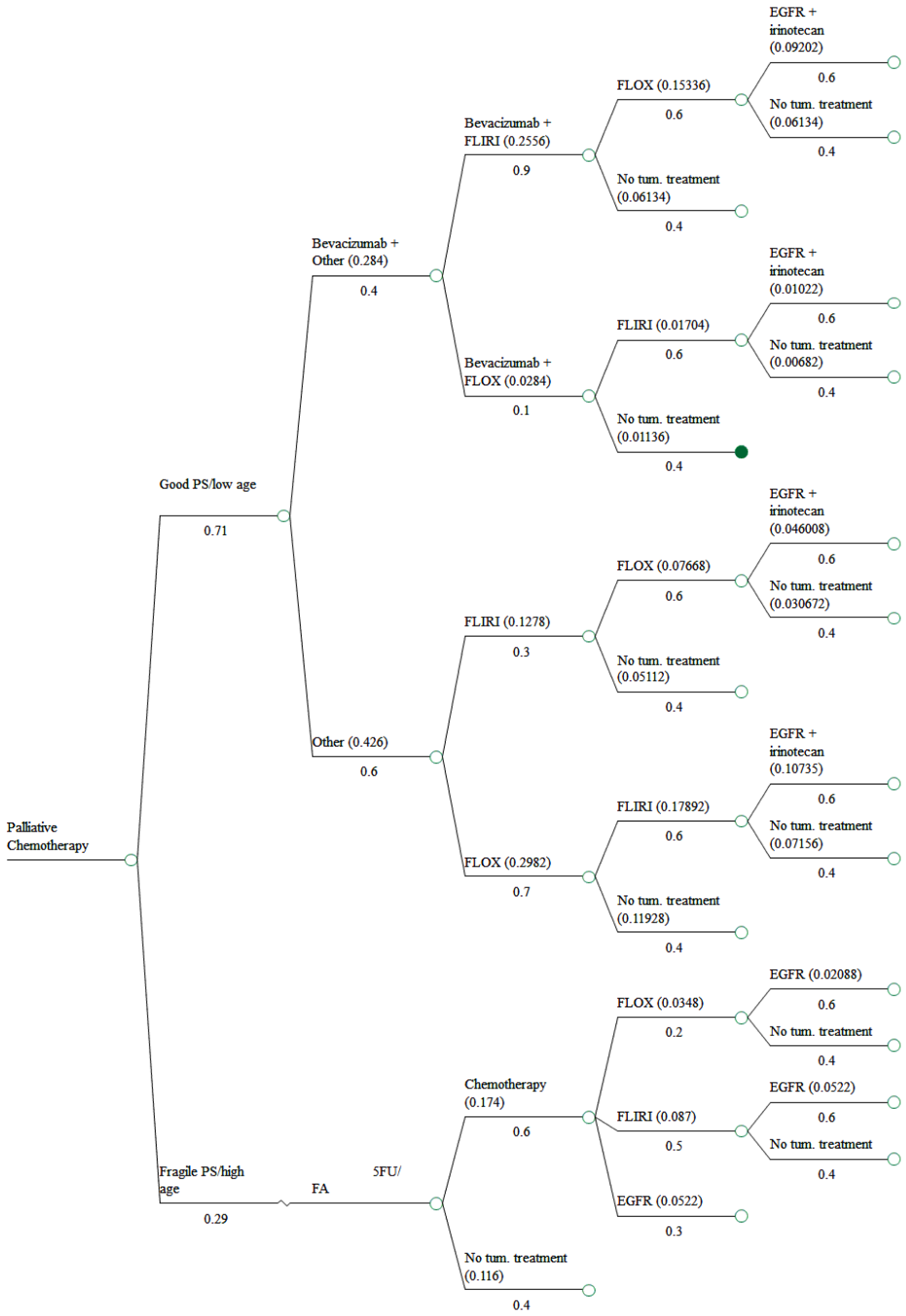

***Figure 5**. The decision tree for patients treated with palliative chemotherapy (61% of all stage-IV patients).*

References

1. Aas E. Cost-effectiveness of screening for colorectal cancer with once-only flexible sigmoidoscopy and faecal occult blood test. HERO On line Working Paper Series: Oslo University, Health Economics Research Programme, 2009.
2. Folprecht G, Cunningham D, Ross P, Glimelius B, Di Costanzo F, Wils J, et al. Efficacy of 5-fluorouracil-based chemotherapy in elderly patients with metastatic colorectal cancer: a pooled analysis of clinical trials. Annals of Oncology 2004;15(9):1330-38.
3. Tournigand C, André T, Achille E, Lledo G, Flesh M, Mery-Mignard D, et al. FOLFIRI Followed by FOLFOX6 or the Reverse Sequence in Advanced Colorectal Cancer: A Randomized GERCOR Study. Journal of Clinical Oncology 2004;22(2):229-37.
4. Goldberg RM, Sargent DJ, Morton RF, Fuchs CS, Ramanathan RK, Williamson SK, et al. Randomized Controlled Trial of Reduced-Dose Bolus Fluorouracil Plus Leucovorin and Irinotecan or Infused Fluorouracil Plus Leucovorin and Oxaliplatin in Patients With Previously Untreated Metastatic Colorectal Cancer: A North American Intergroup Trial. Journal of Clinical Oncology 2006;24(21):3347-53.
5. Glimelius B, Sørbye H, Balteskard L, Byström P, Pfeiffer P, Tveit K, et al. A randomized phase III multicenter trial comparing irinotecan in combination with the Nordic bolus 5-FU and folinic acid schedule or the bolus/infused de Gramont schedule (Lv5FU2) in patients with metastatic colorectal cancer. Annals of Oncology 2008;19(5):909-14.
6. Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C, et al. Principles of Good Practice for Decision Analytic Modeling in Health-Care Evaluation: Report of the ISPOR Task Force on Good Research Practices—Modeling Studies. Value in Health 2003;6(1):9-17.
7. Weinstein MC. Recent Developments in Decision-Analytic Modelling for Economic Evaluation. PharmacoEconomics 2006;24:1043-53.
8. Punt CJA, Buyse M, Köhne C-H, Hohenberger P, Labianca R, Schmoll HJ, et al. Endpoints in Adjuvant Treatment Trials: A Systematic Review of the Literature in Colon Cancer and Proposed Definitions for Future Trials. Journal of the National Cancer Institute 2007;99(13):998-1003.
9. Briggs AH, Claxton K, Sculpher M. Decision modelling for health economic evaluation. Oxford: Oxford University Press, 2006.

10. SøreideJ A, Eiriksson K, Sandvik O, Viste A, Horn A, Johnsen G, et al. Kirurgisk behandling av levermetastaser fra kolorektal kreft. Tidsskr Nor Legeforen 2008;128(1):50-53.
11. Kobayashi H, Mochizuki H, Sugihara K, Morita T, Kotake K, Teramoto T, et al. Characteristics of recurrence and surveillance tools after curative resection for colorectal cancer: A multicenter study. Surgery 2007;141(1):67-75.
12. van Steenbergen L, Elferink M, Krijnen P, Lemmens V, Siesling S, Rutten H, et al. Improved survival of colon cancer due to improved treatment and detection: a nationwide population-based study in The Netherlands 1989-2006. Annals of Oncology 2010;21(11):2206-12.
13. Scholefield JH, Moss SM, Mangham CM, Whynes DK, Hardcastle JD. Nottingham trial of faecal occult blood testing for colorectal cancer: a 20-year follow-up. Gut 2012;61(7):1036-40.
14. Delcò F, Egger R, Bauerfeind P, Beglinger C. Hospital health care resource utilization and costs of colorectal cancer during the first 3-year period following diagnosis in Switzerland. Alimentary Pharmacology & Therapeutics 2005;21(5):615-22.
15. Lemmens V, de Haan N, Rutten H, Martijn H, Loosveld O, Roumen R, et al. Improvements in population-based survival of patients presenting with metastatic rectal cancer in the south of the Netherlands, 1992–2008. Clinical and Experimental Metastasis 2011;28(3):283-90.
16. Khattak MA, Townsend AR, Beeke C, Karapetis CS, Luke C, Padbury R, et al. Impact of age on choice of chemotherapy and outcome in advanced colorectal cancer. European Journal of Cancer 2012;48(9):1293-98.
17. Jacob S, Ng W, Asghari R, Delaney GP, Barton MB. Estimation of an optimal chemotherapy utilisation rate for colon cancer: An evidence-based benchmark for cancer care. European Journal of Cancer 2009;45(14):2503-09.
18. Körner H, Söreide K, Stokkeland P, Söreide J. Systematic follow-up after curative surgery for colorectal cancer in Norway: A population-based audit of effectiveness, costs, and compliance. Journal of Gastrointestinal Surgery 2005;9(3):320-28.
19. Sorbye H, Pfeiffer P, Cavalli-Björkman N, Qvortrup C, Holsen MH, Wentzel-Larsen T, et al. Clinical trial enrollment, patient characteristics, and survival differences in prospectively registered metastatic colorectal cancer patients. Cancer 2009;115(20):4679-87.
20. Elferink MAG, van Steenbergen LN, Krijnen P, Lemmens VEPP, Rutten HJ, Marijnen CAM, et al. Marked improvements in survival of patients with rectal cancer in the

Netherlands following changes in therapy, 1989–2006. European journal of cancer (Oxford, England : 1990) 2010;46(8):1421-29.
21. Sjövall A, Järv V, Blomqvist L, Singnomklao T, Cedermark B, Glimelius B, et al. The potential for improved outcome in patients with hepatic metastases from colon cancer: a population-based study. European Journal of Surgical Oncology (EJSO) 2004;30(8):834-41.
22. Wolin K, Yan Y, Colditz G, Lee I. Physical activity and colon cancer prevention: a meta-analysis. British journal of cancer 2009;100(4):611-16.
23. Aaserud M KIS, Neilson A , Norum J , Sørbye H , Aas E , Gjertsen M. . Helseøkonomisk evaluering av bevacizumab ved metastatisk kolorektalcancer. Rapport fra Kunnskapssenteret Oslo: Kunnskapssenteret, 2007.

Appendix 2 (paper II): More on application of the model

(App. 2: Published in a Web-only format)

Some applications and advantages of the model should be emphasized. First, the most obvious and direct use of the model is to estimate the cost and survival time for an average CRC patient according to disease stage. The CRC costs can be divided into different cost components, such as primary treatment, follow-up and palliative treatment. Different survival distributions can be estimated by using different endpoints and can be performed with stratification by stage and R-classification after the primary treatment or after different types of recurrence.

Second, the model can *estimate final outcomes from change*s in intermediate outcomes. The model estimates changes in costs and survival by applying different rates of recurrence or mortality, such as a decline in recurrence and mortality rate due to improvements in preoperative diagnostics, surgery and other treatment modalities for patients treated at a specific hospital. Incremental costs per patient due to marginal changes in resource use can also be estimated, such as an increased use of bevacizumab therapy or increasing unit prices, e.g., the price for drugs. Based on intermediate outcomes from randomized controlled trials - like the percentage of the population diagnosed with CRC; distribution between stages; recurrence or the survival rate after, for instance, three or five years; or relative risk - the model can estimate final outcomes like treatment costs and overall survival during the lifespan until the age of 100 years.

Thirdly, the model can be used in *economic evaluations*. By applying modest adjustments and further developments, the model is suitable for performing economic evaluations of different types of screening and prevention and follow-up. The model can also evaluate the effects of present or future variations in treatment strategies, including new surgical techniques and technology, an increased and changing use of chemotherapy, the indication of a treatment shift, increased treatment of elderly and the cost of implementing a new drug treatment. The general structure of the model enables comparative analysis of different types of CRC interventions within the same model, like comparing CRC screening with curative interventions. Almost all published cost-effectiveness analyses of screening compares one kind of CRC screening with another and not with other kinds of interventions in the health service.

Fourthly, the model can *estimate resource use*. By including the use of resources, like labor, instruments, blood, medication and beds for each procedure, the model could be used to predict the need for extra personnel or instruments for new or extended CRC treatment.

Fifthly, the model can *adjust for change in parameters over time (time-dependency)*. The model can simultaneously take into account the time since CRC treatment, consequences of the age of the CRC patients (mortality, recurrence, primary treatment cost, etc.), progress in the treatment (changes in the recurrence rate and survival, etc.) and change in cost and resource-use over time. For this, we use eight tunnels and a three-dimensional data matrix.

Sixthly, the model is *transferable* to other countries that have access to the same types of data, like the OUS data. Calibration has not been used in this model; thus, applying data from another country and building the model with the recommendations and assumptions in this article, the model should, in principle, have a similar goodness of fit.

Additionally, by using the widespread software Excel, modified versions of this type of model could also be used for other purposes. E.g. *decision makers* could use the model by changing the model inputs and gain preliminary insight about potential health benefits and costs of new emerging treatment strategies.

Paper III

Survival and costs of colorectal cancer treatment and effects of changing treatment strategies - a model approach

Paal Joranger, Arild Nesbakken, Halfdan Sorbye, Geir Hoff, Arne Oshaug, Eline Aas.

Paal Joranger, associate professor, Norwegian University of Life Sciences, Norway/Faculty of Health Sciences, Oslo and Akershus University College of Applied Sciences, NO-0130 Oslo, Norway.
Telephone: 0047 92293776/0047 64849145. E-mail: paal.joranger@hioa.no

Arild Nesbakken, professor, Dept of Gastrointestinal surgery, Oslo University Hospital, NO-0424 Oslo, Norway/K.G. Jebsen Colorectal Cancer Research Centre, Oslo University Hospital, NO-0424 Oslo/Norway Faculty of Medicine, University of Oslo, Norway/Centre for Cancer Biomedicine, University of Oslo, Norway. Telephone: 0047 91141398/ 0047 22117491.
E-mail: arild.nesbakken@medisin.uio.no

Halfdan Sorbye, professor, Dept of Oncology, Haukeland University Hospital, Bergen, Norway/University of Bergen, NO-5020 Bergen, Norway. Telephone: 0047 55973589. E-mail: halfdan.sorbye@helse-bergen.no

Geir Hoff, professor, Telemark Hospital, Norway/The Cancer Registry of Norway, Oslo, Norway/University of Oslo, NO-0316 Oslo, Norway. Telephone: 0047 91866762.
E-mail: hofg@online.no

Arne Oshaug, professor, Faculty of Health Sciences, Oslo and Akershus University College of Applied Sciences, NO-0130 Oslo, Norway. Telephone: 0047 99458142/0047 64849193.
E-mail: arne.oshaug@senior.hioa.no

Eline Aas, associate professor, Department of Health Management and Health Economics, Institute of Health and Society, University of Oslo, Oslo, Norway. Telephone: 0047 41141924/0047 22845036.
E-mail: eline.aas@medisin.uio.no

Correspondence

Correspondence to: Paal Joranger
Email: paal.joranger@hioa.no
Telephone: 0047 92293776/0047 64849145.

Keywords

Cost Effectiveness, Colorectal Cancer, chemotherapy, surgery, prevention.

Word count for the text of the manuscript:

Abstract: **333** words. Main text: **4,060** words.

Guarantors of the paper

PJ and EA are guarantors of the paper.

Ethics committee approval

The observational study from 1993 to 2010, including 2049 patients diagnosed with CRC at Oslo University Hospital was approved by the Regional Ethics committee (Norway) for Medical Research (REK approval 1.2005.1629).

The study collecting data based on National Patient Register (referred to as NPR data) was approved by Regional Ethics committee (Norway). The reference number is **S-02113 (2013/83)**

Funding and all other required statements

Financial support for this study was provided entirely by the authors' employer mentioned on the front page. The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report.

Data sharing

The Markov model and the data used are presented in a separate article [1].

