Linköping University Medical Dissertations No. 1578

Health Economic Evaluations of Screening Programs

- Applications and Method Improvements

Mattias Aronsson

Department of Medical and Health Sciences
Linköping University, Sweden
Linköping 2017

© Mattias Aronsson, 2017

Cover design: Trevor Kirschenmann

Published articles have been reproduced with permission:

Paper I; permission granted by John Wiley & Sons, Inc on behalf of the British Journal of Surgery. Paper II; permission granted by Oxford University Press on behalf of Europace. Paper III; permission granted by Wolters Kluwer Health, Inc. on behalf of Neurology. Paper IV; permission granted by Oxford University Press on behalf of Europace.

Printed in Sweden by LiU-Tryck, Linköping, Sweden, 2017

ISBN 978-91-7685-491-4

ISSN 0345-0082

‘Medicine is a science of uncertainty and an art of probability.’

- Sir William Osler

CONTENTS

ABSTRACT	1
LIST OF PAPERS	3
ABBREVIATIONS	5
ACKNOWLEDGEMENTS	7
1. INTRODUCTION	9
1.1. Aim	11
1.2. Research questions	11
1.3. Outline of the thesis	11
2. BACKGROUND	12
2.1. Screening context	12
2.2. Current situation in Sweden	16
2.3. Case studies	18
2.4. Methodological context	22
3. METHODS AND DATA	30
3.1. Long-term cost-effectiveness of screening	30
3.2. Finding the optimal design of screening	42
3.3. Importance of changes in the screening setting	45
4. RESULTS	51
4.1. Long-term cost-effectiveness of screening	51
4.2. Finding the optimal design of screening	56
4.3. Importance of changes in the screening setting	61
5. DISCUSSION	63
5.1. Case discussion	63
5.2. Methodological discussion	66
6. CONCLUSIONS	73
7. REFERENCES	74

ABSTRACT

Screening to detect diseases early is attractive as it can improve the prognosis and decrease costs, but it is often a problematic concept and there are several pitfalls. Many healthy individuals have to be investigated to avoid a disease in a few, which results in a dilemma because to save a few, many are exposed to a procedure that could potentially harm them. Other examples of problems associated with screening are latent diseases and over-treating. The question of optimal design of a screening program is another source of uncertainty for decision-makers, as a screening program may potentially be implemented in very different ways. This highlights the need for structured analyses that weigh benefits against the harms and costs that occur as consequences of the screening.

The aim of this thesis is, therefore, to explore, develop and implement methods for health economic evaluations of screening programs. This is done to identify problems and suggest solutions to improve future evaluations and in extension policy making.

This aim was analysed using decision analytic cost-effectiveness analyses constructed as Markov models. These are well-suited for this task given the sequential management approach where all relevant data are unlikely to come from a single source of evidence. The input data were in this thesis obtained from the published literature and were complemented with data from Swedish registries and the included case studies. The case studies were two different types of screening programs; a program of screening for unknown atrial fibrillation and a program to detect colorectal cancer early. Further, the implementation of treatment with thrombectomy and novel oral anticoagulants were used to illustrate how factors outside the screening program itself have an impact on the evaluations.

As shown by the result of the performed analyses, the major contribution of this thesis was that it provided a simple and systematic approach for the economic evaluation of multiple screening designs to identify an optimal design.

In both the included case studies, the screening was considered cost-effective in detecting the disease; unknown atrial fibrillation and colorectal cancer, respectively. Further, the optimal way to implement these screening programs is dependent on the threshold value for cost-effectiveness in the health care sector and the characteristics of the investigated cohort. This is because it is possible to gain increasingly more health benefits by changing the design of the screening program, but that the change in design also results in higher marginal costs. Additionally, changes in the screening setting were shown to be important as they affect the cost-effectiveness of the screening. This implies that flexible modelling with continuously updated models are necessary for an optimal resource allocation.

LIST OF PAPERS

- I. Mattias Aronsson, Per Carlsson, Lars-Åke Levin, Jakob Hager, Rolf Hultcrantz. Cost-effectiveness of population screening for colorectal cancer in Sweden: a comparison of high sensitive faecal immunochemical test and colonoscopy. *Br J Surg*, 2017, 104(8),1078-1086.
- II. Mattias Aronsson, Emma Svennberg, Mårten Rosenqvist, Johan Engdahl, Faris Al-Khalili, Leif Friberg, Viveka Frykman-Kull, Lars-Åke Levin. Cost-effectiveness of mass screening for untreated atrial fibrillation using intermittent ECG recording. *Europace*, 2015, 17, 1023-1029.
- III. Mattias Aronsson, Josefine Persson, Christian Blomstrand, Per Wester, Lars-Åke Levin. Cost-effectiveness of endovascular thrombectomy in patients with acute ischemic stroke. *Neurology*, 2016, 86, 1053-1059.
- IV. Mattias Aronsson, Emma Svennberg, Mårten Rosenqvist, Johan Engdahl, Faris Al-Khalili, Leif Friberg, Viveka Frykman-Kull, Lars-Åke Levin. Designing an optimal screening program for unknown atrial fibrillation: a cost-effectiveness analysis. *Europace*, 2015, 17(7), 1023-9.
- V. Mattias Aronsson, Martin Henriksson, Lars Bernfort, Lars-Åke Levin. Economic evaluations of screening programs - systematically broadening the evaluation approach to optimal screening designs. *Submitted*.

ABBREVIATIONS

A	Action variable
$a_{x,y}$	Column y in row vector x
AF	Atrial fibrillation
CRC	Colorectal cancer
ECG	Electrocardiography
EVPI	Expected Value of Perfect Information
EQ-5D	EuroQoL-5 Dimension Questionnaire
FASS	Pharmaceutical Specialties in Sweden
FIT	Faecal immunochemical test
FN	False negative screening result
FP	False positive screening result
ICER	Incremental cost-effectiveness ratio
j	Years since the first screening occasion in a potentially optimal program
k	Threshold value for a QALY
MDP	Markov decision process
mRS	modified Rankin Scale
NICE	National Institute for Health and Clinical Excellence
NOAC	Novel oral anticoagulants
OAC	Oral anticoagulants
OC	Optical colonoscopy
POMDP	Partially observable Markov decision process
PSA	Probabilistic sensitivity analysis
QALY	Quality-adjusted life-year
ROCC	Receiver operating characteristics curve
RCT	Randomized Controlled Trial
SCREESCO	Screening study of Swedish Colons
STROKESTOP	Screening study of atrial fibrillation in Halland and Stockholm
TN	True negative screening result
TP	True positive screening result
t_y	Youngest age of screening in an optimal screening program
UK	The United Kingdom
WHO	World Health Organization
WTP	Willingness to pay
λ	Willingness to pay for a QALY

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to the following persons that have supported, inspired and encouraged me during my work with this thesis:

Lars-Åke Levin, my head-supervisor, for sharing your immense knowledge in the field of health economics. I am grateful for the guidance and that you always have been available. Also, thank you for all the fun we had while traveling to conferences and other meetings.

Lars Bernfort, my supervisor and good friend, for all the guidance and motivation you have provided. I would also like to express my gratitude to you for being the wise and down-to-earth person I needed. Your company is already deeply missed.

Lena Hector, my colleague, for all the kindness, patience and unfailing support.

Martin Henriksson, my colleague, for the constructive and excellent advice you have provided during the last years of the work with this thesis. Especially, thank you for the support in regard to the structure of this thesis.

Jenny Alwin, my boss, for being understanding and always providing great conditions for success. I would also like to express my deepest gratitude for the electronic meeting invites.

Mikeala Riddelberg, my roommate and PhD-colleague, for being available and providing good advices during the first confusing years of the work with this thesis.

Martina Lundqvist, my roommate and PhD-colleague, for being an excellent roommate, for all the kindness and for making the work with this thesis much more fun.

Therese Eriksson, my roommate and PhD-colleague, for being such an upbeat and warm person. Also, thank you for making the work with this thesis more fun.

Kasper Munk Johannessen, my colleague and PhD-colleague, for sharing your broad knowledge and for interesting discussions. Also, thank you for all things you helped me learn during my first year working in the pharmaceutical industry.

Almina Kalkan, Johanna Wiss and Jonathan Siverskog, my PhD-colleagues, for inspiration and making the work and breaks during the work with this thesis much more enjoyable.

All my colleagues at HSA, for the laughter and irrelevant discussions we had over the years. Also, thank you for providing excellent input on my research when I have asked for advice. It has been an amazing experience to work with such a great group of people.

Mårten Rosenqvist, Emma Svennberg and the others in the STROKESTOP-study group, for providing support, data and valuable knowledge during the work with the studies of atrial fibrillation.

Rolf Hultcratz, Jakob Hager and the others in the SCREESCO-study group, for providing me with support, data and immense knowledge of the field during the work with the studies of colorectal cancer.

Per Wester, Christian Blomstrand and Josefine Persson, research colleagues, for excellent teamwork and for sharing your knowledge about stroke.

Bristol-Myers Squibb, my current employer, for being understanding during the final year of the work with this thesis.

Ingela and Mikael Aronsson, for always believing in me and encouraging me to do my best. I am thankful for you always being there for me and showing so much interest in my work and life.

My Aronsson, for being an amazing sister and providing moral and emotional support.

Marcus Aronsson, my brother and best friend, for being very supportive during the work with this thesis. Especially, thank you for all help you have provided when I have had technical, mathematical and language questions.

1. INTRODUCTION

Screening to detect diseases early seems attractive as it improves the prognosis and could decrease the costs, but screening is a problematic concept and there are several pitfalls. Firstly, with the purpose of avoiding a disease or decreasing the mortality in a few, many individuals have to be investigated, where most are healthy. This results in a dilemma, because, to save a few, many are exposed to a procedure that could potentially harm them. Secondly, there is a problem with latent or pseudo disease, not everything that is detected will have caused symptoms or death to the individual. [1] The risk of over-treating has been a problem in the cases of screening for various types of cancer but was also the major reason for the negative recommendation from the National board of health and welfare on screening for unknown atrial fibrillation (AF), despite positive results in the clinical studies. [2-5] This highlights the need for structured analyses that weigh benefits against the harms and costs that occur as consequences of screening.

Economic evaluations of screening programs aim to assess the costs and consequences of broad management pathways encompassing the invitation and attendance to the screening test, interpretation of test results, decision to treat or put patients under surveillance, and implementation of chosen treatments. In order to provide good value for money, the cost of inviting an often large number of individuals to the screening tests must be reasonable in relation to the health benefits gained by treating a limited number of individuals. Decision-analytic modelling has been widely utilised to analyse the cost-effectiveness of screening programs. [6-8] It is well suited for this task given the sequential management approach where all relevant data are unlikely to come from a single source of evidence (e.g. a clinical trial) as parameters commonly involve screening specific information such as attendance rates, test characteristics, decision-alteration parameters, and the effect of implemented treatments themselves. Hence, the tools to evaluate a given screening program are available and used presently. Furthermore, a structured approach to handle uncertainty has been developed, in particular statistical uncertainty where probabilistic sensitivity analysis (PSA) can reflect how the precision with which the input parameters have been estimated impact the uncertainty in the cost-effectiveness results. [9] However, given the complex management pathways associated with the implementation of screening programs there is substantial non-statistical uncertainty associated with the implementation of screening programs. In particular, the question of optimal design of the screening program itself is a large source of uncertainty, as a screening program may potentially be implemented in very different ways. Design issues concern the timing of initiation of screening, whether screening should be repeated,

and if so, at what time intervals. [10, 11] This vastly increases the number of possible screening program designs that have to be studied well beyond the reach of clinical studies, where normally only one, or a few, particular designs are studied. Decision-analytic models of screening programs can provide a mean to address this thorny methodological, and for healthcare policy, important issue by assessing numerous screening designs in order to find the optimal design of a certain program. [12]

The widespread use of decision analytic models to synthesise evidence and evaluate the cost-effectiveness with respect to a few different designs of the screening programs means that good knowledge of the progression of the disease often is available. Still, only on rare occasions has this knowledge been used to systematically evaluate what an optimal screening program may look like. [12-14] Failure to systematically evaluate the optimal screening design may result in a scenario where the most effective programs are never studied clinically or from a health economic perspective, implying they will never be implemented, ultimately to the disadvantage of population health.

Another problem associated with evaluations of screening programs is the lack of a direct effect from the screening procedure itself. The effectiveness, and hence, the cost-effectiveness, of the screening, is strongly dependent on available treatments and how the underlying conditions develop and changes over time. Evaluation must consider the continuously changing world around screening programs. This requires flexible modeling of the disease and treatments but also continuously updated analyses.