Competing interest statement

All authors have completed the ICMJE uniform disclosure form at http://www.icmje.org/coi_disclosure.pdf and PJ, EA, GH and AO declare: no support from any organisation for the submitted work [or describe if any]; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years [or describe if any], no other relationships or activities that could appear to have influenced the submitted work [or describe if any].

HS reports grants and personal fees from Merck, grants and personal fees from Roche, grants and personal fees from Amgen, personal fees from Sanofi, during the conduct of the study.

AN has received funding from Helse Sør-Øst for the clinical studies on colorectal cancer (OUS-Aker series), and joins a research Group in OUS who has patent on one diagnostic and two prognostic Genetic tests for colorectal cancer.

Transparency declaration

The lead authors* affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

"The paper adds box"

What is already known on this subject

- The pricing of a new generation of cancer drugs, in combination with limited health care resources, has highlighted the need for improving the methodology to estimate outcomes and economical effects of different treatment options.
- Compared to estimating the cost of CRC treatment empirically, which is the dominant way to do it up to now, model-based CRC cost estimates have several advantages, and to the best of our knowledge this is the first general CRC model estimating both treatment cost and survival.

What this study adds

- The presented and used CRC model is flexible and capable for modifying many aspects of health care costs for CRC treatment, such as prices, type and intensity of treatment and follow up, recurrence rates, CRC and non-CRC mortality.
- Health care cost of successful CRC treatment can be many times higher than unsuccessful, and changed palliative treatment practice can increase the average CRC cost substantially.
- The cost of CRC treatment appears to be modest compared with the years at stake (alternative cost), and reducing recurrence rate by better surgery and achievable preventive efforts like screening of asymptomatic persons, could have a considerable cost-effectiveness potential.

Abstract

Background: Colorectal cancer (CRC) is the third most common cancer in the world and the cause of major morbidity and mortality for many patients with high costs for the health care system. The aim of this study was to estimate the total lifetime health care cost of CRC treatment and survival and explore the consequences of altered prevention or surgical and medical treatment.

Methods: We applied a semi-Markov model with 70 health states and tracked patients' age and time since specific health states. The parameters were based on an observational study (2 049 CRC patients), the National Patient Register, literature and expert opinions.

Results: According to our model, the cost for an average CRC patient was €41 550, which varied from €23 390 to €61 400 depending on the disease stage at diagnosis. The cost of CRC was much higher in patients with both a recurrence after primary surgery and receiving bevacizumab as part of the palliative treatment, with €116 100 and €137 470 for stages I and IV, respectively. A 20% cost change for palliative drugs have only a minor effect on average CRC costs (< 2%), while altered indications for use of palliative chemotherapy increased the cost by up to 29%.
A 5% reduction in recurrence for stages I-III would reduce the cost by €2 280 per patient (5.5%) and increase the overall survival by 0.80 year per patient. This could be attained by e.g., better surgery or possibly by post-cancer lifestyle interventions. Applying the suggested threshold for a QALY gained, the health sector's willingness-to-invest in a 5% reduction in recurrence rate would be €61 306 per average CRC patient and €506 380 per average CRC case prevented in the first place.

Conclusions: Cost of CRC treatment appears to be modest compared with the years at stake. Altered decisions about palliative treatment can increase the average CRC cost substantially. Reducing the recurrence rate by better surgery and achievable preventive efforts like screening of asymptomatic persons could have a considerable cost-effectiveness potential.

1. Introduction

Cancer is a major cause of morbidity and mortality in the Western world, with colorectal cancer being the second most common cancer in women and third in men[25]. Norway is among the countries in the world with the highest incidence of CRC, higher than any other Nordic country [26]. The treatment of colorectal cancer is becoming a significant financial burden to health-care systems within economically developed countries. A current challenge for oncologists and health-care payers is the integration of new, often high-cost, therapies into clinical practice. The pricing of a new generation of cancer drugs, in combination with limited health care resources, has highlighted the need for improving-the methodology to estimate the outcomes of different treatment options. Inherent to this process is the consideration of cost-effectiveness. Several studies have analysed the cost-effectiveness of new medical treatments [27-32], screening [2 33-35], surgical techniques and perioperative care [36-39] for CRC.

There are, however, wide variations in the clinical management of CRC patients with advanced disease [40]. Descriptions of current treatment pathways are necessary for economic evaluations. Variations in clinical practice must be reflected in a model to ensure that findings from an economic evaluation are sufficient to inform policy regarding an optimal use of resources. Increases in CRC prevalence in the future, combined with scarce resources, highlight the importance of such studies for informing health-care policy and program planning. In our present CRC model, all aspects of health care costs for CRC treatment are included and could be modified, such as prices, type and intensity of treatment, recurrence rates, and CRC and non-CRC mortality. In a recent paper, the model was presented and validated [1]. This model is useful for (i) estimating CRC costs and survival, (ii) estimating final outcomes from intermediate outcomes, (iii) performing health economic evaluations, (iv) estimating resources required, (v) capturing the effect of duration since diagnosis on age-related mortality and recurrence, (vi) estimating the cost of changing treatment strategies, and (vii) calculating the opportunity cost of reducing the investment in CRC care.

The aim of the present article was to use the Decision Analytic Model developed by Joranger et al [1] to estimate the cost and survival of colorectal cancer treatment and explore the potential consequences of altered treatment or prevention strategies.

2. Methods

2.1 The model

In brief, the estimation of costs and survival in this paper is based on a semi-Markov model, and details are published in Joranger et al[1]. The flow of CRC patients was simulated through the model from CRC diagnosis, through periods of treatment and healthy periods until the patients were 100 years of age or had died from CRC or other causes (Figure 1). Each arrow reflects the probability of an average CRC patient moving from one health state to another during one cycle or maintaining in the same health state (follow the loops). In the model, the duration of one cycle is set to one year. For each health state, there was a cost model estimating the cost of the health service provided per person per year.

(Figure 1)

We estimated total CRC costs and the survival of an average CRC patient diagnosed at the age of 70 years.

The present version of the model has the health sector budget perspective. Included in the costs were all CRC treatments, including diagnostic and staging investigations, surgery, follow up, treatment for complications, treatment of recurrence and advanced disease (including radiotherapy and chemotherapy), and visits to general practitioners.

The patient *enters* the model at the time of primary diagnosis in one of the TNM stages (I, II, III and IV), and the first step includes the cost of primary work-up and treatment during the first year after diagnosis. The following year, the patient may move to the health state "disease free", which means that the tumour has been resected, and there is neither evidence of macro- or microscopic residual tumour locoregionally nor clinical or radiological evidence of distant metastases. Alternatively for this year, the patients may move to palliative care, recurrence, or to one of the two death states.

For most patients with stage IV, the treatment intention is palliative. A large proportion of these patients do not receive any specific anti-cancer therapy but the best supportive care until entering "Dead by CRC".

A separate decision tree (Figure 2) for estimating the cost of palliative chemotherapy was developed and included in the Markov Model (Figure 1). The first number at each branch indicated the conditional probability, and in the brackets the jointly (total) probability for obtaining a certain type of treatment for a patient starting with palliative chemotherapy[1]. For each treatment in the decision tree, separate cost models were developed, which included the cost of drugs, CT scanning, complications, and the time spent on therapy by the nurse, pharmacist and medical practitioner. The model adjusted for non-compliance and discontinuation of chemotherapy.

The majority of the patients entering one of the three recurrence stages received palliative chemotherapy. Some patients received resection with curative intent, adjuvant chemotherapy, radiotherapy or only the best supportive care. The probabilities of receiving the treatments depended on the type of recurrence, either local recurrence, distant recurrence or both.

Time-dependency in the calculation of probabilities for recurrences or death was captured in the model by including tunnel states.

We applied a 4 percent discount rate for the cost and both zero and 4 percent for overall survival [41 42]. All cost results were in 2011 euros (used average 2011: 1 euro = 7.79 NOK).

2.3 Data sources

We used Norwegian data when possible. An important source of data for estimating recurrence, disease-free survival, and surgical treatments was an observational study from 1993 to 2010, including 2049 patients diagnosed with CRC at Oslo University Hospital – Aker (referred to as OUS data) [43 44]. Information from the National Patient Register (referred to as NPR data), based on data (years 2003 and 2004) related to an analysis by Aas [2], was used to quantify some types of treatments. Other data sources were national life tables, international published papers (such as overall survival for patients receiving palliative chemotherapy) and expert opinions by one colorectal surgeon, one oncologist and one gastroenterologist. Calibration is often used as a

complement to data sources, but it was concluded that calibration was not needed due to a good fit of the model [1].

The estimation of the rate of recurrence, disease-free survival and overall survival was based on an average CRC patient diagnosed at 70 years of age. The treatment and resource use data were based on an average CRC patient, normally diagnosed at the age of 55-85 years. All cost estimates included colon and rectum cancer jointly but were weighted by their share of CRC patients according to the TNM stage.

2.3 Validation and uncertainty analysis

The model validation by Joranger et al. [1] for face-, internal-, cross- and external validity showed a satisfactory match with other models and real-life observations for both cost and survival time, without any preceding calibration of the model.
We used a one-way and multi-way sensitivity analysis for exploring parameter-, methodological- and model-structure uncertainty.

3. Results

3.1 Stage cost

The total lifetime health care CRC cost and loss-of-life years were reported on average for all patients and according to stages (Table 1). Based on our model, a 70-year-old patient has an expected lifetime CRC cost of €41 550. The expected cost increased with the TNM stage at €23 390, €33 500, €49 900 and €61 400 for stages I, II, III and IV, respectively. Table 1 shows the costs and survival for the *base case* scenario, which later will be compared with changing treatment strategies.

3.2 Type and phase of treatment

The treatments with the greatest impact on total lifetime costs (Table 1) were the operation of the primary tumor (€17 910) and palliative chemotherapy (€9 590). Costs related to diagnostic

examinations, adjuvant treatment and follow-up in general were modest for all stages. For stage IV, the costs for "surgery – major resection" (€16 890), "surgery - other" (€19 030) and palliative chemotherapy (€22 190) were dominating. "Surgery – major resection" was the major cost component for stages I and II. Variations between stages depend on differences in treatment, the mix of colon and rectum cases, and the proportion having cancer recurrence.
When categorizing treatment costs according to the clinical pathway, starting with primary examinations and ending with palliative chemotherapy (Table 1), costs varied according to TNM stage at the time of diagnosis. A stage IV patient had the highest expected cost both for primary treatment (€36 940) and palliative treatment (€22 190), while a stage III patient had the highest expected cost of treatment for recurrence (€5 590).

3.3 High- and low-cost scenarios

Above, we used average cost estimates, i.e., the average of all treatment scenarios that the CRC population could enter. Treatment costs for a patient could also be estimated conditioned on certain low- and high-cost scenarios. A stage I patient receiving a major resection without (radio) chemotherapy and who did not experience recurrence, represented a normal low-cost scenario, with €14 490 in expected treatment costs. The high-cost scenario is represented by patients experiencing recurrence after surgery and receiving the maximum of palliative chemotherapy with bevacizumab during the last year of palliative treatment. Expected costs for such patients were on average €127 930 and €137 470 for a stage I and IV patients respectively (Table 1). This finding indicated that a relatively common high-cost treatment scenario was nine times more costly than a relatively normal low-cost treatment scenario.

3.5 Survival and loss of years

According to the model, life expectancy for an average CRC patient diagnosed at the age of 70 years was 9.3 years (7.0 years assuming a 4% discount), implying 6.3 years lost (4.1 years, discounted) compared to an average 70-year Norwegian (Table 1). Life expectancy for a stage I patient was 14.0 years (1.6 years lost), compared to 1.5 years (14.1 years lost) for a stage IV patient.

(Table 1)

3.6 Cost of palliative chemotherapy scenarios

Palliative chemotherapy is an important and increasing cost component as new and expensive treatment protocols are introduced. Table 2 reports the estimated cost of various full-treatment scenarios based on the palliative treatment strategies presented in Figure 2. FLIRI is an irinotecan-based and FLOX an oxaliplatin-based chemotherapy schedule, and 5-FU/FA (5-fluororuracil/folinic acid) is based on Nordic 5FU/FA treatment protocol (Nordic Flv). The cost difference between the full treatment scenario "*5-FU/FA (1st line) and EGFR-inhibitor +irinotecan (2nd line)"* and the scenario *"bevacizumab and FLIRI (1st line), FLOX (2nd line) and EGFR-inhibitor +irinotecan (3rd line)"* is €38 430 (€46 780 versus €85 210), where the strategy with bevacizumab represents the latter. Further, we found that using "*bevacizumab and FLIRI*" in the 1st line instead of using only "*FLIRI*" implied an extra cost of approximately €29 010 (€85 210 versus €56 200).

(Figure 2)

For the average patient receiving palliative chemotherapy, the most expensive part of the treatment scenarios was the combined treatment of "*bevacizumab and FLIRI*" in the 1st line (€9 880), "*FLOX*" in the 2nd line (€1 070), and the "*EGFR-inhibitor + irinotecan*" in the 3rd line (€3 640). These costs for the average CRC patient were combinations of the price of the treatment regime and the probability of receiving the treatment.

(Table 2)

3.6.1 Altered choice of chemotherapy schedule

To show the importance of uncertainty in the input data and the possible impact of future decisions, we estimated the effect of changes in both prices and probabilities (Appendix 2). Most sensitive to changes in treatment costs were the "EGFR (cetuximab/panitumumab) + irinotecan treatment" and "bevacizumab + *FLIRI* treatment" protocols.

The use of bevacizumab is changing in Norway. In the model, we assumed that 29 percent of patients on palliative chemotherapy were treated with this drug. If all these patients were to receive bevacizumab (scenario 1 in Table 3), then the total cost for an average CRC-patient would increase 13.8 percent (€5 730). This change in treatment regimens would increase the treatment costs in Norway by €20.8 mill per year (3 600 diagnosed CRC patients per year) and €4.16 per capita per year. If no one received bevacizumab, the cost would decrease by 5.4 percent

(€2 240 per patient), and the Norwegian health sector expenditure would decrease by €8.2 mill (€1.67 per capita) according to the model. If those receiving *FLIRI/FLOX* in the 1st line of treatment instead received "*bevacizumab* + *FLIRI/FLOX*", the cost would increase by 8.1 percent (€3 360) per patient and increase the health sector expenditure in Norway by €12.2 mill (€2.43 per capita).

3.6.2 Increased use of chemotherapy in the elderly

Colorectal cancer is common in elderly patients, and approximately 40 percent of the patients are 75 years or older. The number of elderly patients diagnosed with CRC is increasing, and studies suggest a more "fit" population of elderly in the future [45]. Elderly patients with CRC have a poorer outcome than younger patients, prescription of chemotherapy is inversely associated with age [46-48], and when receiving chemotherapy, combined chemotherapy is less-often prescribed to the elderly [20]. There is an ongoing discussion as to whether elderly patients are undertreated or not [20 46 49].
(Table 3)

What would be the effect on CRC costs of treating a greater number of older patients with palliative chemotherapy? One extreme scenario would be to assume that everybody would receive palliative chemotherapy (Scenario 5, Table 3). If we use the current chemotherapy pattern of prescription, the cost for an average CRC patient would increase by 9.4 percent (€3 910). If all of these patients were to receive bevacizumab in the 1st line of treatment (scenario 6), then the total cost for an average CRC patient would instead increase by 28.8 percent (€11 970).

3.7 Reduced recurrence rate

Reduction in recurrence rates would affect both survival and health care CRC costs and might be achieved by better training of the surgeon, new surgical techniques, concentration of CRC treatment to fewer centres with robust multidisciplinary teams or by better methods to find high-risk patients who need adjuvant chemotherapy [44 50-52]. In a meta-analysis, a volume-outcome relationship in colorectal cancer surgery was found, based on hospital and surgeon caseload and specialization [50]. Further, a relationship between surgeon experience and local recurrence for rectal cancer has also been shown [44].