1.1. Aim

The aim of this thesis is to explore, develop and implement methods for health economic evaluations of screening programs. This is to identify problems and suggest solutions to improve future evaluations and in extension policy making.

1.2. Research questions

The aim was investigated in a number of health economic evaluations of screening programs using the following research questions:

- How should long-term costs and health effects of different screening programs be estimated in health economic modelling analyses?
- What methods are suitable and need to be developed for systematically finding optimal designs of screening programs in terms of:
 - what screening technique should be used?
 - when the screening should be scheduled?
- Which impact have changes in the setting of the screening program on health economic evaluations:
 - when there are changes in the condition that we try to detect early or avoid?
 - when new treatments for the condition become available?

1.3. Outline of the thesis

The aim was investigated using these research questions throughout this thesis. Readers are guided through the thesis by first being provided a necessary background to screening, the used case studies and available methods in chapter 2. The methodology behind the performed analyses is explained together with input data in chapter 3 and the results are presented in chapter 4. In chapter 5, strengths and limitations of the analyses and how they answer the research questions are discussed. In the final part of the thesis, the conclusions of this work are stated.

2. BACKGROUND

This background chapter will provide health economists and, to the subject of this thesis related, professionals with a necessary basis to understand and interpret the other sections in this thesis. This will include the most important definitions, context, information and references for further reads. The chapter is divided into four parts; screening context, current situation in Sweden, case studies and methodological context.

2.1. Screening context

Screening

Screening is defined as a medical examination to detect a disease or condition of which the investigated individuals have not experienced obvious symptoms. The lack of symptoms is what distinguishes screening from diagnostics. The aim of screening is that the detection, or early detection, of a condition results in a better prognosis for the screened population. This improvement is most often based on a lower risk in the health states with a detected condition (i.e. screening for cardiovascular diseases) or less progressed diseases (i.e. cancer) due to treatment.

Timing of the screening

The timing of the screening is crucial to maximise the positive effect. Screening can, depending on the nature of the condition, be possible from the moment it is attainable to identify risk factors or other associated elements (e.g. genes) for a particular disease. Nevertheless, screening is often relevant to discuss from the time when it is possible to identify pathological changes in the investigated individual, it then continues to be possible until the time the patient visits the care for their condition, as examinations after this point are instead called diagnostics. The period where screening is possible is called sojourn time. Pinpointing the exact right time with the screening procedure is usually difficult, therefore the true time from screening to when the state would have been detected without screening is called the lead-time. Figure 1 describes the possibilities for timing of the screening.

Figure 1. Possible timing of screening and the different phases of when the disease in the affected individuals can be detected.

Optimising (maximising) this lead time is often crucial for the effectiveness of the screening. If the screening is performed too early, the subjects receive no benefit of the intervention as nothing is found. While, if the screening is performed too late, no benefit is achieved as the condition would have been detected anyway, or even worse, is untreatable.

Screening test

In addition to the timing of the test, the characteristics of the screening test itself are of vital importance. What represent a good screening test may vary widely, which is the reason why different types of tests are used in various programs. A common denominator of such tests is that they are not particularly invasive. This is to minimise the negative effects in the individuals who do not have the disease.

Additionally, of major importance in screening programs is the precision of the screening test. This is measured in terms of how good the test is to detect positive individuals but also how accurate it is in avoiding falsely diagnosing those that do not have the disease. This is referred to as sensitivity and specificity, respectively. Figure 2 shows how the sensitivity ($\frac{TP}{TP+FN}$) and specificity ($\frac{TN}{TN+FP}$) are defined. [15]

Figure 2. Sensitivity and specificity and how these relate to the result of a screening test. The green and red curves show the range of the results of the screening test in those with and without the disease. Cut-off value of the screening test is displayed with the dotted line, which is used to decide if an individual is categorized as positive or negative. TN: True negative, TP: True positive, FN: False negative, FP: False positive.

As shown in Figure 2, the sensitivity and specificity most often have a negative correlation, which is usually displayed using a receiver operating characteristics curve (ROCC). ROCCs shows the effect of the important choice between sensitivity and specificity. Figure 3 presents this correlation based on the hypothetical data presented in Figure 2.

Figure 3. Receiver operating characteristics based on the data in Figure 2. The curve describes the relation between sensitivity and specificity.

The shape of the ROCC presented in Figure 3 is common for a screening test. [16] A search for a higher frequency of positive cases (by choosing a lower cut-off value) often means that the number of false identification in-

creases. Large portions of all true positives can be detected without increasing the number of false cases much, but the number of false cases increases exponentially when approaching a perfect test in regard to finding all positives. This is a difficult consideration that must be made by decision makers and is a dilemma in many screening programs, because only those with true positive results could be positively affected by the screening, while those with true negative, false positive and false negative imply costs for the society and could have negative health effects.

Criteria for screening

In addition to the accuracy and timing of the test, the effect of the screening also depends on the characteristics of the treatments available and the sought-after disease. Important characteristics to consider have been discussed by the World Health Organization (WHO). In 1968, WHO stated ten criteria, called Wilson-Jungner criteria, that should be fulfilled for recommending implementation of screening. [17] These criteria have since received minor updates and clarifications by the organisation. [18] WHO's current criteria for screening are presented in Table 1.

Table 1. World Health Organization's criteria for screening.

Criteria	Description
1	The screening program should respond to a recognized need.
2	The objectives of screening should be defined at the outset.
3	There should be a defined target population.
4	There should be scientific evidence of screening programme effectiveness.
5	The programme should integrate education, testing, clinical services and programme management.
6	There should be quality assurance, with mechanisms to minimize potential risks of screening.
7	The programme should ensure informed choice, confidentiality and respect for autonomy.
8	The programme should promote equity and access to screening for the entire target population.
9	Programme evaluation should be planned from the outset.
10	The overall benefits of screening should outweigh the harm.

Types of screening

To fulfil these criteria, different types of screening programs need to be designed in different ways. These designs may vary widely between different diseases and countries. They are often divided into three main categories:

- *Population screening* (or mass/systematic screening) means that everyone in the population is invited to the screening, often at a certain age. This implies that a large proportion of those with the disease will be identified but will require large resources and a solid organisation of the health care.

- *Target screening*, which means inviting a portion of the population to screening. This often applies to a certain part of the population with a higher risk for a particular condition, or a higher risk of serious events given the condition. This normally results in a more effective screening but also means that serious considerations have to be made, ethical as well as medical, as appropriate individuals must be selected.
- *Opportunistic screening*, which imply screening in conjunction with a healthcare visit for any other reason than the disease sought after in the screening program. This means that opportunistic screening can be similar to targeted screening as only a certain population has the possibility to participate. An advantage of this type of screening is that when patients already are in contact with the care, the screening can be done at low cost and with high participation rates.

Summarising screening

Screening is performed to detect conditions in an early stage to improve prognosis. The nature of screening implies the relatively unique characteristics in relation to other medical interventions that only a few of the participants have positive effect of the procedure. The positive effect in those that benefit from the screening by having a better life than without screening must compensate cost and negative effects in the screened individuals:

- without the condition
- with a too late detected disease for improved prognosis
- with no effect despite an improved general prognosis
- with false positive diagnosis

2.2. Current situation in Sweden

Based on the WHO criteria, the Board of Health and Welfare have published Swedish criteria concerning when to recommend counties to use screening. The Board of Health and Welfare has 15 criteria and all current and future programs are required to go through these before they are implemented. [4] All criteria used in Sweden are listed in Table 2.

Table 2. The Board of Health and Welfares criteria from implementation of screening programs.

Crite- ria	Description
1	The condition should be an important health problem.
2	The natural course of the disease should be known.
3	The condition should have an asymptomatic phase that can be detected.
4	There should be an appropriate test method.
5	There should be measures to better effect in the early stages than at the clinical de- tection.
6	The screening program should reduce mortality, morbidity or disability associated with the condition.
7	The test method and further investigation should be accepted by the target popula- tion.
8	Measures permit must be understood and accepted by the target population.
9	The health benefits should exceed the negative effects of the screening program.
10	The screening program shall be acceptable from an ethical perspective.
11	Screening program's cost-effectiveness should be evaluated and deemed reasonable in relation to the need.
12	Information on participation in the screening program should have been valued.
13	Organisational aspects that are relevant to a national equivalent screening should have been clarified.
14	Screening program's resource requirements and feasibility must be assessed.
15	There should be a plan for evaluation of the screening program effects.

Based on these criteria, several screening programs are currently used in Sweden and more programs are analysed and discussed. There are mainly seven population screening programs that are recommended or dis-
cussed:

- *Breast cancer screening*, which has been applied in Sweden for 25 years and is a well-developed program that is used in the entire coun-
try. This screening is performed using mammography every 18 to 24
months in women between 40 and 74 years. The screening is consid-
ered to be clinical- and cost-effective. [19, 20]
- *Abdominal aortic aneurysm screening*, is recommended and shown
cost-effective in Sweden but often debated. [21] The current recom-
mendation is to screen men at age 65 years but not women. The
screening is performed as a one-time examination using ultra sound.
[4]
- *Cervical cancer screening* has been performed in Sweden since
around year 1960. Screening with Human Papillomavirus testing
and cytology is currently recommended for women aged 23-64 years.
[4]
- *Colorectal cancer screening* is recommended and has been running
in Stockholm County with stool sample test (gFOBT) used as the

screening test. A new large study of colorectal cancer (CRC) screening using faecal immunochemical test (FIT)-test and colonoscopy (OC) is ongoing in Sweden and presented in detail in section 2.3.

Screening for *cystic fibrosis*, *atrial fibrillation* and *prostate cancer* is also discussed. [4] Additionally, wild screening (i.e. patients visiting non-organised screening) is currently common for some diseases, such as prostate cancer.

2.3. Case studies

This section presents the most important information about the case studies used in the thesis. This implies two cases of population screening that were used to investigate the aims and research questions that are described in section 1.1-2. The two cases cover different types of screening programs; screening for (1) atrial fibrillation (AF) and (2) colorectal cancer (CRC), because various types of conditions mean different types of problems in an evaluation. Further, the implementations of treatment with thrombectomy and novel oral anticoagulants (NOAC) are used to illustrate how changes in the screening setting affect the evaluations.

Colorectal cancer screening

CRC is the third most commonly diagnosed cancer in the world and is associated with mortality and sufferings in the affected individuals and high costs for the society. [22] There are approximately 6000 new cases each year in Sweden and the disease specific mortality is over 40%. Its natural history and the polyp-cancer pathway have been well known for several decades. Cancer can be prevented to a large extent by removing the polyps. In addition, the early-stage detection of CRC often implies lower mortality rates. Hence, several guidelines, such as the US preventive service task force, [23, 24] the Gastroenterology multi-society task force, [25, 26] and the American Cancer Society, [27] have recommended CRC screening for asymptomatic average-risk individuals. [28, 29]

Several methods with different attributes, e.g. OC, faecal occult blood test (FOBT), FIT, and sigmoidoscopy, are already in use in screening programs. However, there is a widespread difference in the implementation status and strategies followed between countries. [30] All the screening options are considered to imply clinical benefits and have different impacts on the limited health care resources (e.g. supply of endoscopists). [31-34] Two of the most preferred methods for screening for CRC are OC and FIT. [30] OC is often seen as the 'golden standard' while FIT, followed by OC when FIT is positive, is less sensitive but has higher participation rates and limits the need for OC. Both methods have previously been shown to be clinical- and cost-effective compared to no screening. [8, 35-38] However,

the relative effectiveness of these screening methods is yet to be studied head-to-head for different programme designs.

The SCREESCO study

The Screening of Swedish Colons (SCREESCO)-study (NCT02078804) is a randomised controlled trial (RCT) on CRC screening, which is sponsored by and run in 18 of the 21 Swedish county councils with a total population of 7.5 million. The study is three-armed and has an inclusion time of three years. All participants will be 60 years old when they are included. Individuals 60 years of age will be randomised from the population register and invited to screening by mail. Twenty thousand individuals will be invited to primary OC and 60 000 individuals will be invited to high sensitive FIT and, if positive, to a subsequent follow-up OC. In total 120 000 randomised individuals will not be invited to screening but serve as controls and will be followed in the Swedish Cancer Register. Follow-up time was set to 15 years and the primary endpoints were disease specific mortality and CRC incidence. Secondary outcomes to be studied were quality assurance variables of OC, participants and non-participants experiences of the invitation and the screening procedure and health economics.