According to the model, a 5 percent reduction in a 10-year recurrence rate (from 32.5 to 27.5 percent) for stage I-III would reduce the cost by €2 280 per patient (-5.5 percent) and increase overall survival by 0.80 year (it would be 0.64 year if we discount overall survival by 4% per year). This would imply 2 320 life years saved per year in Norway. According to the cancer registry of Norway, 3 624 were diagnosed with CRC in 2009, and 79.8 percent were diagnosed with stage I, II or III disease (OUS-data). A 5 percent reduction in recurrence would then reduce the health cost for Norway by €6.60 mill per year (€1.31 per capita). In Norway, the threshold willingness-to-pay (WTP) for a QALY gained has been suggested to be €73 783 per QALY gained (in EURO 2011). Given this threshold, the society's willingness to invest in interventions that could contribute to a 5% reduction in recurrence was €61 306 per average CRC patient (€2 280 + [0.80 year * €73 783 per year) in stage I, II or III (or €49 110 when survival was discounted by 4% per year). In total, this sums up to €177 mill per year (€61 306 per patient * 3 624 patients per year * 0.798 in stage I, II or II) and €28.2 per capita.

3.8 Primary prevention

Primary prevention of CRC might be achieved by screening for and removing precursor lesions, by physical activity, modifications in the diet and lifestyle including smoking cessation and prevention of weight gain, and by using anti-inflammatory drugs. Primary prevention reduces the number of CRC cases in all stages. The outcome of preventive intervention for CRC can be estimated by means of the model. The cost saving for the health sector per CRC case prevented is estimated in Table 1 to be €41 550. Additionally, the average CRC patient will lose 6.3 years of life (4.1 years 4% discounted), according to the model. Given the threshold, society's willingness-to-invest in preventive intervention is estimated to be €506 380 per CRC case prevented (€41 550 + [6.3 year * €73 783 per year]) and €344 060 per CRC case prevented if both cost and survival are discounted by 4%.

3.9 Screening – gain from stage migration

There have been randomized controlled trials on CRC screening in several countries. We used the model to analyze the effect when using the result from Denmark and the UK [14 53]. In both trials, faecal occult blood tests (FOBT) were used to discover cancer in an earlier, asymptomatic stage; reduce the cost; and increase the overall survival. Table 4 shows that CRC patients diagnosed in a screening program have a more favourable stage distribution than those in the control groups. The stage migration effect was more pronounced in the UK trial than in the

Danish trial. Patients in the screening group were 50-74 years old and 45-74 years, respectively, in these trials.
According to the model used on the data from Denmark, the health sector saved €13.1 per screened individual and €6 410 per discovered CRC (both excluding the screening cost). The result based on the UK trial was €19.0 and €9 054. The savings were partly a result of less severely staged CRC requiring less treatment and the reduced probability of recurrence or advanced disease.

(Table 4)

3.10 Uncertainty

For most inputs, the model was insensitive to a 20-percent change. The total cost seemed most sensitive to changes in frequency of surgery and the use of bevacizumab in palliative treatment (see Appendices 1 and 2).
Generally, the cost results seemed to be sensitive to changes in treatment algorithms. This is especially important for evaluation studies with long-time horizons, such as for CRC screening and prevention. Due to a lack of data and continuous changes in the use of expensive chemotherapies, uncertainty in palliative chemotherapy seems to be an important area to address.

4 Discussion

The estimated lifetime, health care CRC cost for an average CRC patient was €41 550 and was highly dependent on the disease stage at diagnosis (€23 390 to €61 400). Compared with an empirical ("model-free") Norwegian study by Aas [2], our overall cost estimate was 39 percent higher, but only 1.3 percent higher after adjusting for differences in the included cost and time horizon (see more in [1]). The increase in costs according to the disease stage was in line with Ladabaum et al. [54] and Frazier et al. [34], while Brown et al. [55] found an increase in cost from stages I to III but a decrease from stage III to IV. However, comparing our CRC costs with non-Norwegian studies is difficult because of differences in unit costs and assumptions for the analyses [56]. Nevertheless, we compared our results with a recent Irish study by Tilton et al. that described the treatment regime and other important conditions that allowed for adjustment based on relevant differences [57]. When adjusting for the exchange rate, annual Irish inflation 2008-2011,

and important differences in unit prices and treatment regimens between the two studies, the cost difference between Tilton's and our model was -3.0, -1.3, 3.6 and -1.2 percent for stages I, II, III and IV, respectively, all within the estimated confidence intervals of the former study (see more in [1]).

The cost of cancer treatment estimated by the model generally seemed to be modest when comparing the cost of treating an average CRC patient with the number of years saved by the treatment, especially for a patient with tumor stage I, II or III. For these stages and given the WTP threshold mentioned, the willingness to invest for the years saved was 29.8 times the cost of a successful treatment in the first place (no recurrence)[2], and if we discounted future years, the result would be 20.5 times. For the average stage I-III patient (inclusive recurrences) the willingness to invest was 12.6 times the total average cost of CRC treatment per patient (7.7 times if discounting the years). For all these calculations we assumed that the CRC patients diagnosed to stage I, II or III at the age of 70 years, on average would have a life expectancy of 5 years without treatment. This implies that the general cost to society of taking resources from CRC cancer treatment and use them for other purposes could be very high (high alternative cost).

There is considerable uncertainty related to the assumption about the treatment regimens for palliative chemotherapy. The regimens are changing over time and differ between regions. Because of a lack of national data, we had to rely on a combination of published studies and expert opinions, the latter usually considered an uncertain source of data. On the other hand, experts may adjust for the expected change over time and between regions, so the in-data used could be brought closer to the present reality than estimates in the literature.

We found that a 5 percentage point *reduction in a 10-year recurrence rate* for stages I-III would reduce the CRC cost by €2 280 per patient and increase overall survival by 0.80 year per patient. Based on these findings and the declared, suggested threshold of €73 783 per QALY gained, the Norwegian health sector should be willing to invest €177 mill in total per year to achieve this reduction in recurrence rate (see Section 3.7). Our assumption about 5 percentage point reduction seems to be a moderate change. A study indicates that the recurrence rate for patients operated on by different surgeons can vary considerably [44]. Approximately 3 000 colorectal

[2] For the CRC patients treated at an age of 70 year and who got no recurrence (successful treatment), we assumed that they would still live 15.56 years on average. Further we only included treatment cost related to the primary examination, the primary treatment, and the follow up of this first treatment (see Table 1).

resections for malignancy are performed each year in Norway. Assuming that each colorectal surgeon should perform at least 15 resections each year to maintain competence, a maximum of 200 surgeons is needed in this field [58]. A comprehensive training program (initial colorectal surgery training and yearly follow-up training) using modern educational tools (such as simulators, operations on animals, etc.), accompanied with workshops and lectures by highly experienced and skilled colorectal surgeons, radiologists and pathologists, probably has the potential of improving the results far more than indicated above and to a cost far below €177 mill per year. Such training programs could therefore be highly cost-effective. Assuming that a comprehensive training program would cost €300 000 per surgeon and the effect would be a reduction in recurrence rate by 5 percent point, the investment would be paid back after only six operations for CRC. As a simplification, we used the value for a QALY as the value for a life-year saved. By adjusting for QALYs, we expect that the estimated gain from reduced recurrence would increase.

The estimates for 5 percentage point *reduction in a 10-year recurrence rate* could also be relevant for estimating possible gains from post-cancer prevention like lifestyle interventions (nutrition, physical activity, etc.). Some studies show significant effects of such interventions [59-67], but these effects are highly uncertain because of the scarcity of high-quality, randomized controlled trials [66 67]. When evaluating post-CRC cancer prevention, we also have to take into account possible changes in the quality of life, physical functioning, and the ability to tolerate treatment, as well as reduce fatigue, co-morbidity and non-CRC death [66 67].

In Section 3.8.1, we estimated the total willingness-to-invest per CRC case avoided by *primary prevention* to be €526 140. We assume that the chance for getting CRC could be reduced by 20 percent as a result of preventive interventions – like lifestyle interventions – and that the chance of getting CRC is 6 percent during a life. Given the suggested threshold, we would be willing to invest €6 080 per average person in Norway to avoid a CRC case (€506 380 * 6% * 20%). Additionally, lifestyle interventions have other positive effects, which also have to be accounted for in the estimates.

For the screening analysis, the estimates did not consider that some of the persons diagnosed with CRC in the screening group would have died of something else before their CRC had given symptoms to be diagnosed without screening. This implies overtreatment for the screening group, where some of the CRCs were unnecessarily discovered, which adds extra cost for the

screening group that was not included in our estimates. To include this in the analysis, we need data for the percent of the population with undiagnosed CRC who die of non-CRC causes.

For a discussion of the model's weakness and further development see Joranger et al [1].

5 Conclusions

Comparing the cost of treating an average patient with the number of years at stake, the health care cost of colorectal cancer seems generally to be modest. The lifetime health sector CRC cost is increasing, along with the stage of the disease and whether or not the patient experiences recurrence after an apparently curative resection.

Changes in the use of palliative chemotherapy will have a major impact on the average CRC cost. Reducing the recurrence rate by better surgery and achievable, preventive efforts like screening of asymptomatic persons, could have a considerable cost-effectiveness potential.

The different applications of the model illustrate how the model could be applied to evaluate a broad range of interventions (general model), making the model useful for health decision makers and health authorities.

Acknowledgments

We acknowledge the Department of Health, Nutrition and Management (HEL), The Faculty of Health Sciences, Oslo and Akershus University College of Applied Sciences for funding Paal Jorangers' doctorate.

References

1. Joranger P, Nesbakken A, Hoff G, Sorbye H, Oshaug A, Aas E. Modeling and validating the cost and clinical pathway of colorectal cancer. Medical Decision Making. 2014.
2. Jemal A, Center MM, DeSantis C, Ward EM. Global patterns of cancer incidence and mortality rates and trends. Cancer Epidemiology Biomarkers & Prevention. 2010 August 1, 2010;19(8):1893-907.
3. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JWW, Comber H, et al. Cancer incidence and mortality patterns in Europe: Estimates for 40 countries in 2012. European Journal of Cancer. 2013;49(6):1374-403.
4. Starling N, Tilden D, White J, Cunningham D. Cost-effectiveness analysis of cetuximab//irinotecan vs active//best supportive care for the treatment of metastatic colorectal cancer patients who have failed previous chemotherapy treatment. Br J Cancer. 0000;96(2):206-12.
5. Hisashige A, Yoshida S, Kodaira S. Cost-effectiveness of adjuvant chemotherapy with uracil-tegafur for curatively resected stage III rectal cancer. Br J Cancer. 2008;99(8):1232-8.
6. Hedden L, Kennecke H, Villa D, Johnston K, Speers C, Kovacic L, et al. Incremental cost-effectiveness of the pre- and post-bevacizumab eras of metastatic colorectal cancer therapy in British Columbia, Canada. European Journal of Cancer. 2012;48(13):1969-76.
7. Hassan C, Rex DK, Cooper GS, Zullo A, Launois R, Benamouzig R. Primary prevention of colorectal cancer with low-dose aspirin in combination with endoscopy: a cost-effectiveness analysis. Gut. 2012 August 1, 2012;61(8):1172-9.
8. Ercolani G, Cucchetti A, Cescon M, Peri E, Brandi G, Gaudio MD, et al. Effectiveness and cost-effectiveness of peri-operative versus post-operative chemotherapy for resectable colorectal liver metastases. European Journal of Cancer. 2011;47(15):2291-8.
9. Tappenden P, Jones R, Paisley S, Carroll C. The cost-effectiveness of bevacizumab in the first-line treatment of metastatic colorectal cancer in England and Wales. European Journal of Cancer. 2007;43(17):2487-94.
10. Lansdorp-Vogelaar I, van Ballegooijen M, Zauber AG, Habbema JDF, Kuipers EJ. Effect of rising chemotherapy costs on the cost savings of colorectal cancer screening. Journal of the National Cancer Institute. 2009 October 21, 2009;101(20):1412-22.
11. Frazier AL, Colditz GA, Fuchs CS, Kuntz KM. Cost-effectiveness of screening for colorectal cancer in the general population. JAMA: The Journal of the American Medical Association. 2000 October 18, 2000;284(15):1954-61.
12. Sharp L, Tilson L, Whyte S, O'Ceilleachair A, Walsh C, Usher C, et al. Cost-effectiveness of population-based screening for colorectal cancer: a comparison of guaiac-based faecal occult blood testing, faecal immunochemical testing and flexible sigmoidoscopy. Br J Cancer. 2012;106(5):805-16.
13. Aas E. Cost-effectiveness of screening for colorectal cancer with once-only flexible sigmoidoscopy and faecal occult blood test: Oslo University, Health Economics Research Programme2009.
14. Lee L, Li C, Landry T, Latimer E, Carli F, Fried GM, et al. A systematic review of economic evaluations of enhanced recovery pathways for colorectal surgery. Annals of Surgery. 2014;259(4):670-6.
15. Braga M, Vignali A, Zuliani W, Frasson M, Di Serio C, Di Carlo V. Laparoscopic versus open colorectal surgery: cost-benefit analysis in a single-center randomized trial. Ann Surg. 2005;242(6):890-5.

16. Murray A, Lourenco T, de Verteuil R, Hernandez R, Fraser C, McKinley A, et al. Clinical effectiveness and cost-effectiveness of laparoscopic surgery for colorectal cancer: systematic reviews and economic evaluation. Health Technol Assess. 2006;10(45):1-144.
17. Franks PJ, Bosanquet N, Thorpe H, Brown JM, Copeland J, Smith AMH, et al. Short-term costs of conventional vs laparoscopic assisted surgery in patients with colorectal cancer (MRC CLASICC trial). Br J Cancer. 2006;95(1):6-12.
18. Shabaruddin FH, Elliott RA, Valle JW, Newman WG, Payne K. Understanding chemotherapy treatment pathways of advanced colorectal cancer patients to inform an economic evaluation in the United Kingdom. Br J Cancer. 2010;103(3):315-23.
19. Helsedirektoratet. Economic evaluation of health intervention - a guide. Oslo: The Norwegian Directorate of Health; 2012.
20. Finansdepartementet. Guid for cost-benefit analysis. Oslo: The Treasury Department; 2005.
21. Sjo OH, Lunde OC, Nygaard K, Sandvik L, Nesbakken A. Tumour location is a prognostic factor for survival in colonic cancer patients. Colorectal Disease. 2008;10(1):33-40.
22. Nesbakken A, Nygaard K, Westerheim O, Mala T, Lunde OC. Local recurrence after mesorectal excision for rectal cancer. European Journal of Surgical Oncology. 2002;28(2):126-34.
23. Christensen K, Doblhammer G, Rau R, Vaupel JW. Ageing populations: the challenges ahead. The Lancet.374(9696):1196-208.
24. Sanoff HK, Carpenter WR, Stürmer T, Goldberg RM, Martin CF, Fine JP, et al. Effect of adjuvant chemotherapy on survival of patients with stage III colon cancer diagnosed after age 75 years. Journal of Clinical Oncology. 2012 July 20, 2012;30(21):2624-34.
25. van Gils CWM, Koopman M, Mol L, Redekop WK, Uyl-de Groot CA, Punt CJA. Adjuvant chemotherapy in stage III colon cancer: Guideline implementation, patterns of use and outcomes in daily practice in The Netherlands. Acta Oncologica. 2012;51(1):57-64.
26. Klint A, Engholm G, Storm H, Tryggvadóttir L, Gislum M, Hakulinen Tea. Trends in survival of patients diagnosed with cancer of the digestive organs in the Nordic countries 1964–2003 followed up to the end of 2006. Acta Oncologica. 2010;49(5):578-607.
27. Sorbye H, Pfeiffer P, Cavalli-Björkman N, Qvortrup C, Holsen MH, Wentzel-Larsen T, et al. Clinical trial enrollment, patient characteristics, and survival differences in prospectively registered metastatic colorectal cancer patients. Cancer. 2009;115(20):4679-87.
28. Seymour MT, Thompson LC, Wasan HS, Middleton G, Brewster AE, Shepherd SF, et al. Chemotherapy options in elderly and frail patients with metastatic colorectal cancer (MRC FOCUS2): an open-label, randomised factorial trial. The Lancet.377(9779):1749-59.
29. Archampong D, Borowski D, Wille-Jørgensen P, Iversen L. Workload and surgeon´s specialty for outcome after colorectal cancer surgery. The Cochrane Collaboration. 2012.
30. Rogers SOJ, Wolf RE, Zaslavsky AM, Wright WE, Ayanian JZ. Relation of surgeon and hospital volume to processes and outcomes of colorectal cancer surgery. Annals of Surgery. 2006;244(6):1003-11 10.97/01.sla.0000231759.10432.a7.
31. Lenzi J, Lombardi R, Gori D, Zanini N, Tedesco D, Masetti M, et al. Impact of procedure volumes and focused practice on short-term outcomes of elective and urgent colon cancer resection in Italy. PLOS ONE. 2013;8(5).
32. Scholefield JH, Moss SM, Mangham CM, Whynes DK, Hardcastle JD. Nottingham trial of faecal occult blood testing for colorectal cancer: a 20-year follow-up. Gut. 2012;61(7):1036-40.
33. RCPH. Screening for colorectal cancer in Vejle and Copenhagen county: Research Centre for Prevention and Health (RCPH); 2007.