Screening for atrial fibrillation

The second case used in this thesis is screening for previously unknown AF. AF is the most common type of heart arrhythmia. The arrhythmia is present when the heart beats too fast, slow or in an irregular way. In Sweden, approximately 210 000 individuals, representing 3% of the adult (≥ 20 year) population, have a hospital diagnosis of AF. [39] The prevalence of diagnosed AF increases with age and has been reported to be 9 % among individuals of 75–79 years, 13.5% and 17.8% among the age groups of 80–84 years and 85 years and higher, respectively. [40] It has been predicted that the future prevalence of AF will increase further due to an ageing population. [41]

AF does not always affect quality of life but is associated with an increased risk of thromboembolic events, [42] especially ischemic stroke, which is associated with decreased quality of life, [43, 44] increased mortality, [45, 46] and high costs. [47] The risk of stroke can successfully be reduced by approximately 70 % in patients with AF using oral anticoagulants (OAC). [48] OAC treatment, according to the guidelines of European Society of Cardiology, is recommended in patients with AF and CHA₂DS₂-VASc score ≥ 2 (congestive heart failure, hypertension, age ≥ 75 years, diabetes mellitus, stroke, vascular disease, age 65–74 years, sex category; scores range from 0 to 9). [49, 50] Prior studies have shown that one third of individuals with AF might be asymptomatic, or have symptoms that are not well recognised as caused by AF. [51–53] As asymptomatic individuals are less likely to seek health care, they might not receive appropriate treatment; hence, they are at a higher risk of sustaining an ischemic stroke.

Importance of screening for identifying asymptomatic AF has already been realised by medical society and target screening for AF is recommended by international guidelines. [52, 54-56] Screening may help in decreasing the risk of stroke, providing positive effects such as reduced mortality and morbidity. However, screening of a large number of individuals also entails high costs. [57, 58] Whether the cost can be considered reasonable for the clinical gain is still a matter of debate.

The STROKESTOP study

The STROKESTOP study is an ongoing study to see if screening for untreated AF and initiation of OAC treatment can reduce the risk of ischemic stroke cost-efficiently over 5-years follow-up. Individuals born in 1936 /1937 and living in Stockholm County (n=23 888) or in the Halland region (n=4869) at the end of 2010 were identified by their personal identity numbers. Participants without a previous history of AF, who were in sinus rhythm on the first visit, were instructed in the use of a handheld electrocardiography (ECG) recorder for intermittent ECG recordings over two weeks. Participants placed their thumbs on the device for 30 seconds twice daily for 14 days, and in case of palpitations. The device has shown higher sensitivity for detection of AF than conventional 24 h Holter recordings. [59-61] A Zenicor model of the handheld-ECG is shown in Figure 4.

Figure 4. The handheld ECG-recorder used in the STROKESTOP-study (photo from www.zenicor.se).

AF was defined as at least one 30-second recording with irregular rhythm without p-waves, or a minimum of two similar episodes lasting 10–29 seconds during two weeks of intermittent recording. [62] Screening started in 2012 and concluded in the end of year 2013. In total, 13 892 inhabitants were invited to take part in screening and 54% participated (6887 participants). A previous diagnosis of AF was present in 636 patients (9.2%) and new previously unknown AF was detected in 210 patients (3.0%). In participants with newly detected AF 93% started OAC treatment.

Thrombectomy and acute treatment for stroke

Stroke is, as mentioned above, a leading cause of serious long-term disability worldwide, [63-65] resulting in decreased quality of life and survival, [45, 66] increased burden on informal caregivers, [67] and high costs to the society. [47] A primary predictor of the costs associated with strokes is the functional outcome after the stroke. [68] Hence, improved acute care is not only relevant from a medical perspective but also from an economic standpoint.

The only major treatment for an acute ischemic stroke that until 2015 was proven effective and recommended in guidelines was thrombolysis intravenous alteplase administered within 4.5 h of stroke onset. [69, 70] However, this treatment is often suboptimal in the case of an ischemic stroke with proximal occlusions of the major intracranial arteries. Therefore, intra-arterial treatment has been suggested for such patients, [71] which can be administered both chemically and with thrombectomy. The first endovascular randomised controlled trials (IMS- III, MR RESCUE, and SYNTHESIS expansion) without neuroimaging criteria of large-vessel occlusion and using less effective devices were not more effective than using intravenous tissue plasminogen activator. [72-74] However, with the use of effective stent retrievers, more rapid door-to-groin puncture, and neuroimaging criteria of proximal vessel occlusion, five randomised studies namely ESCAPE, EXTEND-IA, MR CLEAN, REVASCAT, and SWIFT PRIME showed clear benefits of adding endovascular treatment (primarily thrombectomy) to standard care. [71, 75-78] These studies presented short-term (3-months follow-up) outcome while the long-term effects of the treatment are yet to be studied. As preventing stroke and the damage it results in are the primary purposes of treatments (i.e. warfarin and NOAC) and screening programs (screening for AF, see section 2.3) the effect of such paradigm shifting treatment must be studied and considered when evaluating screening for AF.

Summary of case studies

In Table 3, the most important aspects of the cases presented in section 2.3 are summarised.

Table 3. Summary of case studies included in this thesis.

	Screening for CRC	Screening for unknown AF
Strategy under evaluation.	Screening with OC and FIT-test to find CRC or polyps in an early stage.	Screening with handheld-ECG to identify unknown AF.
Comparator	No screening and other screening techniques.	No screening and other screening options.
Primary patient population	60-year-old men and women.	75-year-old women and men.
Screening study name	SCREESCO	STROKESTOP
Status of intervention at the time for the evaluation	Screening is used in some counties in Sweden.	Population screening is not used in Sweden.
Recommendation at the time of the evaluation	Screening for CRC recommended in national and international guidelines.	Mass screening not recommended in international or national guidelines.
Method to evaluate the screening	Simulation model of Markov-type.	Simulation model of Markov-type.
Analysed research question	How should long-term costs and effects from screening be estimated? What screening technique should be used in the screening for CRC?	How should long-term costs and effects from screening be estimated? How is an optimal screening design identified? How is the evaluation of screening affected by changes in the screening setting?

2.4. Methodological context

This fourth section of the background chapter provides the context to the general methodology used in this thesis. This includes some cornerstones in health economic theory, methods and applications of these in the case of screening.

Health economic evaluations

A premise that exists in all economic problems is that resources are limited. This applies to all levels of the society and implies that we need to make choices about how to get most value for money. This is also the underlying assumption in health economics; all investments in the health care have an alternative use. The use of ineffective treatments ultimately means that resources must be taken from other parts of the society, for instance effective treatments. The net effect is a lower health in the society and lower social welfare. This is a rationale why the cost-effectiveness must be analysed; decision makers must have information about the costs and effects associated with treatments to make good and informed decisions. A tool to describe if a treatment is cost-effective and should be used is the cost-effectiveness plane (Figure 5).

Figure 5. A cost-effectiveness plane including the four possible areas in which interventions can be categorised.

A general and simple rule is that treatments that imply lower costs and better outcome than alternative treatments should always be implemented as they improve the welfare (Figure 5, D). Interventions that imply higher costs but worse outcome should never be considered as they lower the welfare (Figure 5, A). However, the relation between cost and health effects must be analysed further and judged in interventions where both factors increase or decrease (Figure 5, B and C).

A well-used tool to investigate this relation is the incremental cost-effectiveness ratio (ICER). The ICER is a measurement for this relation (the cost-effectiveness) and is defined as the differences between the cost of the investigated intervention and an alternative treatment divided by the differences in the outcomes between the same treatments:

$$ICER = \frac{Cost_{intervention} - Cost_{alternative\ T}}{Effect_{intervention} - Effect_{alternative\ T}}$$

Costs are measured in monetary terms while the effects can be expressed in different ways. To enable comparisons in the health care sector between different types of treatments, Quality-Adjusted Life-Years (QALYs) are often used. QALYs are calculated by multiplying the number of life-years in a health state with the health-related quality of life in that given state. A low cost per gained QALY of a new intervention represents good value for money and implies that it should be considered for implementation.

Methodological problems

However, performing health economic evaluations to analyse the cost-effectiveness of screening programs are associated with several analytical problems. Examples of such problems are:

- Long time-horizons in the studies and analyses are required to capture all relevant costs and effects.
- There are many possibilities in terms of how the screening could be conducted.
- Screening lacks a direct positive effect and is instead dependent on available treatments.

Some of the issues are common to health evaluations in general while other problems are strongly associated with evaluation of screening.

Long time-horizons

The long time from implementation of an intervention to the potential benefits is no specific problem for screening, but is prevalent there since it is often a long time between action and outcome in screening programs. This makes it almost always necessary to extrapolate the results of clinical studies as they rarely have a time horizon long enough to capture all relevant costs and effects. This also often implies that the calculation of costs and effects becomes very complex. A method to systematically structure this kind of calculation is decision analytic simulation models. The model type that is most commonly used in health economic research is Markov models. [79, 80] In health economic research, Markov Models means Markov processes and are characterised by typical Markov properties. This means a memoryless stochastic process that is based on chains of health states (called Markov chains).

Markov models can use continuous time but discrete time is most common in the health economic models, as this better suit the data that is typically available. In discrete time models, the Markov chains are repeated at fixed time intervals until the appropriate time horizon is reached. Individuals or other objects simulated in a Markov chain have at any point of time certain probabilities to move to other states in the chain. The next time period the individuals have new probabilities to move to the other states or remain in the current state (for example see Figure 6). An advantage of models in general is that probabilities for transitions, risk for events, utilities and cost can come from several different sources. Sometimes the model is based on a single clinical study (piggyback) while other times published data from several studies and registries are obtained from the published literature. At the start of the model, there is usually a decision problem, and

depending on the decisions taken, the simulated subjects will end up in different chains. Figure 6 shows a schematic figure of a typical three state Markov-model.

Figure 6. Example of structure of a basic three stage Markov model. Probabilities of transitions (p) to other health states, shown by arrows, are in this example dependent on the treatment given (T) and the time since the start of the model chain (t).

Long-term (often lifelong) costs and QALYs are then calculated for all decisions possible in the model. This is done by calculating costs and QALYs based on the distribution between health states and events in every time cycle in the model and then summing it up over the simulated individuals' lifetime in the model.

A related model, which is used primarily for sequential decisions, is Markov decision process (MDP). [81] These models carry the same properties as Markov models but include, unlike these, decision nodes in the Markov chain and not only in the start of the model chain. This implies that not only probabilities, but also decision actions, affect the simulated individuals in the model. Different decisions generate different reward functions, which is the basis for this type of models. These functions can then be solved mathematically so that we do not have to simulate all alternative decisions. Sometimes the current health state of the simulated individual is not known, in such cases we have to use partially observable MDPs (POMDP). [81]

An alternative to build models with lifelong effects is awaiting clinical studies with adequate follow-up. This often implies that we have to wait for 15-20 years, which means that major changes may occur in other areas, such as new screening methods, new treatments and changes in prevalence, incidence or severity of the requested state. These factors often make the results of such trials useless for the purpose of the evaluation and imply a long time lag until interventions become available for the public.

Many possibilities

Another difficulty for evaluation of screening is that the intervention is not clearly defined. In comparison with the evaluations of pharmaceuticals or medical technologies that only can be used in one or perhaps a few different ways, screening for a certain condition or disease can be designed in an almost infinite number of ways. Examples of such issues that need to be considered are:

- What screening test should be used?
- At what age should screening be initiated?
- How often should screening be repeated?

A design of a screening program must answer all these (and additional) questions, and a specific combination of answers to these is throughout this thesis called *screening program designs*. There are a large number of answers to the questions when discussing screening and there are inherent that decisions must be made when more than one option is possible. These decisions have to be made immediately and, in the case of screening, these are completely or partially mutually exclusive.

Simple general decision rules how this should be handled in such cases has, for example, been set up by Karlsson et al. and Johannesson et al. (Figure 7). [82, 83]

The first step is to construct a rank-list with all screening designs ordered according to the clinical effectiveness, i.e. based on what design that gains the most QALYs. This implied giving the design providing the fewest QALYs the lowest rank and the one providing most QALYs the highest rank. If a design provided less QALYs to a higher cost than any other design, the less effective intervention is dominated and thus should be excluded from the analysis and removed from the rank-list. This is represented by D in Figure 7.

The next step is to calculate the ICER for each successively more effective screening design compared to the previous design in the rank-list. In the cases where a higher ranked design provides a lower ICER, the less effective design should be ruled out by extended dominance. This is represented by B in Figure 7 as this design is dominated by a hypothetical combination of A and C (represented by E). [84]

Figure 7. Dominated and extended dominated programs in a cost-effectiveness plane.

The result of this is a rank-list with successively more effective design with increasingly higher ICERs.

Unfortunately, we cannot just implement the highest ranked of the remaining designs, as we must consider the alternative use of limited resources in the health care sector. As previously mentioned, the design not only competes with mutually exclusive screening designs but also against other types of consumption. Hence, we have to decide which intervention should be implemented based on available resources and, in extension, the health care decision makers' perception of the maximum willingness to pay (WTP) for gaining a QALY (λ). [83]

For an intervention to be considered cost-effective and be used, the ICER (cost of producing QALYs) for an intervention should be lower than what we are prepared to pay for gaining a QALY. What we are prepared to pay for gaining a QALY is called the threshold value. What this represents vary between different countries and healthcare systems. [85] In some systems, it represents what society considers a QALY to be worth (WTP_{QALY} , k , e.g. Sweden), while other countries aim to maximise the health budgets (λ , e.g. the United Kingdom (UK)). These values are the same in an optimal world, but this is rarely the case because of sluggish and inefficient transfers between different budgets in the society. In Figure 8, a treatment is considered cost-effective ($ICER < \lambda$) if it is located to right side of the orange line.

Figure 8. Threshold value for a QALY. The orange line and slope of this curve is often called the threshold value of a QALY. WTP: Willingness to pay.

No global threshold exists for what should be considered as cost-effective to pay for gaining a QALY. The National Institute for Health and Clinical Excellence (NICE) in UK uses a cost-effectiveness threshold ranging between £20 000 and £30 000 for a gained QALY. [86] The National Board of Health and Welfare, Sweden, states that a cost between €11 000 and €56 000 is a moderate cost per QALY while a cost over €56 000 should be considered a high cost for gaining a QALY. [87] The Council for Public Health and Health Care in the Netherlands recommend €80 000 per gained QALY as the threshold. [88]

These thresholds have been set relatively arbitrarily without systematic analysis. Attempts have been made to calculate the threshold for different countries, [89] but few have included a proper calculation as it is difficult methodologically and requires extensive work. In one of the most rigorous studies, Claxton et al. showed the ‘true’ threshold in the UK was significantly lower than those used as NICE guidelines. [89] In Sweden, Svensson et al. conducted a study based on subsidy decisions. [90] However, both the studies by Svensson et al. and Claxton et al. contain many assumptions, which limits the conclusions that can be drawn based on these studies.

Screening lacks a direct positive effect in itself

An additional problem is that screening is a special type of health intervention, as the intervention only seeks to identify a condition and not in itself improving the health of the subjects. This means that screening is highly dependent on the underlying disease and the treatments available.

The effectiveness of the available treatments has a major impact on the effectiveness of screening. Changes in the treatment that is used when the condition is identified also affect the efficiency of the screening. Unlike interventions such as pharmaceuticals that always have a “fixed” effect on the subject, no matter how far in the future we look, the effect of screening changes almost continuously since the screening depends on the current and future treatment regimes. Furthermore, this means that parameters such as adherence to treatments have a major impact on the efficiency and cost-effectiveness of screening.

The condition, which the screening aims to reduce or improve, can also change over time. An example of this is the case of screening for abdominal aorta aneurysm, where the prevalence of the condition has been reduced by 70% over the past 30 years. This has affected the usefulness of screening and, hence, the cost-effectiveness. [21] This can be caused by an addition of a new treatment to the treatment flora or changes in demographic factors (such as a reduction in smoking). [91] The usefulness of the screening will be affected if the severity of a screened for condition decreases. It could also be the other way around; if a treatment improves the ability to treat a condition if caught early, this may improve the efficiency of the screening.

3. METHODS AND DATA

This section presents how the methods described in Section 2.4 were used and what data sources that were included. As stated earlier, the general methodology used in this thesis were decision analytic cost-effectiveness analyses. These analyses were performed using simulation models constructed as Markov models (more details about this model type are presented in section 2.4). All models were developed in Microsoft Excel, deployed life-long time horizon and had a health-care payer perspective. The models consisted of a number of different health states which the simulated individuals could experience. Moreover, the models deployed fixed-time one-year cycles, and in every cycle, the simulated individual had a certain probability to move to another health state. The probabilities of staying or transitioning into other health states and risks of various events, were obtained from the published literature and complemented with data from the case studies described in section 2.3 and Swedish registries. We mainly used cost and population data from a Swedish setting in order to make the analyses as consistent as possible. Tunnel-states were used to model memory and incorporate time-dependency in our models. We used a discount rate of three percent for both effects and costs in all the base-cases. The time (2014-2016), when the analyses were conducted, was used for exchange rates and prices; therefore, certain results and costs will differ slightly in different parts of this thesis.

Two-way sensitivity analyses were performed to assess robustness of the models and the importance of uncertain parameters by varying the parameter values. The statistical uncertainty of the results was studied with PSA using Monte-Carlo simulations. The results of the simulation were analysed using cost-effectiveness planes and acceptability curves.

3.1. Long-term cost-effectiveness of screening

This first section of the Methods and Data chapter describes the specific methodology and data used for evaluating the cases of screening for CRC and AF.

Cost-effectiveness of screening for colorectal cancer

Analytic approach

The cost-effectiveness analysis of CRC screening was based on a decision-analytic Markov model. In the model, the remaining life of a hypothetical

population of Swedish 60-year-olds was simulated with an emphasis on CRC and its disease progression, polyp-pathway, detection and treatment. Figure 9 shows a schematic figure of the model structure, the most important health-states and other components.

Figure 9. Schematic figure of the CRC-model with its core health-states and possible annual transitions. Squares represent health states and arrows possible transitions with a certain probability to move from one health state to another. The left part of the figure, with dotted lines, represents the initial screening while the right part of the figure with solid lines represents health states and transitions after the screening. During this part transitions are repeated until all simulated individuals are dead. The lower part of the figure presents the different health states in the natural disease progression of known or unknown adenoma/CRC that is used in the model. CRC: Colorectal cancer.

Simulation overview

The cornerstone in the model is the simulation of the disease progression and the impact of screening. When the simulation starts at the age of 60 years, all simulated individuals are invited to screening, in which they may or may not participate. At that time, all individuals are either in the health state “Normal bowel” or in a health state with currently undiagnosed adenoma or CRC: “Low-risk adenoma”, “High-risk adenoma”, “Local CRC”, “Regional CRC” or “Distant CRC”. Per Swedish clinical guidelines, low-risk adenomas are defined as ≤ 10 mm, tubular growth, and low-grade of dysplasia. High-risk adenomas are defined as having one of the following char-

acteristics: >10 mm, tubovillous or villous growth, and high-grade of dysplasia. Health states representing CRC are defined according to American Joint Committee on Cancer, where Stage I-II are represented by “Local CRC”, Stage III by “Regional CRC” and Stage IV by “Distant CRC”. Over time, there is a certain annual risk for individuals with a normal bowel to develop low-risk adenomas. Low-risk adenomas can with a yearly probability develop into high-risk adenomas, which can develop into undiagnosed CRC (the progression is described in the lower part of Figure 10). These progression rates are not well studied in the scientific literature. Therefore, this analysis had to rely on indirect methods that used real-world prevalence/incidence data and mathematical methods for estimating the yearly progression rates. [92] Further, in the model, there is a certain yearly probability that an unknown condition, e.g. adenoma or CRC, can be diagnosed; this can be done either by screening (left part of Figure 10) or spontaneous detection (opportunistic screening or by symptomatic presentation).

If adenomas are detected, they are assumed to be removed with further follow-up, and when CRC is detected, the disease is treated according to treatment guidelines. Patients with CRC have a risk of dying from the disease and also from procedure-related events if treated or screened. The annual mortality rates for patients diagnosed with local, regional, and distant CRC for the first five years were obtained directly from the Swedish national CRC registries (data from 2009–2013). [93] CRC-related mortality was dependent on the age of the patient, the stage of cancer, and the time in the current cancer stage. From the 5th year and afterwards, we extrapolated the mortality rates with cancer based on the mortality in the previous years by using Weibull survival curves. The procedure-related mortality from OCs was obtained from a meta-analysis of complications observed from the examinations. [94] Irrespective of the bowel-status, all simulated individuals had a risk of dying from other causes (non-colorectal related). This standard mortality was dependent on age and gender and was retrieved from the national statistics of the general population in Sweden. [95]

Screening strategies

In the model, we simulated five different screening strategies. In all the below program strategies, screening is initiated at the age of 60 years in both men and women. Three strategies were adopted from the SCREESCO-study:

- I. No screening
- II. Screening twice with FIT
- III. Screening once with OC

As a reference, two additional screening strategies with repeated screening up to the age of 80 years were also simulated.

- IV. Screening biennially with FIT
- V. Screening every ten years with OC

This selection was based on what we identified to be the most commonly used programs globally for screening with OC and FIT. [30]

Screening parameters

Data regarding the participation rates in the first round of screening were projected based on early experience in the SCREESCO-study and are shown in Table 4. In the case of repeated screening, in the population of hypothetical individuals, we assumed that 31% could be categorised as ‘never attendees’ for FIT and 38% for OC. Further, 40% were categorised as ‘always attendees’ for FIT and 31% for OC. These numbers were calculated based on the early experience in the SCREESCO-study and a study of repeated screening in the UK. [96] This means that these two subpopulations will never, respectively always, attend the screening, irrespective of the number of times they were invited to the screening. However, the attendance rate at consecutive invitation to screening was assumed to be the same as at the first invitation to screening. [96, 97] The specificity and sensitivity of OC were retrieved from a meta-analysis presented by Telford et al. [98] and the corresponding parameters for FIT (cut-off value 10 μ g) were obtained from Wijkerslooth et al. [37]

Utility weights

Utility weights were applied to the simulated individuals in the model to create QALYs. These utility weights included values ranging from 1 (representing full health) to 0 (representing death). The age-dependent baseline utility weights, based on the general population of Sweden, were obtained from Burström et al. [99] The utility was adjusted for disease severity in individuals with cancer based on Ness et al. [100] In the base-case scenario,

screening procedures were not expected to affect the health-related quality of life. However, in the sensitivity analyses, published data about the willingness-to-pay to avoid OC were used in order to incorporate the potential discomfort of the procedure into the model. [101]

Unit costs and resource usage

Parameter data regarding resource usage, in the case of cancer, were obtained from the Swedish Colorectal Cancer registry and Tilson et al. [93, 102] The unit costs for pharmaceuticals in patients treated with chemotherapy were gathered from pharmaceutical specialties in Sweden (FASS). [103] Unit costs for visits and procedures were based on NordDRG. [104] Individuals attending OC were expected to miss one day of work. This implied a production loss of €221 (€53 000/ 240 workdays), as obtained from national statistics of all individuals aged 60–64 years in the general population of Sweden. [95] All unit costs were adjusted to the year 2016 and converted to Euro using the exchange rate on June 1, 2016 (€1 = 9.3 SEK). Details of the cost data used in the model are presented together with other input data in Table 4.

Table 4. Important parameter estimates used in the CRC- model.

Parameters	Mean value	Reference
General parameters		
Age (base-case)	60 years	Early experience in the SCREECO-study
Female gender (base-case)	50% *	Early experience in the SCREECO-study
Time-horizon	Life-long	
Discount rate	3%	[105]
Attendance rate OC	0.38	Early experience in the SCREECO-study
Attendance rate FIT	0.50	Early experience in the SCREECO-study
Attendance rate colonoscopy after FIT	0.81	[106]
Screening parameters		
False positive FIT	0.01 [†]	
Negative FIT	0.89 [†]	
True positive FIT	0.10 [†]	
Specificity FIT	0.93	[37]
Specificity OC	1.00	[107]
Sensitivity OC		
Low-risk adenoma	0.92	[98]
High-risk adenoma	0.97	[98]
Colorectal cancer	0.97	[98]
Sensitivity FIT		
Low-risk adenoma	0.06	[108]
High-risk adenoma	0.35	[37]
Colorectal cancer	0.88	[37]

Complications		
Perforations OC	0.09%	[94]
Bleedings OC	0.30%	[109]
Perforations polypectomy	0.09%	[94]
Bleedings polypectomy	0.87%	[110]
Spontaneous detection rate (yearly)		
Local	0.16	[111]
Regional	0.49	[111]
Distant	1.00 *	
Parameters	Mean value	
Prevalence at age of 60 years		
Normal bowel	76.12%	[112]
Low-risk adenoma	18.48%	[112]
High-risk adenoma	5.15%	[112]
Local cancer	0.18%	[112]
Regional cancer	0.05%	[112]
Distant cancer	0.02%	[112]
Transitions (yearly)		
Normal-to-low-risk adenoma (age 50–64)	0.007	[92]
Normal-to-low-risk adenoma (65–74)	0.008	[92]
Normal-to-low-risk adenoma (75+)	0.004	[92]
Low-risk-to-high-risk adenoma	0.02	[92]
High-risk adenoma-to-Local cancer	0.05	[92]
Local-to-regional cancer	0.20	[92]
Regional-to-distant cancer	0.65	[92]
New adenoma after polypectomy		
Removed high-risk adenoma ≤1year	0.25	[26, 113]
Removed high-risk adenoma >1year	0.04	[114]
Removed low-risk adenoma ≤1year	0.18	[26, 113]
Removed low-risk adenoma >1year	0.04	[114]
Costs		
Invitation to screening	25.00 *	
Material FIT-test *2	100.00	[115]
Lab FIT *2	100.00 *	
OC	6 065	[104]
Polypectomy	7 384	[104]
Quality of life weights		
Decrement local	0.08	[100]
Decrement regional	0.30	[100]
Decrement distant	0.55	[100]

*Assumption, OC: Colonoscopy, FIT: Faecal Immunochemical Test.