34. Ladabaum U, Phillips KA. Colorectal cancer screening: Differential costs for younger versus older Americans. American journal of preventive medicine. 2006;30(5):378-84.
35. Brown ML, Riley GF, Potosky AL, Etzioni RD. Obtaining long-term disease specific costs of care: Application to medicare enrollees diagnosed with colorectal cancer. Medical Care. 1999;37(12):1249-59.
36. Yabroff KR, Borowski L, Lipscomb J. Economic studies in colorectal cancer: Challenges in measuring and comparing costs. JNCI Monographs. 2013 August 1, 2013;2013(46):62-78.
37. Tilson L, Sharp L, Usher C, Walsh C, S W, O'Ceilleachair A, et al. Cost of care for colorectal cancer in Ireland: a health care payer perspective. The European Journal of Health Economics. 2012;13(4):511-24.
38. Norderhaug I TH. Pasientvolum og kvalitet ved koloncancerkirurgi. Oslo: Nasjonalt kunnskapssenter for helsetjenesten2009.
39. Meyerhardt JA, Giovannucci EL, Holmes MD, Chan AT, Chan JA, Colditz GA, et al. Physical activity and survival after colorectal cancer diagnosis. Journal of Clinical Oncology. 2006;24(22):3527-34.
40. Meyerhardt JA, Heseltine D, Niedzwiecki D, Hollis D, Saltz LB, Mayer RJ, et al. Impact of physical activity on cancer recurrence and survival in patients with stage III colon cancer: findings from CALGB 89803. Journal of Clinical Oncology. 2006;24(22):3535-41.
41. Meyerhardt JA, Giovannucci EL, Ogino S, Kirkner GJ, Chan AT, Willett W, et al. Physical activity and male colorectal cancer survival. Arch Intern Med. 2009 December 14, 2009;169(22):2102-8.
42. Lynch BM, Cerin E, Owen N, Aitken JF. Associations of leisure-time physical activity with quality of life in a large, population-based sample of colorectal cancer survivors. Cancer Causes & Control. 2007;18(7):735-42.
43. Haydon AM, MacInnis RJ, English DR, Giles GG. Effect of physical activity and body size on survival after diagnosis with colorectal cancer. Gut. 2006;55(1):62-7.
44. Meyerhardt JA, Niedzwiecki D, Hollis D, Saltz LB, Hu FB, Mayer RJ, et al. Association of dietary patterns with cancer recurrence and survival in patients with stage III colon cancer. JAMA: The Journal of the American Medical Association. 2007 August 15, 2007;298(7):754-64.
45. Huxley RR, Ansary-Moghaddam A, Clifton P, Czernichow S, Parr CL, Woodward M. The impact of dietary and lifestyle risk factors on risk of colorectal cancer: A quantitative overview of the epidemiological evidence. International Journal of Cancer. 2009;125(1):171-80.
46. Rock CL, Doyle C, Demark-Wahnefried W, Meyerhardt J, Courneya KS, Schwartz AL, et al. Nutrition and physical activity guidelines for cancer survivors. CA: A Cancer Journal for Clinicians. 2012;62(4):242-74.
47. Ravasco P, Monteiro-Grillo I, Camilo M. Individualized nutrition intervention is of major benefit to colorectal cancer patients: long-term follow-up of a randomized controlled trial of nutritional therapy. The American Journal of Clinical Nutrition. 2012 December 1, 2012;96(6):1346-53.

Tables

Table 1. Percent in each stage, lifetime cost (euro) of CRC and survival time for a 70-year old CRC patient.

	All-stages	Stage I	Stage II	Stage III	Stage IV
Percent in each stage at diagnosis	100.0	17.8	36.3	25.7	20.2
Total lifetime cost	41 550	23 390	33 500	49 900	61 400
Different types of treatment					
Preop diagnostics and staging	2 050	1 900	2 110	2 350	1 690
Surgery - Major resection	17 910	16 640	17 500	20 180	16 890
Surgery – other	7 230	940	2 850	8 510	19 030
Adjuv./neoadj. Chemotherapy	1 340	24	530	4 100	450
Radiotherapy	1 620	690	1 580	2 850	950
Follow up, in total	1 810	690	2 730	2 530	200
Palliative chemotherapy	9 590	2 500	6 210	9 380	22 190
Different phases in the treatment					
Primary examination	1 650	1 700	1 640	1 650	1 630
Primary treatment	25 330	16 950	19 150	30 740	36 940
Follow up first treatment	1 690	640	2 590	2 320	180
Examination and treatment of recurrence (1st year with diagnosed recurrence)	3 170	1 540	3 780	5 590	440
Follow up after recurrence	120	58	140	210	17
Palliative chemotherapy	9 590	2 500	6 210	9 380	22 190
Cost scenarios					
Low (no complications and recurrence)		14 490	17 200	23 590	
High (Full treatment incl. recurr. and bevacizumab)	127 930	116 100	123 150	135 010	137 470
Survival (years)					
Years survived after diagnosis	9.3	14.0	11.5	9.0	1.5
Years survived after diagn., discounted (4%)	7.0	10.3	8.6	7.0	1.4
Life years lost	6.3	1.6	4.1	6.6	14.1
Life years lost, discounted (2%)	5.1	1.2	3.2	5.2	11.6
Life years lost, discounted (4%)	4.1	0.9	2.6	4.2	9.7

Table 2. Costs (€) of palliative chemotherapy for various full-treatment scenarios (receiving all treatments in the scenario), and the cost of the average patient starting palliative chemotherapy (assuming base-case conditional probabilities in Figure 2).

Treatment scenarios	Full treatment cost per patient	Mean cost per patient with palliative treatment			
		1st line	2nd line	3rd line	All lines
5-FU/FA, FLOX, EGFR + irinotecan	53 780	420	240	800	1 470
5-FU/FA, FLIRI, EGFR + irinotecan	53 970	1 050	630	2 010	3 680
5-FU/FA, EGFR + irinotecan	46 780	630	2 010	-	2 640
bevacizumab and FLIRI, FLOX, EGFR + irinotecan	85 210	9 880	1 070	3 640	14 590
bevacizumab and FLOX, FLIRI, EGFR + irinotecan	84 450	1 070	120	410	1 600
FLIRI, FLOX, EGFR + irinotecan	56 200	1 230	540	1 820	3 590
FLOX, FLIRI, EGFR + irinotecan	56 130	2 790	1 290	4 250	8 330
5-FU/FA, FLOX, EGFR + irinotecan	-	17 060	5 900	12 930	35 880

EGFR= Epidermal Growth Factor Inhibitor.

Table 3. Scenarios of palliative chemotherapy treatment show percent change and change in the cost of an average CRC patient, compared with base case.

Selected treatment scenario	Cost change (percent)	Cost change (€)
1. All patients getting palliative chemotherapy receive bevacizumab[2]	13.8	5 730
2. No patients receive bevacizumab	-5.4	-2 240
3. Those with FLIRI/FLOX in 1st line treatment in base case, receive instead bevacizumab and FLIRI/FLOX	8.1	3 360
4. Bevacizumab price from pharmacy reduced 50%	-2.3	-960
5. "All" patients not disease-free after treatment, receive palliative chemotherapy (see text)	9.4	3 910
6. All patients in scenario 5 above receive bevacizumab in 1st line treatment (see text)	28.8	11 970
7. Ten percent point moved from 5FU/FA-treatment (often old patients) to combination chemotherapy with bevacizumab	2.0	820
8. Ten percent point more get palliative chemotherapy	2.3	960

Table 4. Shows how the CRC patients are distributed in the screening and control group.

	Denmark		UK (Nottingham)	
	Screened	Control	Screened	Control
Stage I	0.370	0.148	0.506	0.151
Stage II	0.277	0.338	0.205	0.346
Stage III	0.272	0.300	0.241	0.285
Stage IV	0.081	0.214	0.048	0.218

Figures

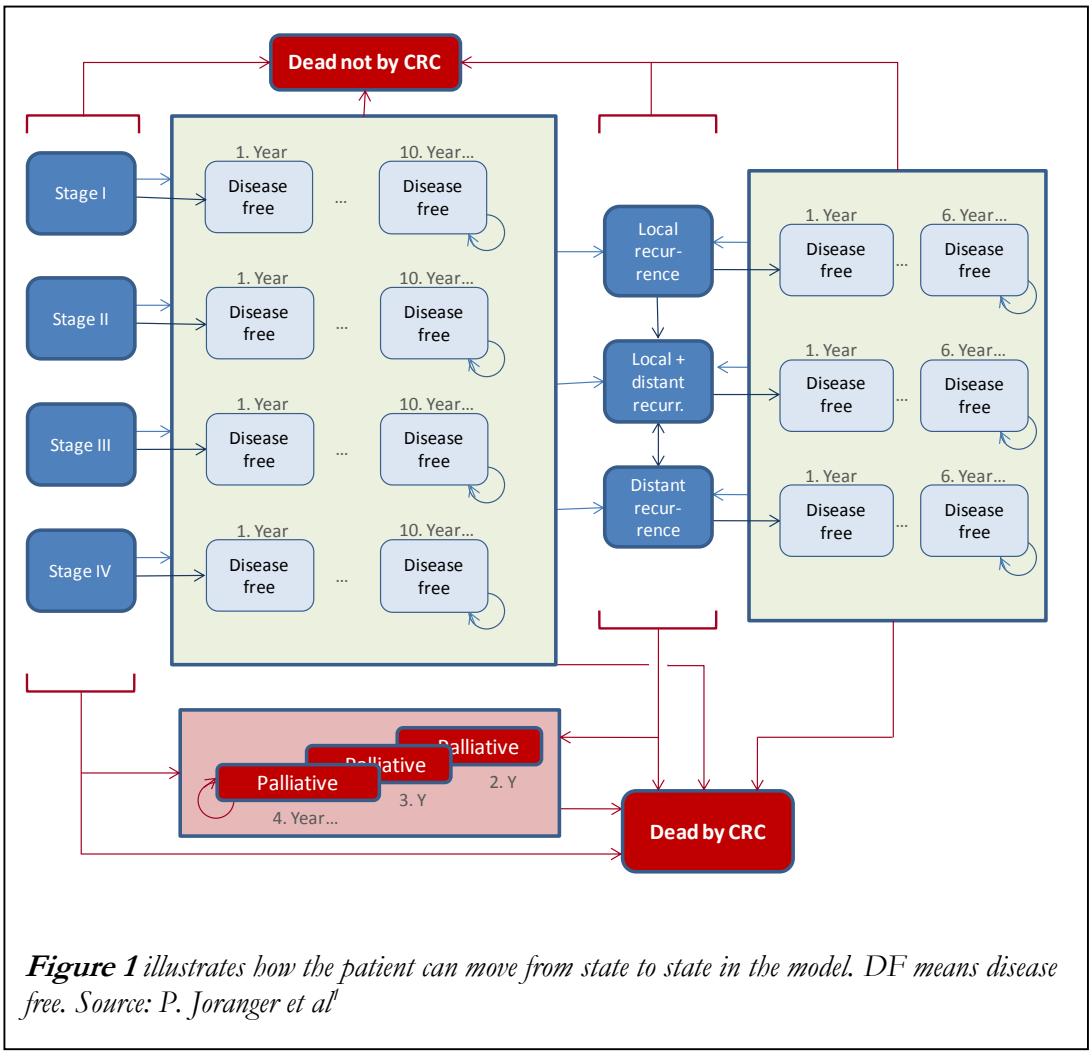

Figure 1 *illustrates how the patient can move from state to state in the model. DF means disease free. Source: P. Joranger et al*[1]

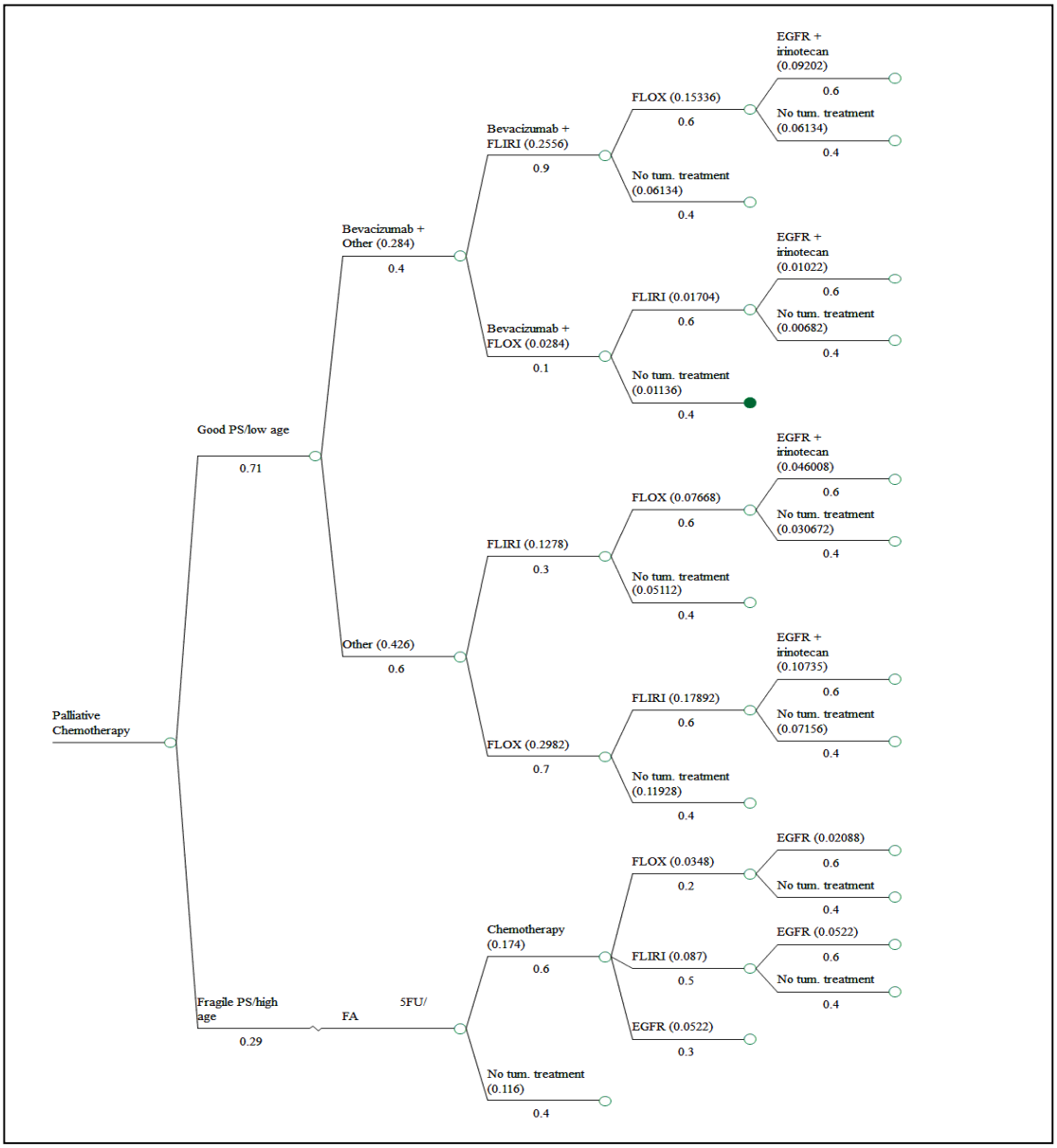

Figure 2. The decision-tree for palliative chemotherapy. Conditional probabilities without brackets.

Source: P. Joranger et al[1]

EGFR=Epidermal Growth Factor Inhibitor.

Appendix 1 (paper III):
One-way and multi-way sensitivity analyses

(App. 1: Published in a Web-only format)

We did a one-way sensitivity analysis where we increased the relevant parameter by 20 percent. The most important parameter was selected and shown as blue columns in Figure A1.1. These columns can both be analyzed as a result of price change or change in the use of resources. The green columns show selected changes of parameters normally decided by the government as partly empirically based, and the dark gray are different scenarios (see more in Tables 3 and A2.1 in appendix 2).

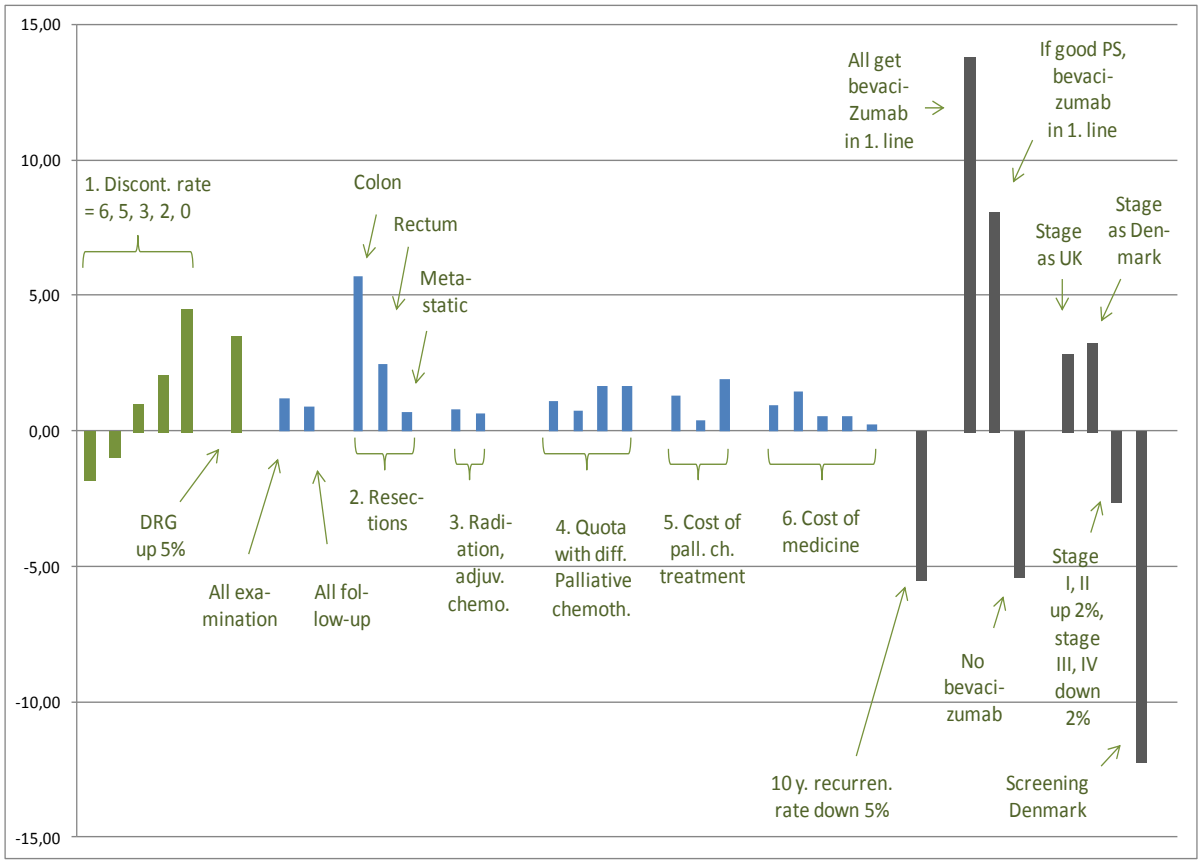

Figure A1.1 Percent change in total cost (all stages) when parameters are increased by 20 percent (the blue column) or changed as shown in the figure or in the text.

From the group to the left in Figure A1.1 ("Discount rate"), we see that the cost of an average CRC patient changes just about +/- 1 percent if the discount rate changes from 3, 4 or 5 percent,

which is normally the alternative value of the rate. The last green columns show that the cost changes by 3.5 percent if the value pr DRG increases by 5 percent.

In the blue column, the resection of the colon (5.7 percent) and rectum (2.5 percent) has the largest effect on the total cost (group 2). Our data are reliable for the probability for the different CRC patients having these resections, so the increase of 20 percent seems to be large compared to the real uncertainty for these parameters. The cost estimate used per resection is based on the DRG score system and is a common method in health evaluation today, but it is nevertheless criticized for having low reliability (Drummond et al . 2005, s. 59).
In group 3, we see that a 20 percent increase for all radiation (0.8 percent) or for all kinds of neoadjuvant or adjuvant chemotherapy treatments (0.7 percent) has less than a percent effect on the total CRC cost for all-stages.

For group 4, we analyze the effect of changing the probability of receiving a certain treatment by 20 percent and see that the results are affected by more than one percent for three of the elements. There is a lack of relevant statistics for this parameter, and we relied partly on expert opinion. Further, this parameter does change over time. Some possible effects of change are shown by the three dark gray columns. Palliative chemotherapy seems to be an important area for controlling uncertainty in the cost analysis, both because of the scarcity of data and the changing use of expensive drugs.

For "*6. Cost of medicine*" (price from pharmacy), we expect the parameter to be close to the prices the hospital paid for medicine in 2011. However, these prices often change over time and contribute with important uncertainty to the study of long-time horizons (study of screening). The first three dark-gray columns to the right show the effect on the cost when different transition probabilities are changed. The first column shows the change in the 10-year recurrence rate of 5 percent down, reducing the cost by 5.5 percent for stages I, II and III as a whole. This seems to be a test of both the uncertainty for the level of the parameters present value, and a relevant change for future years in the real value of recurrence.

Also, the stage distribution will influence the all-stage CRC cost. If we increase stages I and II by two percentage points and reduce stage III and IV by two percent, the cost will decrease by 2.6 percent. Further, if we change our distribution to that similar to the control group in the UK (Nottingham) study or the Danish study (ref), then the cost will increase by 2.8 and 3.2 percent,

respectively. This indicates that comparing all-stage CRC costs between populations can be disturbed by a different stage distribution. This can be important when some countries have screening programs and others do not. The last column shows the cost reduction (12.2 percent) if the stage distribution was changed to the screening group in the Danish study [53].

Generally, the cost results seemed to be sensitive to changes in treatment algorithms (e.g., palliative chemotherapy and screening). This is especially important for evaluation studies with long-time horizons, such as for CRC screening and prevention. Due to a lack of data and continuous changes in the use of expensive chemotherapies, uncertainty in palliative chemotherapy seems to be an important area to address.

Appendix 2 (paper III):
Change in the cost of palliative chemotherapy

(App. 2: Published in a Web-only format)

Table A2.1. The change in cost for an average CRC patient when increasing the input variable by 20 percent or 10 percentage point.

Treatment	20 percent increase		0.1 quota increase [1]
	Percent change	Cost change	Percent change
Change in the probability of receiving			
5FU/FA in 1st line (5FU/FA-scenario)	-0.48	-200	-0.83
Chemotherapy, 2nd line in "5FU/FA-scenario"	0.72	300	0.60
Bevacizumab, 1st line, assume in "no-5FU/FA-sc."	1.08	450	1.35
Chemotherapy, 2nd line, assume in "no-5FU/FA-sc."	1.66	690	1.38
Chemotherapy (EGFR+irinotecan), 3rd line	1.63	680	1.36
Change in cost for the treatment			
Nordic Flv	0.27	110	
Bevacizumab+FLIRI	1.29	540	
Bevacizumab+FLOX	0.14	60	
FLIRI 1st linje	0.16	70	
FLOX 1st linje	0.37	150	
EGFR (Cetuximab + irinotecan)	1.89	780	
Change in the cost of the medicine			
Bevacizumab	0.92	380	
FLIRI	0.51	210	
FLOX	0.54	220	
EGFR (Cetuximab + irinotecan)	1.47	610	
5FU/FA	0.24	100	

To show the importance of uncertainty in the input data, we estimated the effect of changes in both prices and probabilities (Table A2.1). Most sensitive to the 20 percent change in treatment cost were the EGFR (cetuximab) + irinotecan treatment with a 1.89 percent change (€780) and the "*bevacizumab + FLIRI*" treatment with a 1.29 percent change (€540).

When we only took into account a 20-percent increase in drug costs from the pharmacy, EGFR (cetuximab + irinotecan) had a 1.47 percent change (€610) and bevacizumab a 0.92 percent change (€380). The price of 5FU/FA was least sensitive (0.24 percent, €100) to a 20 percent change.

Paper IV

Adobor et al. Scoliosis (2014) 9:21
DOI 10.1186/s13013-014-0021-8

Scoliosis

RESEARCH Open Access

A health economic evaluation of screening and treatment in patients with adolescent idiopathic scoliosis

Raphael D Adobor[1*], Paal Joranger[2,3], Harald Steen[2,4], Ståle Navrud[3] and Jens Ivar Brox[5]

Abstract

Summary of background data: Adolescent idiopathic scoliosis can progress and affect the health related quality of life of the patients. Research shows that screening is effective in early detection, which allows for bracing and reduced surgical rates, and may save costs, but is still controversial from a health economic perspective.

Study design: Model based cost minimisation analysis using hospital's costs, administrative data, and market prices to estimate costs in screening, bracing and surgical treatment. Uncertainty was characterised by deterministic and probabilistic sensitivity analyses. Time horizon was 6 years from first screening at 11 years of age.

Objective: To compare estimated costs in screening and non-screening scenarios (reduced treatment rates of 90%, 80%, 70% of screening, and non-screening Norway 2012).

Methods: Data was based on screening and treatment costs in primary health care and in hospital care settings. Participants were 4000, 12-year old children screened in Norway, 115190 children screened in Hong Kong and 112 children treated for scoliosis in Norway in 2012. We assumed equivalent outcome of health related quality of life, and compared only relative costs in screening and non-screening settings. Incremental cost was defined as positive when a non-screening scenario was more expensive relative to screening.

Results: Screening per child was € 8.4 (95% CrI 6.6 to10.6), € 10350 (8690 to 12180) per patient braced, and € 45880 (39040 to 55400) per child operated. Incremental cost per child in non-screening scenario of 90% treatment rate was € 13.3 (1 to 27), increasing from € 1.3 (−8 to 11) to € 27.6 (14 to 44) as surgical rates relative to bracing increased from 40% to 80%. For the 80% treatment rate non-screening scenario, incremental cost was € 5.5 (−6 to 18) when screening all, and € 11.3 (2 to 22) when screening girls only. For the non-screening Norwegian scenario, incremental cost per child was € -0.1(−14 to 16). Bracing and surgery were the main cost drivers and contributed most to uncertainty.

Conclusions: With the assumptions applied in the present study, screening is cost saving when performed in girls only, and when it leads to reduced treatment rates. Cost of surgery was dominating in non-screening whilst cost of bracing was dominating in screening. The economic gain of screening increases when it leads to higher rates of bracing and reduced surgical rates.

Keywords: Cost minimisation analysis, Scoliosis screening, Scoliosis treatment, Health related quality of life

* Correspondence: radobor@ous-hf.no
[1]Section for Spine Surgery, Department of Orthopedic Surgery, Oslo University Hospital-Rikshospitalet, Pb 4950, Nydalen, N-0424 Oslo, Norway
Full list of author information is available at the end of the article

Introduction

Adolescent idiopathic scoliosis (AIS) is a complex three dimensional deformity of the spine, characterized by lateral curvature >10° and axial rotation, which affects 2-3% of otherwise healthy teenagers [1-3]. The deformity usually progresses with rapid growth of the spine and can affect health related quality of life of the patients [4]. Conventional treatment options are bracing and surgery [1-3]. Bracing is normally recommended for progressive curves of 20-40° in immature patients to prevent progression and reduce surgery, whilst surgery is considered for curves >45°-50° to stop progression and correct the deformity [1]. In patients with AIS, only a minority have progressive curves requiring treatment [5], and 90% of those treated are girls [6,7]. Treatment outcomes are usually measured by radiographic changes of the curves, but increasingly also by changes in health related quality of life. Early detection by screening allows for monitoring curve progression, and timely initiation of bracing. A recent randomised study found bracing to reduce curves which progress to the threshold of surgery [5].

Screening is controversial and practices vary worldwide [8-10]. Opponents cite mainly increased costs and lack of effectiveness of the programs. Some previous studies have supported whilst others have discouraged screening [11,12]. The United States Preventive Services Task Force neither supported nor opposed screening in 1993 [12,13], but recommended against routine screening in 2004 [14]. Discontinuation of screening programs has led to late detection and high rates of surgeries in various countries [15-17]. Currently, most international scoliosis and child health societies support and recommend screening [18]. The Scoliosis Research Society's International task force recently reported even before the BRAIST study [5] was published, that screening was effective in technical, clinical, program, and treatment efficacy, but could not make a statement on cost effectiveness due to lack of studies evaluating costs and health economic analyses [19].

Reviews and long-term studies suggest that health related quality of life of patients treated with brace or surgery are not different [1,2,6]. The aim of the present study was therefore to perform a cost minimization analysis (CMA) comparing only costs in screening and non-screening settings, while assuming equal long term health related quality of life of patients whose scoliosis are detected through screening or without.

Methods

We used a model approach to compare costs in screening with non-screening scenarios. The main mathematical equation on which the model was based is shown in Additional file 1. Input model parameters were collected from screening and hospital care. Screening in Norway was performed once in 12-year old children, and did not detect patients suitable for bracing [20]. We assumed similar epidemiology and natural history of AIS in Hong Kong and Norway, and used suitable data from a large population-based cohort longitudinal screening study by Lee et al. from Hong Kong in 2010 as model input for screening [21]. In this study, 115190 children were screened: 3158 received X-rays, 59 had out-patient visits for further assessment only, 264 were braced, 10 had surgery, and 29 had both brace and surgery (85% brace and 15% surgery). The percent treated in Hong Kong was thus 2.63 per 1000 children.

Screening is no longer performed in Norway. According to administrative data from the three scoliosis clinics in Norway, 122 adolescents were treated for scoliosis in 2012, of which 51(42%) were braced and 71(58%) had surgery, with about 10% of them having both brace and surgery. These 122 children, aged 11 to 17 is the number of patients out of the cohort of 63421 children who were the target group for scoliosis treatment in Norway for that year. Thus, the percent of children treated in Norway in 2012 was 1.92 per 1000 children.