Sensitivity analyses

The probability that the treatment is cost-effective was calculated using a conservative threshold of €10 000 per gained QALY.

Cost-effectiveness of screening for atrial fibrillation

Analytic approach

The cost-effectiveness analysis of screening in 75/76-year old individuals for AF was based on a decision analytic Markov model. The developed model had an emphasis on AF, stroke, and OAC-treatment and how screening affects individuals for the rest of their lives. Using the simulation model, we analysed 1000 hypothetical individuals who matched the population of the STROKESTOP-study (according to sex, age, and CHA₂DS₂-VASC-score). The simulation of the natural disease progression and the effect of screening in men and women required data including prevalence, incidence, risk of events, morbidity, and mortality. Figure 10 depicts the core of the AF- model.

Figure 10. A basic description of the structure in the decision analytic Markov model for AF. The decision problem and screening procedure are described in Part 1 of the model, while Part 2 shows how the risk of thromboembolic events and bleedings depends on AF-status and CHA₂DS₂-VASC-score. All individuals may also suffer death from non-cardiac reasons. Part 2 was repeated every month for the rest of the life of the hypothetical individuals. AF: Atrial fibrillation.

Screening procedure

Handheld electrocardiography (ECG) was used as the screening technique to identify previously untreated AF. As in the STROKESTOP-study, [5, 116] the hypothetical individuals were asked to do a 30 second ECG recording twice daily for two weeks. We assumed that the attendance rate of 53% found in the STROKESTOP study was also relevant when the screening was repeated. However, we assumed that 25% of the population never attended

screening regardless of the number of screening occasions. This assumption is supported by the findings of screening studies in other areas. [96]

Prevalence of atrial fibrillation

In this analysis, we used the data from the STROKESTOP-study for determining the number of individuals with AF detected from the screening procedure. Previously undiagnosed AF was found in approximately 3% of the participants. The total prevalence including newly detected (3%) and previously known AF (9%) in the screened population was 12%. [5] The most important finding of the STROKESTOP-study was that previously undetected AF represented approximately 25% of the total prevalence of AF in 75/76-year-old individuals. In the analysis, this ratio was assumed to be applicable to all individuals aged between 55 and 85 years, as no other data were available.

The prevalence of diagnosed AF and the mean CHA₂DS₂-VASc in individuals diagnosed with AF was obtained from a Swedish registry study (Figure 11). [39, 117] As shown in Figure 11, the CHA₂DS₂-VASc score increases with age.

Figure 11. The prevalence of AF and the mean CHADS₂-VASc in men and women diagnosed with AF in Sweden. [39, 117]

The nature of asymptomatic AF is yet unknown. We, therefore, made assumptions regarding the proportion of the asymptomatic AF that would be detected by coincidence every year (for instance, from primary care or hospital visits). In the base-case scenario, we assumed that 5% of all unde-

tected AF would be detected every year without screening based on estimates from cardiologists. For the control group with no screening, the prevalence and incidence of diagnosed AF were modelled according to real-world data from a Swedish registry study (Figure 11). [39, 117]

Treatment of atrial fibrillation

For a screening program to be effective, there should be an accepted treatment for patients with recognised disease that affect mortality or morbidity. [17] In the case of identifying AF-patients, there is strong evidence supporting that OAC treatment reduces the risk of stroke. [48]

As per international guidelines, patients with a CHA₂DS₂-VASc score of ≥ 2 should be treated with OAC. All individuals with AF in the STROKESTOP-study had a CHA₂DS₂-VASc ≥ 2 and OAC should, therefore, be considered if AF was found. Considering the contraindications and concordance for OAC treatment, based on the STROKESTOP-study, 93% with detected AF and 1.5% with undetected AF were expected to be treated with OAC. The discontinuation rate was in the base case scenario modelled according to the ARISTOTLE-trial. [118] Additionally, based on the results of the same study, Apixaban (Eliquis, Pfizer) was expected to be used in 93% of the individuals if AF was detected, [5] and the characteristics of the individuals fulfilled the criteria for OAC treatment recommendation in international guidelines. [119] A discontinuation rate of 11% per year was applied to the model for the first 2.5 years, and after that, we used a yearly rate of 8%. [118, 120]

The novel OAC treatments (NOAC), including apixaban, rivaroxaban, and dabigatran, have been increasingly used in most industrialised countries. We expected this trend to continue; for example, Swedish authorities (The National Board of Health and Welfare) recommend that NOAC should be considered in all new treatment initiations. [121] In the base-case scenario, we assumed that apixaban (Eliquis, Pfizer) was used as the OAC treatment, because the available data from the ARISTOTLE-trial fitted our model better than data of dabigatran and rivaroxaban. [118, 122, 123] This assumption was not considered critical, because the cost of the three NOACs is similar, [103] and no direct or indirect comparisons, to the best of our knowledge, have found any significant difference between the drugs in terms of effect on ischemic stroke. However, warfarin is currently the most commonly used OAC in Sweden; hence, we included the results for warfarin in the sensitivity analysis.

Risk of events

The primary objective of the screening program was to limit the number of ischemic strokes. AF, as an independent risk factor, causes a four- to five-fold increase in the risk of stroke. [119] Apart from AF, congestive heart failure, hypertension, age, diabetes, prior stroke/TIA, vascular disease, and

female gender are also expected to affect the risk of stroke. The widely-used CHA₂DS₂-VASc scoring system using those variables, was applied in the model to predict the risk of stroke. The CHA₂DS₂-VASc risk score system is presented in Table 5.

Table 5. Description of the CHA₂DS₂-VASc scoring system

	Condition	Points
C	Congestive heart failure (or Left ventricular systolic dysfunction)	1
H	Hypertension	1
A₂	Age ≥ 75 years	2
D	Diabetes Mellitus	1
S₂	Prior stroke, TIA or thromboembolism	2
V	Vascular disease	1
A	Age 65–74 years	1
Sc	Sex category (female sex)	1

In the base-case scenario, we assumed that asymptomatic, often paroxysmal episodes of AF implied the same risk of stroke as longer symptomatic episodes. [124, 125] This assumption was tested in the sensitivity analyses.

The risk of major bleeding and the risk reduction for stroke of apixaban compared to warfarin were obtained from the ARISTOTLE-trial for patients with AF treated with NOAC. [118, 126, 127] The risk reductions for stroke from warfarin compared with no treatment was obtained from a meta-analysis of Hart et al. [48] The risk of stroke and bleedings for individuals with undetected and untreated AF was obtained from a Swedish registry study. [117] Moreover, the non-stroke related mortality based on gender and age was obtained from statistics of the overall Swedish population. [95]

Resource usage and unit costs

Cost data were mainly obtained from the Southeast Healthcare region of Sweden. The cost for the screening procedure was obtained from a previously published cost-effectiveness study of screening using handheld ECG stroke survivors. [128] Resources used in the screening procedure included invitation to screening, staff costs, materials, equipment, and additional examinations due to difficulties in diagnosing AF.

The mean cost of a stroke was obtained from a calculation of stroke costs in Sweden. [129] The monthly drug costs were gathered from FASS. [103] Other unit costs were obtained from published literature, and Table 6 lists all of them. Three percent was used as the discount rate for both costs and effects in the base-case scenario. All unit costs were adjusted to the year 2016 and converted to Euro using the exchange rate on January 10, 2016 (€1 = 9 SEK).

Utility weights

The QALY-weights used in the model were attributed to the participants' age based on the utility in the overall population of Sweden (Collected with EuroQoL-5 Dimension Questionnaire (EQ-5D)). [130] In the study, health-related quality of life was collected with EQ-5D and was translated into utility weights, i.e. QALYs, using the British value-set published by Dolan. [131] Ischemic and hemorrhagic strokes were expected to decrease the quality of life for the individuals. [43] Table 6 presents the quality of life and utility decrements used in the model together with the other parameters.

Table 6. Important parameters in the AF screening model.

Parameter	Mean	Reference
Baseline characteristics		
Age	75	[5]
Female gender	55.9%	[5]
CHA ₂ DS ₂ -VASc score	3.45	[5]
Prevalence unknown AF	3.0%	[5]
- Persistent AF	0.5%	[5]
- Paroxysmal AF	2.5%	[5]
ECG recordings per subject	26	[5]
Probabilities		
Mortality stroke (CHADS ₂ = 2)	0.269	[45]
Yearly spontaneous detection of asymptomatic AF (base case)	5%	Assumption
AF detected after stroke	88.2%	[52]
Intracranial bleedings	0.6	[117]
Major bleedings warfarin	5.2	[117]
Major bleedings no OAC	2.3	[117]
Yearly stroke risk in AF		
With apixaban		
CHA ₂ DS ₂ -VASc = 2	0.008	[48, 117, 118]
CHA ₂ DS ₂ -VASc = 3	0.012	[48, 117, 118]
CHA ₂ DS ₂ -VASc = 4	0.018	[48, 117, 118]
CHA ₂ DS ₂ -VASc = 5	0.027	[48, 117, 118]
CHA ₂ DS ₂ -VASc = 6	0.037	[48, 117, 118]
With no OAC		
CHA ₂ DS ₂ -VASc = 2	0.024	[117]
CHA ₂ DS ₂ -VASc = 3	0.036	[117]
CHA ₂ DS ₂ -VASc = 4	0.054	[117]
CHA ₂ DS ₂ -VASc = 5	0.083	[117]
CHA ₂ DS ₂ -VASc = 6	0.113	[117]
Proportion treated with (AF-patients)		
Warfarin	0%	Base-case
Apixaban	93%	[5]
Aspirin	0%	Base-case
Costs (€)		
Screening handheld ECG	106	[132]
Invitation screening	2	[132]
24-h ECG	266	a
Apixaban (yearly)	844	[103]
Stroke ≤1 year	18 172	[129]
Stroke >1 year	4336	[129]
Severe bleeding	2927	[133]
Minor bleeding	40	Assumption
Quality of life		
Age 50–59 years	0.839	[130]
Age 60–69 years	0.808	[130]
Age 70–79 years	0.794	[130]
Age 80–88 years	0.733	[130]
QALY-loss ischemic stroke (yearly)	0.15	[43]
QALY-loss bleeding stroke (yearly)	0.30	[43]
Proportion unknown AF in all ages		
	25%	[5]
Yearly spontaneous detection of asymptomatic AF		
	5%	Assumption
Accepted OAC therapy after screening	93%	[5]
Discontinuation rate <2.5 years after initiation	11%	[118]
Discontinuation rate >2.5 years after initiation	8%	[120]

^aCost at department of cardiology, Linköping University Hospital. OAC: oral anticoagulants, AF: Atrial fibrillation, ECG: Electrocardiography, QALY: Quality-adjusted life-year.

3.2. Finding the optimal design of screening

This second section of the Method and Data chapter presents the methodology and data used for studying the research questions of how an optimal screening program design can be identified. The model developed for evaluating the screening for unknown AF was used also as a basis for this analysis.

Finding the optimal design of screening for atrial fibrillation

In Sweden, approximately 80% of all patients diagnosed with AF are between 55 and 85 years. [39] Thus, it seems reasonable to assume that screening for unknown AF is most appropriate for individuals within this age interval. We have also assumed that a maximum of one screening occasion every year would be adequate. Using these assumptions for screening initiation and repeated screenings, there are 2 147 483 648 different possible designs of programs with population screening for unknown AF. These programs are defined in Table 7.