Model input for the non-screening scenarios were based on Norwegian data when available. Otherwise, inputs were estimated from the Hong Kong data.

Study perspective in relation to costs

We used a health sector budget perspective focusing on the costs related to orthopaedic treatment in hospital care [22], and in addition, we included costs for the society due to transportation and parents' opportunity cost of time during treatment of their children.

Strategies being compared

Screening for scoliosis may lead to over-referrals to X-rays and outpatient evaluations, increased rates of bracing, but reduced surgical rates compared to settings when children are not screened [23,24]. In non-screening settings, many children are diagnosed late when they are matured, with curves not suitable for bracing [15-17,23]. We therefore assumed that reduced numbers of children are treated for scoliosis in non-screening settings and estimated reduced treatment rates of 90%, 80%, and 70%, respectively of those treated in screening by Lee et al. We compared costs in these reduced treatment rates to costs in the screening setting in Hong Kong. Treatment in this context includes the percentage of children who have X-rays for diagnosis, those treated with brace or surgery, and those who have further follow-ups. The estimated treatment rate of non-screening in Norway 2012 was 73% of that in Hong Kong. We also compared costs in non-screening scenario

in Norway 2012 with the costs in the screening setting in Hong Kong. Since AIS is more prevalent in girls, and 90% of those treated for AIS are girls [5,6], we performed separate analyses in girls.

In all non-screening scenarios, we simulated different distribution rates of brace and surgery based on the available non-screening data from Norway (58% surgery and 42% brace), since this is the only available data on the distribution of brace and surgery in a non-screening setting. We used data from Hong Kong to estimate the frequency of X-ray examination and referrals since non-screening Norwegian data was not available (see Additional file 1). Based on this study, we estimated that about 15% of children required referrals to X-ray and to specialist's examinations. In all non-screening scenarios, these rates were adjusted accordingly.

Incremental cost was defined as the cost of treatment in a non-screening scenario minus the cost of treatment and cost incurred in conducting the screening. A positive incremental cost therefore implies that screening is more cost saving compared to the non-screening scenario. How incremental cost changes by varying the ratio of bracing to surgery was estimated for all the non-screening scenarios. The probability of the incremental cost being positive was estimated for all cases.

Time horizon for cost estimations, discount rate

The time horizon for estimating costs was six years from the first screening at 11 years of age. We assumed two screenings per child, based on the recommendations of the Scoliosis Research Society [18] at the age of 11 and 13 years, and anticipated that 60% of the scoliosis cases were detected at the first screening and the rest at the second. We based our assumption on the knowledge of age and gender- specific prevalence of scoliosis, as well as the length of time between detection and treatment. Since screening tests are not fully accurate, it has also been suggested that scoliosis screening programs should be planned as a continuous process and not just a once and for all project as there is a possibility of missing out on some cases if screening is performed only once. For the non-screening scenarios we also assumed a dispersion of the expected cost (bracing and surgery) of 10%, 15%, 20%, 20%, 15% 10%, and 10% for each age group from 11 to 17 respectively. The literature is scarce with regards to the true dispersion of expected costs in scoliosis treatment, but shows a peak of treatment around 13–14 years of age. We therefore assumed 25% expected costs before, and 35% after the peak years [2,5,6,25]. When aggregating costs over time, we used an annual social discount rate of 4% (as recommended by the Norwegian Directorate of Health [26]) to calculate the present value of costs. The social discount rate is an interest rate used to bring future value into the present when considering the time value of money [22].

Estimating costs and resources

We used hospital's costs and administrative data, and market prices to estimate the cost of screening, bracing and surgery.

Screening

Screening was performed once in 4000 twelve year old children as part of a vaccine and physical examination program from autumn 2006 to spring 2007 [20]. Community nurses and physical therapists performed the screening. All activities directly involved in the screening and follow-up of patients were identified, measured, and costs estimated (Table 1).

Bracing and surgery

We estimated the costs of bracing and surgery based on data from hospital records. For bracing, we estimated the costs of the brace equipment, transportation, radiographic and clinical examinations during the period of brace wear, 3 days hospital hotel services for the child and one parent during brace fitting. Additionally, the costs of reimbursements for wear and tear of clothing and beddings from the National Insurance Scheme were included. For surgery, we estimated the costs of implants, salaries of the staff at the theatre, intensive care, intermediate postoperative care, regular ward costs, and costs of re-operations (Table 1).

Surgery was usually performed using either a hybrid construct with an average of 5 pedicle screws, 8 hooks, and 5 to 6 sublaminar wires or an all pedicle-screw construct using 15 to 17 pedicle screws. Two surgeons usually performed the surgery using an estimated average time of 180 minutes. One anesthesiologist, one anesthesiology nurse and two scrub nurses assisted them working on average for 300 minutes. After surgery, patients stayed in hospital for an average of 10 days. No braces were used postoperatively. During the first postoperative year, patients had two follow-up consultations. In addition, costs of radiological examinations, outpatient visits for follow-ups, transportation, and costs of complications and re-operations during the first year were measured.

With the public universal healthcare system in Norway, there are no hospital fees for parents when children are braced or surgically treated. Cost per hour for different health professionals was estimated by adding social costs of employment (pension, insurance, sick-leave, and training) and overhead to the salary (inclusive income tax). The salary and social costs for hospital staff were estimated using the mean salary at the Oslo University

Table 1 Resource unit used, cost (€) per unit, number of units and the uncertainty interval used for the cost estimation in the probabilistic sensitivity analysis (PSA)

No.	Variables	Unit cost (€)	Range (±%), cost	Units	Range (±%), units
	Screening				
1	Examiners (minutes)	47	20	9	20
2	Materials and supplies	0.03	20		
3	Scoliometer	1.4	20		
	For confirmation of scoliosis				
4	Transportation to X-ray exam	22	50		
5	Radiographs	63	30		
	For confirmed scoliosis >20°				
6	Transport to specialist evaluation	182	50		
7	Specialist evaluation	62	30		
8	Radiographs	128	30		
	Brace treatment				
9	Boston brace	3020	20	1.5	30
10	Reimbursement for wear and tear of clothes and linen/year	725	20	2	20
11	Hospital hotel, days (child and 1 parent)	212	30	3	30
12	Out-patient consultations	62	30	4	20
13	Physical therapy	55	30	1	20
14	Radiographs	128	30	4	20
15	Time used by one parent (days)	289	30	4	30
16	Transportation	137	50	4	50
	Surgery				
17	Implants/utilities (per operation)	9390	20		
18	Out patients consultations	62	30	4.5	10
19	Surgeons (hours)*	118	30	6	20
20	Anesthesiologists (hours)	118	30	5	20
21	Anesthesiologist nurse (hours)	71	30	5	20
22	Scrub nurses (hours)*	71	30	10	20
23	Intensive care (days)	4190	30	1	
24	Postoperative care unit (per day)	1872	30	2	25
25	Regular ward (days)	1541	30	8	25
26	Physical therapy	55	30	10	20
27	Radiology examination	160	30	6	20
28	Time used by one parent (days)	289	30	15	30
29	Taxi from home to school after treatment (days)	63	50	10	50
30	Transportation (days)	104	50	6	30
31	Transportation home after surgery	508	50		

*Two surgeons and two scrub nurses were involved in each surgery.
All items in each category of interventions were identified, measured, and costs estimated. Percentage of uncertainty was estimated for each item.
The percentages of the uncertainty of the PSA's are also given.

Hospital and the estimates of the overhead costs were based on data from the Norwegian Central Bureau of Statistics [27]. Salary and social costs of public health nurses were based on data from the Norwegian Nurses organization, and local community administrations.

Currency, price date and conversion

All prices and costs were converted from 2006 to 2012 NOK (Norwegian kroner) by using an inflation rate of 3.21% per year based on the yearly rate of change of one unit value within the Diagnosis-Related Group (DRG)

System in Norway. The exchange rate used was 8 NOK =1 € (Euro).

Statistical analysis

Values are given as numbers, percentages, means and mean differences. Results are presented with a 95% credibility interval (CrI), which show the 2.5th and 97.5th percentile of the outcome distribution. The uncertainty of input variables was assessed by one-way and multi-way sensitivity analyses. Parametric uncertainty was analyzed by probabilistic sensitivity analysis (PSA), where all uncertainties in the relevant parameters were accounted for simultaneously [22,28]. The PSA was used to analyse the distribution of incremental cost estimations in all scenarios (100000 interactions) and to estimate the CrI for total incremental costs, which forms the basis for the Tornado diagram in Figure 1. In the PSA, we used gamma distributions for estimation of unit costs, beta distributions for the number of hours used and their probabilities. Poisson distributions were used for the number of children treated.

The screening study was approved by the Regional Ethical Committee for Medical Research in Norway.

Results

Cost estimations

For all the relevant scenarios, the total estimated costs were € 8.4 (95% CrI 6.6 to 10.6) per child screened, € 10350 (8690 to 12180) per patient braced, and € 45880 (38040 to 55400) per surgery (re-operations included). The average time used to screen a child was 9 minutes (Table 1).

Incremental costs and outcomes

The incremental cost per child in a non-screening scenario of 90% treatment rate compared with screening was € 13.3 (1 to 27). The probability of the incremental cost being positive was 99%. In the 80% treatment rate non-screening scenario, incremental cost was € 5.5 (−6 to 18) with the probability of the incremental cost being positive was 82%. When comparing non-screening scenarios to screening for girls only: the incremental cost was € 11.3 (2 to 22) for the 80% treatment rate scenario and € 4.3 (−4 to 14) for the 70% treatment rate scenario. The probability of the incremental cost being positive was 99% and 82%, respectively. The incremental cost per child in the non-screening Norwegian scenario compared with screening was € 0.1 (−14 to 16), and the probability of the costs being positive was 50% (Table 2).

Comparing the undiscounted cost per child in the 80% treatment rate non-screening scenario, to screening, the cost of bracing per child of € 26.0 (21 to 33) was dominating in the screening scenario, whilst the cost of surgery per child of € 60.2 (48 to 75) was dominating in the non-screened scenario.

Incremental cost in the non-screening 90% treatment rate scenario varied from € -6.3 (−13 to 3) to € 27.6 (14-42) as the percentage of surgery increased from 30% to 80%. For the 80% treatment rate scenario with 30% surgery, and 70% bracing, incremental cost was € -11.0 (−19 to −3) favouring non-screening. With 80% surgery, and 20% bracing, incremental cost was € 18.2 (6 to 33) favouring screening (Table 3).

Characterizing uncertainty

The expected incremental cost estimates are shown in Figure 2. In the 90% treatment rate non-screening scenario, the probability of a positive incremental cost was close to 100%. Results comparing non-screening scenarios to screening in girls are shown in Figure 3. Uncertainty is also illustrated in the tornado diagram for the non-screening scenario of 80% treatment rate. The most important contributor to uncertainty was the percent braced, followed by the probability of being re-operated (Figure 1).

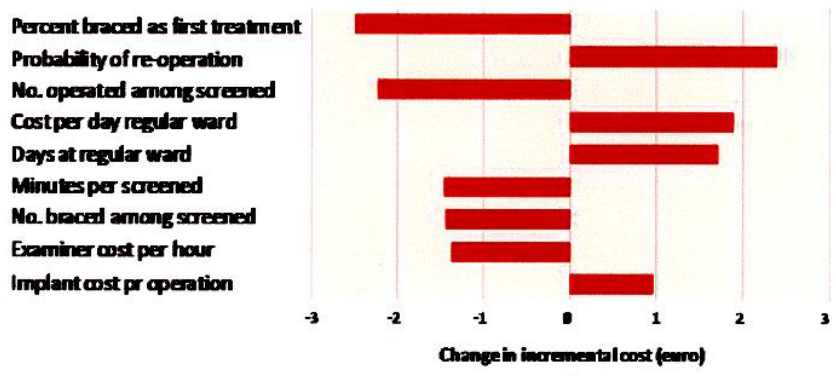

Figure 1 Tornado diagram (sensitivity analysis) for comparing the 80% treatment rate of Lee et al. non-screening scenario to screening.

Table 2 Cost (€) per alternative (screening boys and girls combined vs girls only) and incremental cost relative to screening in four non-screening scenarios with a 95% Credibility Interval (CrI)

	Screening boys and girls			Screening girls only		
	Cost per child	Incremental cost per child	Probability incremental cost >0	Cost per child	Incremental cost per child	Probability incremental cost >0
Screening	57.0 (49 to 66)	-		50.6 (44 to 58)	-	
Non-screening Norway	57.1 (44 to 73)	0.1 (−14 to 16)	50%	57.1 (44 to 73)	6.5 (6 to 21)	84%
Non-screening 90% treatment rate of Lee et al.	70.3 (59 to 84)	13.3 (1 to 27)	99%	70.3 (59 to 85)	18.4 (8 to 30)	>99%
Non-screening 80% treatment rate of Lee et al.	62.5 (52 to 75)	5.5 (−6 to 18)	82%	62.5 (52 to 75)	11.3 (2 to 22)	99%
Non-screening 70% treatment rate of Lee et al.	54.7 (46 to 66)	−2.3 (−13 to 9)	33%	54.7 (46 to 66)	4.3 (−4 to 14)	82%

The incremental cost was highest in the 90% treatment rate non-screening scenario with probability of being > 0 close to 100%. Incremental cost in non-screening Norway 2012 is close to the 70% treatment rate scenario. Incremental costs were higher in all non-screening scenarios when comparing screening of girls only than when comparing to screening of both boys and girls. The probabilities of incremental costs being >0 are also higher when comparing non-screening scenarios to screening of girls only than for both boys and girls combined.

Discussion

Scoliosis screening programs are considered to be beneficial from a clinical point of view [19], but are criticized for high costs due to high referral and treatment rates [8,11,13]. In the present study we used data from a large longitudinal screening study, and detailed costing of all activities in performing the analyses. Results suggest that screening is cost saving, unless both treatment rates and surgical rates are very low in comparative non-screening scenarios. In agreement with previously published studies reporting that discontinuation of screening has led to late detection and high rates of surgery [15-17], the model applied in the present study indicates that costs increase in non-screening scenarios with high rates of surgery and lower rates of bracing.

The effectiveness of a screening program thus depends on the costs involved and the number of cases detected early that result in bracing and less surgery compared to a non-screening setting. In a recent clinical trial, bracing reduced the number of children with curve progression to the threshold of surgery [5].

The results of the present study show that, screening has a large potential of cost saving if only girls are screened. Selective screening of girls is most cost saving because they constitute about 90% of those treated for scoliosis. In Table 2, we showed that there is a high probability of cost saving when only girls are screened compared to non-screening scenarios with treatment rates widely ranging from 70% to 100% of those of screening.

Table 3 shows that in the extreme non-screening scenario where treatment rates are approaching those of screening, screening both boys and girls was not cost saving. Likewise in the extreme non-screening scenario where treatment rates were very low approaching 60% of those treated in screening, non-screening becomes cost saving. However, these scenarios are the least likely to occur. In the non-screening scenarios where treatment levels are 90-100% of those in screening, patients are probably younger at detection, and likely to be recommended bracing according to guidelines and the results of the recent RCT study on bracing [5]. This implies that the ratio of bracing/surgery is likely to be >1 and bracing will be the dominating treatment option. On the contrary, when treatment levels in non-screening scenarios are in the 60% to 70% range of that of screening, patients are likely to be older and curves too large and not suitable for

Table 3 Incremental costs in non-screening scenarios compared with screening

		Ratios of brace/surgery in non-screening scenarios					
		20/80	30/70	40/60	50/50	60/40	70/30
Treatment rates in non-screening scenarios compared to screening	100%	**37.0** (22 to 55)	**29.7** (16 to 45)	**22.4** (10 to 36)	**15.1** (4 to 27)	**7.8** (−2 to 18)	**-0.5** (−8 to 9)
	90%	**27.6** (14 to 44)	**21.0** (9 to 35)	**14.5** (3 to 27)	**7.9** (−2 to 19)	**1.3** (−8 to 11)	**- 5.3** (−13 to 3)
	80%	**18.2** (6 to 33)	**12.4** (1 to 25)	**6.5** (−4 to 18)	**0.7** (−9 to 11)	**- 5.2** (−14 to 4)	**- 11.0** (−19 to −3)
	70%	**8.8** (−3 to 22)	**3.7** (−7 to 15)	**- 1.4** (−11 to 9)	**- 6.5** (−15 to 3)	**- 11.6** (−20 to −3)	**- 16.8** (−24 to −9)
	60%	**- 0.6** (−11 to 11)	**- 5.0** (−15 to 5)	**- 9.3** (−18 to 0)	**- 13.7** (−22 to −5)	**- 18.0** (−26 to −10)	**- 22.5** (−30 to −15)

Mean 95% CrI are given for non-screening scenarios with treatment rates from 60% to 100% combined with different ratios of bracing to surgery from 20/80 to 70/30.
Non-screening is more expensive with higher treatment rates and higher surgical rates compared with screening. Non-screening is less expensive with lower treatment rates and higher bracing rates compared to screening.

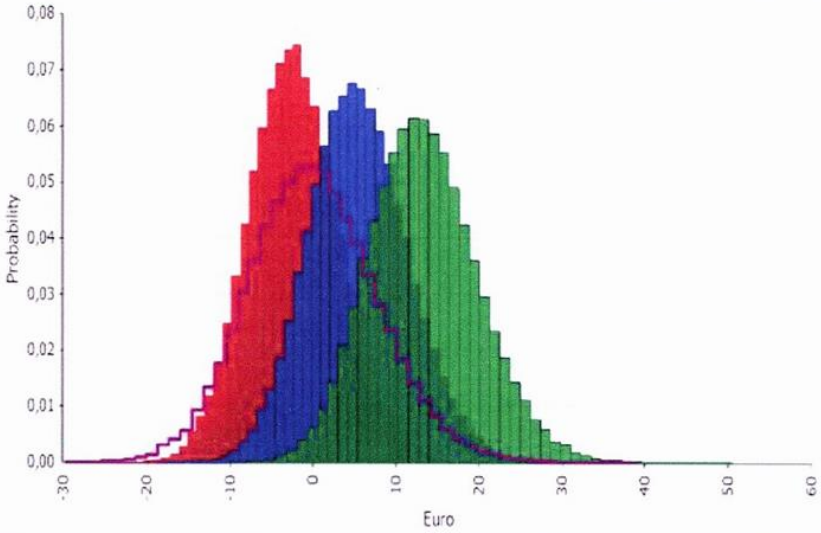

Figure 2 Incremental cost estimations in four non-screening scenarios compared to screening both boys and girls. Incremental costs increase from left to right looking at the top of the curves. Incremental cost was lowest in non- screening 70% treatment rate of Lee et al (red), followed by Norway (purple) 80% treatment rate of Lee et al (blue), and 90% treatment rate of Lee et al (green). Incremental costs were highest with higher treatment rate non-screening scenarios and lower in low treatment rate non- screening scenarios compared to screening of both boys and girls. The areas under the curves to the right of zero equals the probabilities of incremental costs being >0.

bracing [15], and surgery is most likely to be the dominating treatment option (i.e. ratio of brace to surgery likely to be <1).

In the Hong Kong study, about 15% of those detected by screening ended up having surgery compared to about 60% in non-screening Norway. Obviously, screening is not cost saving if the number treated in non-screening approximates that with screening and the surgical rate is 15%. However, this scenario is very unlikely to occur and was therefore not included in our analyses.

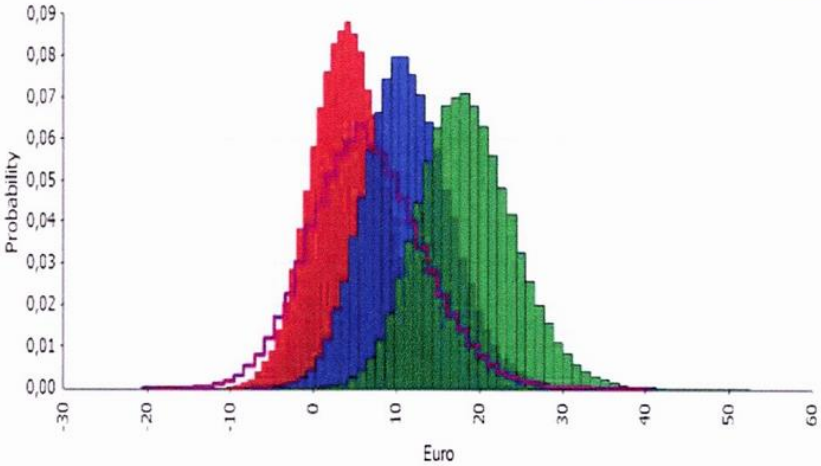

Figure 3 Incremental cost estimations in four non-screening scenarios compared to screening of girls only. Incremental costs increase from left to right looking at the top of the curves. Incremental cost was lowest (cost saving) in the 70% treatment rate of Lee et al (red), followed by non-screening Norway (purple), the 80% treatment rate of Lee et. al non screening scenario (blue), and the 90% treatment rate of Lee et al. non-screening scenario (green) compared to screening girls only. The areas under the curves to the right of zero equals the probabilities of incremental costs >0 which are considerably higher when comparing non-screening scenarios to screening of girls only than when comparing non-screening scenarios to screening of both boys and girls (Figure 2).

An interesting finding according to Table 3 is that screening both boys and girls tends to increase costs if the distribution of brace/surgery is 70/30 or 60/40 in a non-screening scenario. This scenario is also unlikely to occur. According to a previous Norwegian study non-screening scenarios of 30/70 or 40/60 are more likely to occur [15].

Our findings are in agreement with a review [29] on cost effectiveness of screening that found screening to be cost effective in one study [30], and recommended screening only for high-risk groups such as girls at twelve years of age in order to reduce over-referrals and over-treatment. However, the most recent review was not able to conclude whether screening was cost effective or not [31]. None of the studies cited in these reviews, however, applied recommended health economic evaluation principles [32].

Simulations in the present study suggest that the economic gain of screening increases when screening leads to higher rates of bracing and reduced rates of surgery. In a previous study, we reported higher rates of bracing and reduced surgical rates during a period of screening compared to a period without [15]. Similar findings have been reported from the Netherlands, Sweden and USA [23,24,33]. Bracing has been shown to reduce progression of curves to the threshold of surgery. In the recently published RCT study on bracing, the success rate was >70% and about 90% in those with high compliance [5]. Similar results were observed at long-term in a large Norwegian cohort study [6]. The current evidence of efficacy of bracing in the short term and good results at long-term indicates that patients with AIS should be detected early to allow for bracing. In addition, bracing avoids the complications of surgery, keeps the spine mobile, and might have positive long term effects. These benefits should be considered when interpreting the results of the present study. There has however been a lack of enthusiasm for bracing in the past amongst care providers. This is presumably due to the absence of high level of evidence of efficacy on bracing, and concerns of negative psychological impact on the patients. The results from the recent RCT study [5] on bracing do not however support this view.

With the assumptions made in the current study, screening of both boys and girls would neither have increased nor decreased costs compared to the treatment of AIS in Norway in 2012 where the estimated treatment rate was 73% compared to screening in Hong Kong, and 58% had surgery. However, selective screening of girls only would have been cost saving in Norway; as shown in Table 2 above.

Studies in the past have reported varying costs of scoliosis screening, and costs of bringing cases detected on screening to treatment, depending on how costs are measured [30,34-39]. The cost of screening in the current study is comparable to similar programs in Europe where total costs were included [34-36]. The estimated cost was based on two screenings per child, and community nurses performed the screening in conjunction with a vaccine program. Transportation costs and salaries of health professionals would have increased if screening had been performed in a different and isolated setting and not by community nurses. The estimated costs of bracing and surgery are comparable to those reported in the literature [40]. Many factors may influence the validity of our cost estimations. Treatment costs are likely to be underestimated in our study as bone grafts and intra-operative neuromonitoring were not used during surgeries, as compared with a study from the USA [40]. Our study perspective was limited to costs related only to expenses in an orthopedic department. We did not include costs related to primary health care, paramedics and alternative costs in relation to referred patients. In addition, we did not systematically register costs of patients' out- of-pocket expenses like transportation in relation to adjuvant treatment for scoliosis. Though physical therapy and counseling are not routinely offered to AIS patients in Norway, it is estimated that 1/3 of the patients use physiotherapy whilst under brace treatment or postoperatively [6,41].

Several input parameters contribute to uncertainties in our analysis. The cost of regular wards in surgical treatment was difficult to estimate accurately despite considerable effort. AIS patients undergoing surgical treatment require increased nursing resources compared to caring for ordinary pediatric patients at the orthopedic ward. The main analyses may also underestimate the cost of surgery.

The probabilities of positive incremental costs varied widely in the current study. There was however higher certainty in the incremental cost estimates when comparing non-screening scenarios to screening of girls only, as opposed to boys and girls combined. More research is warranted in order to reduce the uncertainties in future health economic evaluations of scoliosis treatment.

Limitations and strengths

Ideally, randomised studies or controlled prospective studies are needed to compare outcome in scoliosis treatment detected through screening or otherwise. However since the prevalence of scoliosis is low, it is difficult to include an adequate study sample even within a large country or internationally. Clinical trials including utility comparisons of bracing and surgery in both short and long terms are lacking. Utility scores may differ in shorter periods during treatment, for example by wearing a rigid brace, or postoperatively.

We assumed similar prevalence and natural history of AIS in Hong Kong and Norway in performing the analysis. Studies, however, show regional variations in the prevalence

of AIS, like higher prevalence in girls, but not boys in higher latitudes than in lower latitudes [42]. However, those differences could be linked to environmental factors such as the difference in the onset of menses in different geographic locations [43], and different cultures and not related to genetics. It is also likely that mechanisms of referral may be very different in the two settings, and in various countries, due to healthcare systems structures and barriers to access. The presentation of AIS has also been reported to be linked to socioeconomic status and race [44]. A recent study however found equal prevalence of AIS in 12- year old children in Malaysia and Norway [20,45].

The main strength of the present work is the application of current recommended standards for reporting health economic evaluations in conducting the study [32]. This gives more transparency and complete reporting of methods and findings which will facilitate interpretation and comparison of similar studies. We also used data from the largest reported longitudinal study of screening cohorts [21]. Analyses were performed to assess the uncertainties. The percentage detected for bracing, costs of surgery, and re-operations were the major contributors to uncertainty. More accurate estimates of these factors could improve the reliability and applicability of future analyses.

Generalisability

The model approach used in the current study could be employed worldwide with local cost estimate variations. Our results provide the missing economic evidence for health policy makers and healthcare providers to consider reintroduction of scoliosis screening.

In providing health services, policy makers are concerned about costs in view of limited healthcare resources, whereas patients and their families value the best treatment option available independent of costs. At present, there is a gap in the knowledge of the patient's preference in choosing treatment options. In a recently published trial, bracing was preferred to observation by patients and their families leading to the interruption of the trial and subsequently continued as a preference study [5].

Conclusions

Early detection through screening leading to bracing and fewer surgeries may save costs. Selective screening of high-risk groups like girls should probably be preferred. Screening is not likely to increase costs unless both treatment and surgical rates are very low in comparable settings where screening is not performed.

Consent

Written informed consent was obtained from all patients for the publication of this report and any accompanying images.

Additional file

Additional file 1: The mathematical model.

Competing interests

None of the authors have received benefits for personal or professional use from a commercial party related directly or indirectly to the subject of this manuscript e.g., royalties, stocks, stock options, decision making positions.

Authors' contributions

RDA, PJ and JIB designed the study. RDA, HS, and JIB were involved in the collection of the data for the manuscript. PJ and SN collected data and performed the health economic analysis. PJ built and ran the simulation model for the study. RDA, PJ, HS, SN and JIB took part in the analysis and the interpretation of results, drafting and critical review of the manuscript. All authors have given final approval to the version to be published.

Acknowledgements

The authors wish to acknowledge spine surgeons Roger Sørensen and Rolf Riise for their support during the research work and preparation of the manuscript. We also wish to acknowledge Health Region Southern Norway, the public community nurses, physical therapists, and school children who took part in the screening program. Authors also acknowledge Kristine Bratholm, Christina Idsøe, and Arild Holme at the economic department at the Oslo University Hospital, Rikshospitalet for their contribution towards the preparation of the manuscript. Authors acknowledge immensely the contribution of public health nurse Silje Rimeslåtten in preparing the community nurses for the screening program, and collection of data for the manuscript.

Author details

[1]Section for Spine Surgery, Department of Orthopedic Surgery, Oslo University Hospital-Rikshospitalet, Pb 4950, Nydalen, N-0424 Oslo, Norway. [2]Oslo and Akershus University College of Applied Sciences, P. O Box 4, St. Olavs plass, N-0130 Oslo, Norway. [3]School of Economics and Business, Norwegian University of Life Sciences, P. O. Box 5003, N-1432 Ås, Norway. [4]Biomechanic lab, Department of Orthopedic Surgery, Oslo University Hospital-Rikshospitalet, Pb 4950, Nydalen, N-0424 Oslo, Norway. [5]Department of Physical Medicine and Rehabilitation, Oslo University Hospital and University of Oslo, Pb 4950, Nydalen, N-0424 Oslo, Norway.

Received: 19 August 2014 Accepted: 18 November 2014
Published online: 06 December 2014

References

1. Asher MA, Burton DC: **Adolescent idiopathic scoliosis: natural history and long term treatment effects.** *Scoliosis* 2006, 1:2.
2. Weinstein SL, Dolan LA, Cheng JC, Danielsson A, Morcuende JA: **Adolescent idiopathic scoliosis.** *Lancet* 2008, 371:1527–1537.
3. Parent S, Newton PO, Wenger DR: **Adolescent idiopathic scoliosis: etiology, anatomy, natural history, and bracing.** *Instr Course Lect* 2005, 54:529–536.
4. Weinstein SL, Dolan LA, Spratt KF, Peterson KK, Spoonamore MJ, Ponseti IV: **Health and function of patients with untreated idiopathic scoliosis - A 50-year natural history study.** *Jama-J Am Med Assoc* 2003, 289:559–567.
5. Weinstein SL, Dolan LA, Wright JG, Dobbs MB: **Effects of bracing in adolescents with idiopathic scoliosis.** *N Engl J Med* 2013, 369:1512–1521.
6. Lange JE, Steen H, Brox JI: **Long-term results after Boston brace treatment in adolescent idiopathic scoliosis.** *Scoliosis* 2009, 4:17.
7. Karachalios T, Sofianos J, Roidis N, Sapkas G, Korres D, Nikolopoulos K: **Ten-year follow-up evaluation of a school screening program for scoliosis. Is the forward-bending test an accurate diagnostic criterion for the screening of scoliosis?** *Spine* 1999, 24:2318–2324.
8. Yawn B, Yawn RA: **Efficacy of school scoliosis screening.** *Orthopedics* 2001, 24:317.
9. Leaver JM, Alvik A, Warren MD: **Prescriptive screening for adolescent idiopathic scoliosis: a review of the evidence.** *Int J Epidemiol* 1982, 11:101–111.
10. Karachalios T, Roidis N, Papagelopoulos PJ, Karachalios GG: **The efficacy of school screening for scoliosis.** *Orthopedics* 2000, 23:386–391.