Simulation

In the model simulation, all these designs of screening were tested and analysed in MathWorks MATLAB. A thousand individuals were simulated for invitation to be screened or not for each possible screening design. The thousand hypothetical individuals simulated without screening served as reference group in the analysis. This was executed using an action variable (A).

The Action variable

A was added to the original model. [7] The variable is essential for this framework and implies that at each annual decision epoch, there is a decision problem with two options; individuals are invited or not invited to screening. Hence, in this framework, A was used to define investigated screening designs with respect to age of screening initiation and repeated screening. It should be noted that the A used in this framework represents the same actions as decision nodes in a POMDP model. However, in this framework, the variable was modelled as a matrix with unique rows containing 1s and 0s, representing screening and no screening, respectively:

$$A_{x,n} = \begin{cases} 1 = \text{Screening} \\ 0 = \text{No screening} \end{cases}$$

We assumed that the screening for AF can only be relevant once per year (t) in a population aged 55 to 85 years. Given this limited annual interval for when the screening can be conducted, the matrix has 31 columns (31 decision epochs). The number of unique rows in the matrix are then by

definition 2 (as a binary number system was used) raised by 31. As we used a systematic approach, A was defined and created by transforming a column vector containing an arithmetic sequence of numbers in decimal form, $0, 1, 2, 3, \dots, (2^{(no. \text{ decision epochs})} - 1)$, into a matrix with a binary number system:

$$A = \text{Binary} [0, 1, 2, 3, \dots, (2^{(no. \text{ decision epochs})} - 1)]$$

Hence, the matrix had 31 columns and 2 147 483 648 (2^{31}) rows, where each row vector (a_x) represented a specific unique design of the screening program, in terms of when the screening was scheduled. In all a_x , the first column from the right in the row vector, $a_{x,31}$, represented whether the simulated individuals were invited to screening at the age of 55 years, the second, third, ..., 31st column from the right represented invitation at the age of 56, 57, ...85 years, respectively. The A used in our study is described in Table 7.

Table 7. Definition of screening designs (a_x) in the Action variable (A) and how these were created in the simulation model.

Program design	Decimal number	Binary number/ Row vector [a_x]	Design definition
a_0	0	[0 ... 0 0 0 0 0 0 0 0 0 0 0 0]	No screening (reference design)
a_1	1	[0 ... 0 0 0 0 0 0 0 0 0 0 0 1]	Screening at the age of 55 years
a_2	2	[0 ... 0 0 0 0 0 0 0 0 0 0 1 0]	Screening at the age of 56 years
a_3	3	[0 ... 0 0 0 0 0 0 0 0 0 0 1 1]	Screening at the ages of 55 and 56 years
a_4	4	[0 ... 0 0 0 0 0 0 0 0 0 0 1 0 0]	Screening at the age of 57 years
a_5	5	[0 ... 0 0 0 0 0 0 0 0 0 0 1 0 1]	Screening at the ages of 55 and 57 years
...	...	...	...
$a_{34603008}$	34 603 008	[0 0 0 0 0 1 0 0 0 0 1 0 ... 0]	Screening at the ages of 75 and 80 years
...	...	...	...
$a_{2147483647}$	2 147 483 647	[1 ... 1 1 1 1 1 1 1 1 1 1 1 1]	Screening every year 55 to 85 years.

Hence, each row vector (a_x) was sequentially and exclusively employed in the simulation model using a syntax starting with a_0 followed by $a_1, a_2, \dots, a_{2147483647}$. With this approach, the lifelong costs and effects of all possible program designs were estimated.

Optimisation

All screening designs were expected to be associated with different estimated QALYs and costs. The designs were considered mutually exclusive describing the course of action that can be taken over a lifetime perspective for 55-year-olds, implying that only one screening program can be implemented. Hence, standardised cost-effectiveness decision rules from mutually exclusive programs were applied in the analysis. [83] This meant that the optimal screening design is when the cost of producing one additional

QALY (the marginal cost, which is technically the same as the ICER) is equal to the marginal value of gaining a QALY (λ).

A relevant ICER-curve was obtained by using the estimated costs and QALYs of all the 2 147 483 648 designs. All designs that were dominated by another program were removed, as these programs can never be considered relevant. Additionally, a large amount of designs was ruled out by extended dominance. The remaining a_x represented the cost-effectiveness frontier. [83] The frontier then included designs that were increasingly more effective but also more costly. The slope of this curve, by definition, represented the ICER-curve. Hence, the ICER was obtained from the derivative of the frontier. Designs with an ICER over €100 000 per QALY gained were removed from the analysis to make the interpretation of figures easier (as these programs would most likely never be considered relevant).

In this analysis, the λ was assumed to be the value for the health care sector to produce one more unit of the outcome (QALY). λ can be interpreted as the threshold value of a QALY or the WTP for a QALY as described in section 2.4.

This optimisation provided the optimal result for 55-year-olds. The youngest age of screening (t_y) in any of the designs included in the cost-effectiveness frontier also provides us with information about what age it is relevant to start analysing screening (i.e. from which starting age the decision problem should be analysed). As screening was irrelevant before this age, the result obtained using a 55-year-old population is also valid for all individuals equal or younger than t_y years, born and not yet born. Hence, in this thesis this screening program is called the long-term optimal screening program.

Handling of the current population

The age of 55 years was chosen as the starting age of the simulation model in the base-case in order to be conservative and confident that no relevant starting age for the screening program was excluded. However, if a screening program is to be implemented based on this analysis, a relatively large proportion of the current population is already older than the youngest age of screening included in the frontier (t_y). In the long run, one can argue that this is not an issue. However, in a short-term perspective, a separate decision problem concerns how to handle the individuals older than t_y ; they may still benefit from screening although the design of the program cannot strictly follow that of individuals younger or equal to t_y years old. Hence, we need to calculate new cost-effectiveness frontiers for these individuals.

A was used also for the purpose of evaluating this decision problem. A specific age group in the current population is either younger than our

starting age of the model (≤ 55 years) or j years older than the starting age ($55+j$ years). For those older than t_y , a new limited action matrix was created only including a_x with 0s in the columns with a column number higher than j (A_j) (column numbers are counted from the left in the matrix; see Table 7.) as these individuals evidently can never be screened at these ages.

As the results for all a_x were already calculated in the previous part of our analysis, it was not required to recalculate the results for all the designs. Instead, by using the row numbers (in the original A) for the designs included in A_j , we could sort out the relevant results obtained by using the original A . The optimisation analysis was then conducted on the designs included in A_j using the same approach as previously presented for A .

Moreover, this was systematically handled for the entire population by calculating the optimal program designs when $j=1, 2, \dots, 30$ (i.e. 56, 57...85-year-old individuals) using a syntax. If t_y is known, only $j > t_y - 55$ needs to be analysed. However, from a programming point of view, it is easier to handle t_y as unknown and calculate the result systematically for all values of j . This additional analysis can thus provide an answer to how screening should be optimally handled for every age group in the current population of above t_y years of age.

Sensitivity analyses

To assess the robustness of our results, deterministic two-way sensitivity analyses were performed on the assumptions made in our model. Furthermore, some PSAs were conducted to study the statistical uncertainty. Further, additional sensitivity analyses are presented in section 3.3.

3.3. Importance of changes in the screening setting

This third section of the chapter describes the evaluation of the impact of changes in the screening setting. This was analysed using two different cases of changes in the screening setting; a change in the treatment of the sought-after condition and changes in the severity of the condition the screening aims to prevent. The basis for both the analyses was the case of screening for previously unknown AF.

Altered treatment options

As previously mentioned, of great importance for the benefit of the screening is the treatment offered in case of a detected disease. Warfarin has traditionally been the treatment for stroke prevention but NOAC has been introduced during the last few years. The stroke prevention treatment is vital

parameter for the effectiveness of screening with the purpose of detected unknown AF. The importance of this change on the cost-effectiveness of the screening needs to be studied. Briefly this was done by populating the model presented in 3.1 with warfarin data. Efficacy data for warfarin was obtained from a meta-analysis by Hart et al. [48] Additionally, unit cost from FASS was used to calculate the changes in costs. [103]

Change in the condition the screening aims to prevent

The only treatment for acute ischemic stroke has for a long time been intravenous thrombolysis. As described in section 2.3, thrombectomy has now been added to the treatment flora, changing the outcome for many stroke survivors. The main purpose of screening for AF is to avoid stroke, and in extension improve potential stroke patients' lives. When thrombectomy improves the outcome of stroke it also decreases the potential benefit of screening for AF in many patients. A separate cost-effectiveness analysis was conducted to include this effect on both costs and health outcomes in the screening model for AF.

Analytic approach

Using a simulation model, we estimated the costs and effects of adding endovascular treatment to hypothetical individuals who matched the population of the five randomised studies previously mentioned in section 2.3 in terms of age and gender distribution. [71, 75-78] Figure 12 describes the core health states and fundamentals of the model.

Figure 12. Core health states and fundamentals of the thrombectomy model.

Outcome after stroke

The short-term clinical outcome, defined in this analysis as the outcome during the first 90 days, was for the base-case scenario entirely based on pooled data from the ESCAPE, EXTEND-IA, MR CLEAN, REVASCAT, and SWIFT PRIME studies. [71, 75-78] The main outcome data applied from these studies were the modified Rankin Scale (mRS) data obtained after 90 days. It is commonly used as a tool to measure the degree of dependency or

level of global disability after a stroke (0 implies no disability; 6 indicates dead). The global disability has a strong connection with the amount of costs and quality of life in patients, which makes it a suitable tool for this analysis. [66, 68]

The randomised, multicentre studies ESCAPE, MR CLEAN, REVASCAT, and SWIFT PRIME all used mRS after 90 days as the primary endpoint while EXTEND-IA had mRS at 90 days as secondary endpoint. Details of the pooled data from the five studies, which were used in the model as the base-case scenario, are presented in Figure 13.

Figure 13. Functional outcome at 90 days in included studies and the pooled data set. The functional outcome according to the score on the modified Rankin Scale. The bars show the primary outcomes from the ESCAPE, EXTEND-IA, MR CLEAN, REVASCAT, and SWIFT PRIME studies. All studies represent pooled data from the five studies. ET: Endovascular treatment, SC: Standard care.

The short-term outcome after the first 90 days was extrapolated to a lifelong time horizon based on the expected stroke recurrence and mortality in the simulated population. The rate of stroke recurrence was obtained from a study of patients in northern Sweden. [134] The global disability after a recurrent stroke was assumed to follow the same distribution as presented in the five previously-mentioned randomized studies. However, the patients could only transition to health states with the same or worse mRS score as compared to the pre-re-stroke health state.

In the base-case scenario, we have assumed that all simulated individuals, irrespective of their mRS scores, had approximately twice the mortality of the general public of Sweden (matched according to gender and age). [95] This assumption was based on two studies: a study of long-term stroke survival in individuals with AF and a study of first-ever stroke survivors in Australia. [135, 136]

Utility weights

Utility weights based on population data were used as a base for the model. [130] Utility decrements based on the global disability (mRS) after stroke and re-stroke were used throughout the model. [66] Since the observational study that was used in our model was only available as an abstract, the sensitivity of this parameter was tested using the only other published source of data we could identify. [136]

Cost

The short-term costs after the initial stroke and re-stroke were obtained from a Swedish study of cost for stroke that included data regarding the costs incurred for acute visit, rehab and follow-up, drugs, home assistance, and residential housing. [129] Cost multipliers (relative costs), as presented by Dawson et al., was used to distribute the costs according to the mRS score. [68]

The long-term costs were based on the costs incurred for stroke recurrence and the global disability (mRS). [129, 137] The cost of adding thrombectomy to standard care was obtained from the university hospitals in Gothenburg, Linköping and Stockholm. All the costs were adjusted according to the inflation rate for the year 2015 prices and the exchange rates applicable as of March 1, 2015. The details of the cost data are presented with other input data in Table 8.

Table 8. Summary of important parameters in the thrombectomy model.

Parameter	Base-case value	Reference
Age (years)	67.2	[71, 75-78]
Male (%)	53.2	
Time horizon	Lifelong	
Discount rate (%)	3	
Stroke recurrence year 1–10 (annual %)	0.02–0.06	[134]
Long-term hazard ratio stroke	2.0	[135]
QALY weights		
General population		
60–69	0.81	[130]
70–79	0.79	
80–88	0.73	
Stroke decrements		
mRS = 0	0.02	[66]
mRS = 1	0.24	
mRS = 2	0.25	
mRS = 3	0.48	
mRS = 4	0.59	
mRS = 5	0.84	
Procedure costs (incl. materials)		
Adding thrombectomy (US\$)	7 908	^b
Cost stroke (US\$)		
Direct cost in year 1	20 108	[129]
Index stroke admission (incl. ambulance)	12 444	
Stroke readmission	362	
Rehab and follow-up	1 108	
Prevention (drugs)	258	
Home assistance and residential housing	5 935	
Direct cost in year 2+	6 247	
Follow-up visits	140	
Prevention (drugs)	229	
Home assistance and residential housing	5 878	
Cost multiplier according to mRS		
Year 1		[68]
mRS = 0	0.25	
mRS = 1	0.33	
mRS = 2	0.52	
mRS = 3	0.63	
mRS = 4	0.82	
mRS = 5	1	[137]
Year 2+		
mRS = 0–3	0.39	
mRS = 4–5	1	
Stroke-related costs first year (US\$)		
mRS = 0	7 821	[68, 129]
mRS = 1	10 267	
mRS = 2	16 064	
mRS = 3	19 512	
mRS = 4	25 538	
mRS = 5	31 083	
Annual long-term stroke cost in year 2+ (US\$)		
mRS = 0–3	3 169	
mRS = 4–5	8 118	

^a95% confidence interval was used for probabilistic analyses. mRS: modified Rankin Scale. ^b Cost obtained from University Hospitals in Linköping, Göteborg and Stockholm.

Implementing the results in the screening model for AF

The thrombectomy- and the AF- model were connected in Microsoft Excel to incorporate these results, obtained from the cost-effectiveness analysis, in the screening analysis. The calculated average loss of QALYs the stroke survivors experienced in the screening model of screening of 75-year-olds without thrombectomy were adjusted based on the result of the thrombectomy. Additionally, cost for a third-party payer perspective for endovascular treatment was included in the AF model.

This was done by calculating costs and QALYs in thrombectomy model for the patients fulfilling the criteria for screening in the STROKESTOP-study. Additionally, the number of the patients experiencing acute ischemic stroke that would benefit from thrombectomy were estimated. This has been estimated to be 15% of all patients with acute ischemic stroke in Sweden. [138, 139]

4. RESULTS

This chapter presents the results of the case studies and the additional analyses. More details of the case studies can be found in the papers and manuscripts that are available at the end of the thesis.

4.1. Long-term cost-effectiveness of screening

The cost-effectiveness of screening for colorectal cancer

Base-case scenario

In the first case, the CRC-model enabled lifelong simulations of 60-year-old hypothetical individuals in each of the five different screening strategies. Depending on the strategies implemented, different participation and detection rate of colorectal-related conditions are expected. The cumulative number of CRC-related deaths, when 1000 hypothetical individuals are invited and simulated through the five different screening strategies, are shown in Figure 14. When only analysing screening program designs that were used in the SCREESCO-study, OC implied the fewest number of CRC-deaths. However, repeated screening with OC after 10 years and FIT biennially resulted in the largest decrease in CRC-related deaths. But as can be seen in the figure, even with frequent screening, CRC related death can be expected.

Figure 14. Number of CRC-related deaths when 1000 hypothetical individuals are simulated in the five treatment arms. All strategies were compared to no screening.

Figure 15 shows the cumulative net cost of the four strategies with screening compared to no screening. As presented in the figure, the screening, performed in the SCREESCO-study, implied a short-term cost increase

but a long-term cost-saving. Repeated screening implied, as expected, higher total cumulative costs as shown by the spikes in the curves in the figure. This is due to costs occurring from the screening procedures.

Figure 15. Cumulative incremental cost of screening with OC and FIT when 1000 hypothetical individuals are simulated in the five treatment arms.

Costs and QALYs per 1000 invited individuals of all the five different screening strategies are presented in Table 9. It should be noted that, compared to no screening, all four programs showed a reasonable cost per QALY but screening once with OC implied 49 QALYs gained to a lower cost, while OC every 10th year resulted in the largest gain of QALYs but to a total cost increase of €124 600. Screening with FIT twice and biennially also implied a low cost per gained QALY but corresponding strategies were associated with smaller gains of QALYs and higher costs compared to screening with OC (Table 9).

Table 9. Results of 1000 individuals invited to screening with OC and FIT compared to no screening.

	Compared to no screening				Incremental cost per QALY versus:		
	Additional cost (€)	Avoided CRC	Added life-years	Additional QALYs	No screening (€)	OC once (€)	FIT twice (€)
OC once	-64 800	16	52	49	-1300	-	-2000
FIT twice	-17 600	7	28	26	-700	Dominated	-
OC 10 th year	124 600	21	59	56	2200	26 900	4800
FIT 2 nd year	136 700	18	54	51	2700	81 500	6200

Sensitivity analysis

The reliability of the results was tested using the sensitivity analysis. Table 10 presents the importance of assumptions and uncertain parameters for the health economic result. The result was sensitive to changes in the discount rate as observed in most screening programs. This is due to the fact that positive effects could be observed several years after the screening. As shown in the table, the cost-effectiveness results were not sensitive to changes in the participation rates.

Table 10. Two-way sensitivity analyses. All results are compared to no screening per 1000 invited individuals.

	Refer- ence	Participation (ab- solute)		Discount rate		Prevalence CRC (age 60y)	
		-10%	+10%	0%	5%	-50%	+50%
Additional costs (€)							
OC once	-64 800	-43 700	-78 500	-245 000	18 900	-54 000	-72 800
OC 10 th year	124 600	86 000	151 900	-32 300	183 600	134 300	115 000
FIT twice	-17 600	-12 500	-22 700	-76 700	10 400	8600	-26 700
FIT 2 nd year	136 700	112 700	160 600	5 100	178 900	145 900	127 500
Additional QALYs							
OC once	49	34	61	87	34	48	50
OC 10 th year	56	40	68	101	38	54	57
FIT twice	26	20	30	45	18	26	27
FIT 2 nd year	51	40	60	96	34	50	52

The Cost-effectiveness of screening for AF

Base-case scenario

In the second case study, 1000 screened 75-year-old individuals were simulated in the base-case. The number of individuals, diagnosed with AF with screening compared to no screening, is presented in Figure 16. Our life-long model simulation showed that mass screening results in earlier and more detection of AF. The purple area, in Figure 16, shows the number of patients with detected AF without screening, while the red line shows the number of patients with detected AF when screening was implemented. The yellow area between the red and the purple line is the incremental number of patient-years with detected AF resulting from the screening. This area is the real effect of the screening and could be considered as patient-years in a lower risk of stroke due to treatment with OAC.

Figure 16. The number of patients experiencing AF when 1000 individuals are simulated in the model. The red and purple lines show the number of individuals with detected AF. AF: Atrial fibrillation.

In the base-case scenario presented in Table 11, screening of 1000 individuals resulted in 263 more patient-years with detected AF (yellow area). Initiation of NOAC treatment in many of the screened individuals resulted in more major bleedings but fewer strokes in the population. This implied 11 more life-years and 12 QALYs gained per 1000 screened individuals. The screening of 75-year-olds implied an incremental cost of €50 012, resulting in a cost of €4 313 per gained QALY and €6 583 per avoided stroke.

Table 11. Base case scenario for 1000 screened individuals in the AF screening model.

	Lifetime costs	Strokes	Life years	QALY	Cost per gained life year	Cost per gained QALY
No Screening	€3 885 879	92	9 835	6646		
Screening	€3 935 891	85	9 847	6657	€4365	€4313

QALY: Quality-adjusted life year.

Sensitivity analyses

In order to study the uncertainty in our results for 75-year-olds, the statistical uncertainty was tested probabilistically 10 000 times. The result of the PSA, presented as acceptability curves in Figure 17, shows that if the WTP for a QALY is higher than approximately €5000, screening with handheld ECG at the age of 75 years is probably cost-effective.

Figure 17. The number of patients experiencing AF when 1000 individuals are simulated in the AF- model. The red and purple lines show the number of individuals with detected AF. AF: Atrial fibrillation.

Table 12 presents the deterministic sensitivity analyses. The results of the simulation model are sensitive to changes in the time horizon, the prevalence of undetected AF, and the stroke risk in asymptomatic AF.

Table 12. Deterministic sensitivity analyses of screening for previously unknown AF.

Scenario	Value	Cost per gained QALY (€)
Time horizon	5 years	78 049
	10 years	19 092
	15 years	7855
Discount rate	0%	34
	5%	7 874
Prevalence of unknown AF in 75-year-old individuals	1%	26 892
	5%	373
Yearly spontaneous detection of AF	0%	2590
	10%	6232
OAC adherence	50%	12 423
	100%	3661
CHA ₂ DS ₂ -VASc (mean)	2,5	16 176
	4,5	-628
Asymptomatic AF stroke risk compared to symptomatic AF	+ 25%	-1467
	- 25%	15 090
	-50%	40 422
Discontinuation rate OAC	20% year 1 (Decreases yearly with 30%)	10175

4.2. Finding the optimal design of screening

This second section of the chapter presents the result of the optimisation analysis on the screening program for unknown AF.

The long-term optimal screening program design

All different designs of the screening program for AF were simulated in the model using A. When all the 2 147 483 648 possible designs of screening programs were considered, screening at the age of 75 years demonstrated the lowest cost per QALY gained (€4800/QALY). The upper part of Figure 1 shows the seven program designs that remained after the optimisation analysis had been conducted. This curve is the cost-effectiveness frontier of different AF screening program designs for 55-year-olds and these are in this thesis called long-term optimal program designs. It can be seen in the figure that it is possible to gain increasingly more QALYs by changing the design of the screening program, and that the change in design results in higher costs. As shown in the lower part of the figure, the derivative of the frontier, the improvement in the QALY gain is achieved at increasingly higher marginal ICERs. In other words, if the number of screening occasions was increased, this resulted in earlier and more detection of AF but at a higher cost. The youngest age of screening initiation in any of the long-term optimal program designs, t_y , was 65 years, and these designs are therefore relevant for all individuals currently 65 years old or younger.

Figure 18. Cost-effectiveness frontier (upper part) and the MC (lower part) of producing QALYs using screening for AF when 1000 individuals are simulated.

The interpretation of cost-effectiveness depends on λ . As indicated in the lower part of Figure 18, we should not implement any screening program (a_0) if λ is €4 800 or lower (the slightly different result compared to the result in section 4.1 is mainly due to differences in the prevalence of AF in the general population of 75-year-olds in Sweden and the STROKESTOP-population and different exchange rates). Screening at the age of 75 years ($a_{1048576}$) is cost-effective if λ is between €4800 and €31 000. Initiating the screening at the age of 75 years and repeating it at the age of 80 years ($a_{34603008}$) is cost-effective if λ is between €31 000 and €37 000. In addition, screening initiated at the age of 68 years and then repeated at the age of 75 and 80 years ($a_{34611200}$) is cost-effective when λ is between €37 000 and €40 000. Hence, the y-axis on the ICER-curve can be interpreted as a cost-effectiveness ranking list of the different screening designs. Depending on the relevant threshold value, different screening designs will be chosen for implementation (Table 13).

Table 13. Outcome of the optimal designs of screening for AF with handheld-ECG.

Age when screening is conducted (years)	Cost per gained QALY compared to no screening	The design is the best option when the threshold is (€):
All		
No Screening	-	<€4800
75	€4800	€4800–31 000
75, 80	€9500	€31 000–37 000
68, 75, 80	€15 800	€37 000–40 000
65, 75, 80	€16 400	€40 000–74 000
65, 75, 78, 82	€19 500	€74 000–77 000
65, 70, 75, 78, 82	€24 500	€77 000–100 000

Current population

The current population above t_j (65 years of age, $j=20$) faces another decision problem as these would not be subjected to all the screening designs identified in the base case results (the long-term optimal program designs). Hence, the cost-effectiveness frontiers for all age groups older than 65 years (all $j \geq 10$) were obtained from the original analysis. For visual reasons, only optimal designs for 75 ($j=20$) and 80 ($j=25$) year-olds in the current population are presented in relation to the long-term optimal results in *Figure 19*. As shown in the figure, the results only differ when the age of the current population ($55 + j$ years) is older than the starting age of a long-term optimal design.

Figure 19. The cost-effectiveness frontiers using a 55 (purple lines), 75 (red lines) and an 80 year old population (green lines).

Sensitivity analyses

The sensitivity analyses of the major assumptions in the model are presented in Table 14.

Table 14. Sensitivity of cost per QALY to major assumptions used in the AF screening model.