11. Burwell G: **The British decision and subsequent events. Spine (Phila Pa 1976)** 1988, **13:**1192–1194.
12. US Preventive Services Task Force: **Screening for adolescent Idiopathic scoliosis: Policy statement. JAMA** 1993, **269:**2664–2666.
13. US Preventive Services Task Force: **Screening for adolescent idiopathic scoliosis: Review article. JAMA** 1993, **269:**2667–2672.
14. **U.S. Preventive Services Task Force. Screening for Idiopathic Scoliosis in Adolescents: Recommendation Statement.** 2004. http://www.uspreventiveservicestaskforce.org/3rduspstf/scoliosis/scoliors.htm. 2004.
15. Adobor RD, Riise RB, Sorensen R, Kibsgard TJ, Steen H, Brox JI: **Scoliosis detection, patient characteristics, referral patterns and treatment in the absence of a screening program in Norway. Scoliosis** 2012, **7:**18.
16. Ali FM, Edgar M: **Detection of adolescent idiopathic scoliosis. Acta Orthop Belg** 2006, **72:**184–186.
17. Beausejour M, Roy-Beaudry M, Goulet L, Labelle H: **Patient characteristics at the initial visit to a scoliosis clinic: a cross-sectional study in a community without school screening. Spine (Phila Pa 1976)** 2007, **32:**1349–1354.
18. Richards BS, Vitale MG: **Screening for idiopathic scoliosis in adolescents. An information statement. J Bone Joint Surg** 2008, **90(1):**195–198.
19. Labelle H, Richards SB, De Kleuver M, Grivas TB, Luk KD, Wong HK, Thometz J, Beausejour M, Turgeon I, Fong DY: **Screening for adolescent idiopathic scoliosis: an information statement by the scoliosis research society international task force. Scoliosis** 2013, **8:**17.
20. Adobor RD, Rimeslatten S, Steen H, Brox JI: **School screening and point prevalence of adolescent idiopathic scoliosis in 4000 Norwegian children aged 12 years. Scoliosis** 2011, **6:**23.
21. Lee CF, Fong DY, Cheung KM, Cheng JC, Ng BK, Lam TP, Mak KH, Yip PS, Luk KD: **Costs of school scoliosis screening: a large, population-based study. Spine (Phila Pa 1976)** 2010, **35:**2266–2272.
22. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL: **Methods for Economic Evaluation of health Care Programmes.** 3rd edition. Oxford (UK): Oxford Medical Publications; 2005.
23. Bunge EM: **Screening for scoliosis: do we have indications for effectiveness? J Med Screen** 2006, **13:**29–33.
24. Montgomery F, Willner S: **Screening for idiopathic scoliosis. Comparison of 90 cases shows less surgery by early diagnosis. Acta Orthop Scand** 1993, **64:**456–458.
25. Lange JE, Steen H, Gunderson R, Brox JI: **Long-term results after Boston brace treatment in late-onset juvenile and adolescent idiopathic scoliosis. Scoliosis** 2011, **6:**18.
26. www.helsedirektoratet.no. Discount rate. 2014. Ref Type: Internet Communication.
27. **Statistical Yearbook of Norway, 1975–85 and 2003–11.** Statistics Norway; 2012. Electronic citation.
28. Briggs AH, Claxton KSM: **Decision Modelling for Health Economic Evaluation.** Oxford: Oxford University Press; 2006.
29. Sabirin J, Bakri R, Buang SN, Abdullah AT, Shapie A: **School scoliosis screening programme-a systematic review. Med J Malaysia** 2010, **65:**261–267.
30. Thilagaratnam S: **School-based screening for scoliosis: is it cost-effective? Singapore Med J** 2007, **48:**1012–1017.
31. Feldman DE, Beausejour M, Sosa JF, Goulet L, Parent S, Labelle H: **Cost effectiveness of school screening for scoliosis:A systemic review. Int J Child Adolesc Health** 2014, **7:**7–13.
32. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, Augustovski F, Briggs AH, Mauskopf J, Loder E: **Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. Cost Eff Resour Alloc** 2013, **11:**6.
33. Yawn BP, Yawn RA, Hodge D, Kurland M, Shaughnessy WJ, Ilstrup D, Jacobsen SJ: **A population-based study of school scoliosis screening. JAMA** 1999, **282:**1427–1432.
34. Grivas TB, Vasiliadis ES, Maziotou C, Savvidou OD: **The direct cost of "Thriasio" school screening program. Scoliosis** 2007, **2:**7.
35. Koukourakis I, Giaourakis G, Kouvidis G, Kivernitakis E, Blazos J, Koukourakis M: **Screening school children for scoliosis on the island of Crete. J Spinal Disord** 1997, **10:**527–531.
36. Ugras AA, Yilmaz M, Sungur I, Kaya I, Koyuncu Y, Cetinus ME: **Prevalence of scoliosis and cost-effectiveness of screening in schools in Turkey. J Back Musculoskelet Rehabil** 2010, **23:**45–48.
37. Yawn BP, Yawn RA: **The estimated cost of school scoliosis screening. Spine** 2000, **25:**2387–2391.
38. Morais T, Bernier M, Turcotte F: **Age- and sex-specific prevalence of scoliosis and the value of school screening programs. Am J Public Health** 1985, **75:**1377–1380.
39. Roubal PJ, Freeman DC, Placzek JD: **Costs and effectiveness of school screening. Physiotherapy** 1999, **85:**259–268.
40. Kamerlink JR, Quirno M, Auerbach JD, Milby AH, Windsor L, Dean L, Dryer JW, Errico TJ, Lonner BS: **Hospital cost analysis of adolescent idiopathic scoliosis correction surgery in 125 consecutive cases. J Bone Joint Surg Am** 2010, **92:**1097–1104.
41. Bjerkreim I, Steen H, Brox JI: **Idiopathic scoliosis treated with Cotrel-Dubousset instrumentation: evaluation 10 years after surgery. Spine** 2007, **32:**2103–2110.
42. Grivas TB, Vasiliadis E, Savvidou O, Mouzakis V, Koufopoulos G: **Geographic latitude and prevalence of adolescent idiopathic scoliosis. Stud Health Technol Inform** 2006, **123:**84–89.
43. Grivas TB, Vasiliadis E, Mouzakis V, Mihas C, Koufopoulos G: **Association between adolescent idiopathic scoliosis prevalence and age at menarche in different geographic latitudes. Scoliosis** 2006, **1:**9.
44. Zavatsky JM, Peters AJ, Nahvi FA, Bharucha NJ, Trobisch PD, Kean KE, **Richard S** Bucello Y, Valdevit A, Lonner BS: **Disease severity and treatment in adolescent idiopathic scoliosis: the impact of race and economic status. Spine J** 2013, **Oct 5** [Epub ahead of print]
45. Khindarli T: Prevalence of Scoliosis in Primary School Students in Marang District, Terengganu. http://www.researchgate.net. 2014. Ref Type: Electronic Citation.

Appendix A: The mathematical model

A1 Introduction

In this supplementary data, we show the core equation on which the simulation model was based. We begun by presenting the equations for estimating the cost of the different interventions: screening, diagnosis of scoliosis, confirming scoliosis > 20º, brace treatment and surgery. Then we estimated the fraction of children receiving the each category of interventions in the various scenarios. In the end we merged the estimated costs and the estimated fractions to estimate the cost pr child for each category of intervention and for the different scenarios.

The methodology used in the cost-minimizing analysis and discounting are presented in the main text of the manuscript and based on general literature on health economic evaluation like Drummond et al[1] or Hunink et al[2]. Methods for performing decision models probabilistic are based on Briggs et al.[3] The simulation model was built in Microsoft Excel. For the probabilistic sensitivity analysis we used the software @risk which is a part of the Decision Tools Suite software. The software @risk works is an extension to Excel.

A2 Estimation of cost of screening, brace treatment and surgery – all scenarios

Estimating the cost of the school screening per examination:

$C_s = (u_1 \bullet uc_1) + m + s$

u_1 = Number of minutes (units) used pr child pr examination (se row 1 in Table 1).
uc_1 = Cost pr minute (unit cost) used pr examination (se row 1 in Table 1).
m = Cost of materials and supplies per examination.
s = Cost of scoliometer pr examination.

Estimating the cost of diagnosing one child for scoliosis:

$C_{con} = t_{con} + rad_{con}$

t_{con} = Cost of transportation to/from X-ray exam (se row 4 in Table 1).
rad_{con} = Cost of radiographs (se row 5 in Table 1).

Estimating the cost per confirmation of scoliosis > 20º:

$$C_{con>20} = t_{con>20} + q_{con>20} + rad_{con>20}$$

t_{con} = Transport to/from specialist evaluation (se row 6 in Table 1).
$q_{con>20}$ = Specialist evaluation (se row 7 in Table 1).
$rad_{con>20}$ = Radiographs (se row 8 in Table 1).
Estimating the cost of brace per treatment:

$$C_b = \Sigma(u_j \bullet uc_j)$$

Where j = 9 to 16 in table 1. For example for j = 11 are $u_{11} \bullet uc_{11}$ equal to 3 hospital hotel days multiplied with €212 per day in hospital hotel, and likewise for the other cost components of brace treatment.

Estimating the cost per operation:

$$C_{su} = im + t + \Sigma(h_i \bullet hc_i) + \Sigma(u_j \bullet uc_j)$$

im = Utilities/implants cost per operation
t = Cost for transportation home after surgery.
h_i = Hour used of health personal in category i.
hc_i = Cost pr hour pr person of health personal in category i.
u_j = Number of units used of category j.
uc_j = Cost pr unit of category j.

Where i = 18 to 21 in Table 1, and j = 22 to 30 in Table 1.

For each child receiving an operation, 15% were assumed to be re-operated. So, per child operated the cost will be 100% + 15% of the costs estimated by the equation above.

A3 Estimating the fraction of children receiving each category of interventions

A3.1 The screened group

The fraction of the screened children receiving the different category of interventions is entirely based on the Hong Kong study:

$Fsc_j = TrHK_j / ChHK$

Fsc_j = The fraction of children in the screening group receiving intervention category j.
$TrHK_j$ = The number of children in the Hong Kong study receiving intervention category j.
ChHK = The number of children participating in the Hong Kong study.

Here, j = 31 to 34, were 31 means diagnosing scoliosis, 32 means confirming scoliosis > 20º, 33 means brace treatment and 34 means surgery.

A3.2 The non-screening group

A3.2.1 Non-screening scenario Norway

The fraction of children receiving surgery or brace treatment:

$FnscN_j = TrN_j/ChN$

$FnscN_j$ = The fraction of children receiving intervention category j.
TrN_j = The number of children 2012 in Norway receiving intervention category j.
Here, j = 33 and 34, were 33 means brace treatment and 34 means surgery. The number of surgery cases is the number of children receiving surgeon only and the number of children receiving surgeon after being braced.
ChN = The number of children in Norway in the age cohort of year 2012.

To estimate the number of children 2012 in Norway receiving surgery (TrN_{34}) we took the number of children receiving surgery as the first treatment option and added the 10% of the children receiving bracing as first treatment option because these children receive in addition surgery later on.
The fraction of the non-screened children confirmed for scoliosis or scoliosis > 20º:

$FnscN_j = (((TrHK_j / ChHK) \bullet ChN) \bullet Fr\text{-}conf) / ChN) = ((TrHK_j / ChHK) \bullet Fr\text{-}conf)$

Fr-conf = The fraction of the screened children confirmed for scoliosis or scoliosis > 20º, who also would be confirmed for scoliosis or scoliosis > 20º if the same group was not screened. Fr-conf was assumed to be 0.15. The treatment rate for the Norwegian scenario was 73%. To change this according to the scenarios with different treatment rate we adjusted the $FnscN_j$ to fit for the 80% scenario by multiplying $FnscN_j$ with 0.8, and for 70% scenario by multiplying with 0.7.

Here, j = 31 and 32, were 31 means confirmed for scoliosis and 32 means confirmed for scoliosis > 20º.

A3.2.2 Non-screening scenario 70%, 80% and 90%

We illustrate by using the 80% non-screening scenario. The same type of equations was used for the 70% and 90% scenarios.
The fraction of the children in a year cohort (or the chance pr child) receiving surgery or brace treatment for the 80% non-screening scenarios:

$Fnsc80_j = Tr80_j / ChN$

$Fnsc80_j$ = The fraction of children receiving category j treatment for the 80% non-screening scenario.
$Tr80_j$ = The number of children receiving category j of treatment in the 80% non-screening scenario.
Here, j = 33 and 34, were 33 means brace treatment and 34 means surgery.

In the 80% non-screening scenario, number receiving brace treatment and surgery, respectively:

$Tr80_{33} = ((TrN_{33}/(TrN_{33} + TrN_{34})) * TrNifHK) * 0,8$

$Tr80_{34} = ((TrN_{34}/(TrN_{33} + TrN_{34})) * TrNifHK) * 0,8$

$Tr80_{33}$ = The number receiving brace treatment for the 80% non-screening scenario.
$Tr80_{34}$ = The number receiving surgery for the 80% non-screening scenario. The number of surgery cases is the number of children receiving surgeon only and the number of children receiving surgeon after being braced.
TrNifHK = Total number treated with brace or surgery in Norway if the group was screened and treated as for the Hong Kong children. This parameter help us linking the fraction of treated when non-screened to the fraction of treated if screened – treatment among the non-screened is here 80% of the treatment among the screened.

$$TrNifHK = \Sigma ((TrHK_j / ChHK) * ChN)$$

Here, j = 33 and 34, were 33 means brace treatment and 34 means surgery.
Note that, when we use the notion "treatment rate" in the main text, we do not "double-count" the cases of surgery. Instead we refer to the rate of children treated by brace or sugary, where those who are both receiving braced and surgery are included among the braced.

A4 Estimating the cost pr child

A4.1 The screened group

Here we estimate the cost pr child in a cohort (defined as the selected one year cohort) for the different interventions.

$$CChS_{sc} = 1 * C_s$$

$$CChS_{con} = Fsc_{31} * C_{con}$$

$$CChS_{con>20} = Fsc_{32} * C_{con>20}$$

$$CChS_{b} = Fsc_{33} * C_{b}$$

$$CChS_{su} = Fsc_{34} * C_{su}$$

$CChS_{sc}$ = Cost of school screening pr child screened.

$CChS_{con}$ = Cost of confirming scoliosis pr child screened.

$CChS_{con>20}$ = Cost of confirming scoliosis > 20º pr child screened.

$CChS_{b}$ = Cost of bracing pr child screened.

$CChS_{su}$ = Cost of surgery pr child screened.

A4.2 The non-screening group

Here we use the 80% scenario as an example.

$CChN\text{-}S_{con} = FnscN_{31} * C_{con} * 0.8$

$CChN\text{-}S_{con>20} = FnscN_{32} * C_{con>20} * 0.8$

$CChN\text{-}S_{b} = Fnsc80_{33} * C_{b}$

$CChN\text{-}S_{su} = Fnsc80_{34} * C_{su}$

$CChN\text{-}S_{con}$ = Cost of confirming scoliosis pr child not screened.

$CChN\text{-}S_{con>20}$ = Cost of confirming scoliosis > 20º pr child not screened.

$CChN\text{-}S_{b}$ = Cost of bracing pr child not screened.

$CChN\text{-}S_{su}$ = Cost of surgery pr child screened.

These cost pr child pr intervention was dispersed over a 6 year period as described in the main text of the manuscript. The *incremental cost* was estimated by subtracting the total discounted cost pr non-screened child from the total discounted cost pr screened child.

Reference List

(1) Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for Economic Evaluation of health Care Programmes. 3 ed. Oxford Medical Publications; 2005.

(2) Hunink M GPSJWJPJEAea. Decision making in health and medicine - integrating evidence and values. Cambridge University Press:New York; 2001.

(3) Briggs AH, Claxton K SM. Decision modelling for health economic evaluation. Oxford: Oxford University Press; 2006.