	Reference (Cost per gained QALY vs. no screening)	Cost of the screening		Spontaneous detection		Discount rate		Initial OAC adherence	
		-25%	+25%	0	10%	0%	5%	50%	100%
75 years	4800	2300	7300	3000	6900	400	8400	14 100	4100
75 and 80 years	9500	6100	12 900	8000	11 200	4500	13 500	22 300	8500
68,75 and 80 years	15 800	11 000	20 500	13 600	18 200	8 400	22 200	31 000	17 000
65, 75 and 80 years	16 400	11 500	21 200	13 800	19 200	8 300	23 700	32 100	15 100
65,75, 78 and 82 years	19 500	13 800	25 100	17 100	21 900	11 200	26 900	37 600	18 100
65, 70, 75, 78 and 82 years	24 500	18 100	31 000	22 400	26 800	14 900	33 100	45 200	22 900

PSAs, where the statistical uncertainty of the model was studied, show the high uncertainty in these analyses (Figure 20). In the figure, the probability of the different optimal screening being the best is presented. For visual reason, the no screening option is not included in the figure.

Figure 20. Probability of being the optimal screening design given the decision to use screening.

4.3. Importance of changes in the screening setting

This section presents the results of two different types of changes in the screening setting.

Altered treatment options

The change in treatment from warfarin to NOAC has affected the cost-effectiveness results of the screening. The ICER of the screening sunk from €4 173 when using NOAC, to €3784 when using warfarin (a decrease by 9%). This was primarily due to lower pharmaceutical costs, from €264 to €174. As shown by the ICER, this increase in cost, due to the introduction of NOAC has not been fully compensated by the increase in QALYs gained (0.0089 to 0.0116). Hence, the ICER of using NOAC instead of warfarin in the model was €5371 per QALY gained.

Change in the condition the screening aims to prevent

Endovascular thrombectomy

The first part of this analysis was to estimate the long-term costs and effects of the introduction of thrombectomy. The results were then applied in the screening model for AF to study how this change affected that screening.

The lifelong model showed that adding endovascular treatment to guideline-based care resulted in a gain of 0.4 life-years and 0.99 QALYs along with a cost saving of €197 using pooled data. Figure 21 shows the proportion of the hypothetical individuals that were alive and the average utility weight in the model. Additionally, the figure shows that the treatment becomes cost neutral for the healthcare payer after approximately 17 years.

Figure 21. The curves show the estimated survival, mean utility weight and cumulative incremental cost in the thrombectomy model. Cumulative incremental cost represents the net cost of adding endovascular thrombectomy to standard care. ET: Endovascular treatment, SC: Standard care.

Effect on the effectiveness of screening for atrial fibrillation

The average loss of QALYs due to stroke, in the screening model of screening for 75-year-olds without thrombectomy, was predicted to be 1.53 QALYs per stroke. If all individuals would be considered as suitable patients for thrombectomy and receive the above presented benefit, this loss would sink to 0.54 QALYs per stroke. Additionally, the incremental cost of screening would increase modestly from €48 to €50 per individual. This would imply that the ICER of screening of 75-year-olds would increase from €4173 to €12 242 per QALY (an increase by 193%).

However, only 15% of all stroke patients are expected to receive thrombectomy treatment due to many being medically non-eligible or arriving too late to a thrombectomy performing hospital. [139] This implies that the QALY loss per stroke would instead only sink to 1.38 QALY per stroke. The screening cost per patient would increase to €49 per screening individual. This implies an ICER of €4645 instead of €4173 (an increase by only 11%).

5. DISCUSSION

In this thesis, I have investigated applications and problems associated with health economic evaluations of screening programs. This was done using two cases: screening for CRC and screening for unknown AF. Further, to investigate the effects of changes in the screening setting, non-screening technologies were included in the thesis as scenario analyses. Hence, the attempt to answer the research questions stated in section 1.2 was based on two case studies and two scenario analyses.

In this chapter, the decision problems relevant for the two case studies are first briefly discussed. In the second section of the chapter, I discuss the methodological problems that were encountered and handled throughout this thesis.

5.1. Case discussion

Cost-effectiveness of screening for colorectal cancer

In the first case study, our analysis showed that screening for CRC is highly cost-effective, which is similar to several previously published studies. [8] However, the screening programs differ between countries in terms of timing of implementation and contextual differences, such as different population preferences. Development of the test and new knowledge of the organisation of programs will continuously change the conditions and demand new assessments when new programs shall be implemented. Unlike previously published cost-effectiveness analyses, our study was partly based on screening data from early experience in a head-to-head study of OC and high-sensitive FIT. When these tests are compared, the often-debated factor is the participation rate. Projected data from early experience in the SCREESCO-study showed a lower adherence to both OC (38%) and FIT (50%) than expected. Even when we used lower participation rates than were estimated in the study design, the investigated screening strategies in the SCREESCO-study implied both clinical and economic gains. OC once was dominant compared to screening with FIT twice, resulting in a better clinical outcome to lower costs when using a lifelong perspective in a Swedish context. This screening strategy also implied the lowest cost for gaining QALYs, but if the health care payer is prepared to pay more than €26,900 per QALY, additional QALYs can be gained to a reasonable cost using screening with OC every tenth year. Cost-effectiveness is one of the several aspects to consider in decisions regarding implementation of a CRC screening program. Also, other factors such as patient acceptability of the

test and availability of personnel have to be taken into consideration. Furthermore, OC is a rather complicated way of bowel examination that requires sufficient access to OC facilities and a good quality of OC procedure. However, this is not the case in many countries. The shortage of endoscopists and variability and quality of the examination are other aspects to consider that require further investigation.

A limitation, also present in most cost-effectiveness studies, is that it is difficult to include all types of effects occurring from the screening. For instance, a parameter that was not included in our analysis was the small but existing effect of false positive results on the quality of life. It is likely that such patients have increased levels of anxiety during a short time period. Including this parameter in our model could have a slightly negative effect on the screening results.

Another limitation was that we only compared OC, high sensitive FIT and no screening because these were used in the SCREESCO-study. Several other screening technologies are available and should be studied in further analyses. For instance, CT-colonoscopy has been shown to be a cost-effective screening alternative. [111] A similar problem was that the SCREESCO-study included only three treatment arms: no screening, screening once with OC and twice with FIT. In our analysis, we added two additional treatment arms with repeated screening, but in a real-world context, there are many more options in terms of when the screening is initiated, how often it is repeated and the possibility of different combinations of screening tests in the same program. As shown in this thesis, finding the best design of a screening program is of vital importance for an optimal resource usage.

Cost-effectiveness of screening for atrial fibrillation

The second case study analysed the screening for previously unknown AF. Our analyses showed that the AF screening program studied in the STROKESTOP-study implied a cost of €4100- 4800 per QALY, depending on the source of input data and the date of analysis. Irrespectively, this ICER is traditionally regarded as low. If this screening program, studied in the STROKESTOP-study, is used in a country with 10 million residents (Sweden), approximately 36 000 (0.4% of the total population) individuals aged 75-years would be suitable for screening every year (participation rate: 53.5%). [5] Based on this analysis, the screening for asymptomatic AF would result in approximately 270 fewer strokes, 410 more life-years, and 420 more QALYs to an incremental cost of €1 800 000.

At the time of our literature review, only a few other cost-effectiveness analyses of screening for AF in relevant populations were available. [55, 132, 140-142] Levin et al. found that screening of individuals, who have previously experienced a stroke, implies lower costs and gained QALYs compared to no screening. [128] In the same study, the authors showed that screening with handheld ECG for 30 days was both less costly and

more efficient in finding AF than the 24-hours Holter ECG. Lowres et al. presented a cost of €3142 per gained QALY when screening for AF in pharmacies using iPhone electrocardiogram. [140] The results from that model-based analysis are similar to the results from our study, despite being conducted in another setting (Australian) and an older population (mean age was 79 years). Also in line with our results, Maeda et al. found that screening with pulse palpation in Japan implied an ICER of \$8000-\$10 000 per QALY gained. [141] The Health Information and Quality Authority in Ireland found that screening every five years from the age of 65 years with pulse palpation implied a cost of €9,471 per QALY gained. [142] In a piggy-back study, Hobbs et al. showed that opportunistic screening of 65-year-old individuals in the UK with pulse palpation implies a cost of £363 per case detected. [55] Further, studies have shown that insertable cardiac monitors and 14-day ambulatory ECG adhesive patch monitors are also effective in identifying previously unknown AF. [52, 140, 143-146] Hence, a limitation of our study was to only include handheld ECG as the technique to find previously unknown AF. Future studies need to investigate the relative effectiveness of these continuous monitoring devices compared to intermittent long-term screening techniques such as handheld ECG, pulse palpation and MyDiagnostick.

Another limitation in the present screening model was that we assumed asymptomatic AF to be associated with the same risks as symptomatic AF. This assumption was based on the findings of trials on the device-treated and type 2 diabetic patients. [124, 125] However, the correctness of this assumption is yet to be established. This uncertainty was also the main reason for the National Board of Health and Welfare to give a negative recommendation for AF screening in Sweden. However, the cost per QALY gained is reasonable even if this risk is 50% compared to the risk in symptomatic AF (Table 12). Long-term follow-up data from the STROKESTOP-study will provide important data for this parameter in future models and will be available in 2018.

The parameters used in this model simulation were mainly based on STROKESTOP-study data from the Stockholm County. As shown in the STROKESTOP-study, the participation and the proportion of previously undetected AF were lower in this area compared to the rural region of Halland. [147] If Halland is more representative of the overall population of Sweden, then the cost per gained QALY and avoided stroke would be lower. A pilot study prior to the STROKESTOP-study also presented relatively large differences in participation between demographic areas. [148] Additionally, this study found a correlation between areas with low participation, high proportion of immigrants, and a high CHA₂DS₂-VASc-score. As shown in Table 12, a higher risk of stroke implied a lower cost-effectiveness ratio. This could motivate tailored efforts to increase the uptake in such areas. Despite not being important for the cost-effectiveness, the participa-

tion rate is considered as an important parameter from public health perspective. Screening programs that are cost-effective should be designed to include as many participants as possible. Simplifying the screening procedure and other methods to increase the participation rate from the STROKESTOP should be discussed in further analyses.

Further, as the cost per gained QALY is relatively low for the program studied in the STROKESTOP-study, the discussion should not be limited to whether a screening program should be applied in the healthcare or not but should also include whom, when and how it should be implemented. Hence, in our optimisation analysis, we systematically studied, evaluated, and identified numerous different designs of screening programs for unknown AF that could be considered optimal. With a threshold value of approximately €50 000 per QALY, which often has been used in the literature for Sweden, [149, 150] the most optimal program would be to screen at the ages of 65, 75, and 80 years. In comparison with the program design used in the STROKESTOP-study, [5] this design is expected to reduce the number of strokes with 1.4 and gain 1.5 additional QALYs to a total cost of approximately €49 800 per 1000 invited individuals. This implies a cost of €31 000 per gained QALY. In a country sized as Sweden, this program design implies 174 gained QALYs at a cost of €5.8 million compared to the screening performed in the STROKESTOP-study (screening once at the age of 75 years).

5.2. Methodological discussion

The points discussed in the previous section are important for making policy decisions in the case studies that were included in this thesis. However, all the results of the cost-effectiveness analyses rely on the used methodology. Hence, in this section, I discuss some of the different methodological questions that were encountered and addressed during the work.

Modelling analyses

The first topic concerns the use of modelling analyses in general. Modelling analyses are considered as an alternative but are primarily a complement to clinical studies for analysing medical decision problems (where RCTs are considered as the golden standard). RCTs with sufficient follow-up time for capturing all relevant cost and effects occurring from an intervention are rarely available when economic evaluations are undertaken. Waiting for long-term follow-up data from the RCTs before making decisions about an implementation would often imply an unacceptable time lag for screening technologies. Hence, as in the cases covered in this thesis, when long-term real-world or study data are absent and policy decisions are required to be made, well-grounded models based on the overall knowledge of the area are useful tools for decision makers. In these analyses, it is important that

the best available data are used and obtained through systematic and methodological reviews of the existing literature. But, it is important to bear in mind that simulation models often have certain limitations due to it being impossible to find accurate data on lifelong costs and effects for calculating the cost-effectiveness. This problem increases with the increased requirement of data when multiple screening designs are evaluated. In the studies included in this thesis, national registries with good coverage were employed to reflect the best evidence available. [39, 117, 130] However, obtaining data of certain parameters, for instance, prevalence and other disease progression parameters for different age groups, can be challenging due to the lack of clinical evidence in all relevant age-groups. An increased use of data from registries collected in routine clinical practice may, as in this thesis, provide an important source of evidence for this type of modelling.

Further, in order to draw conclusions, the impact of uncertain parameters in this thesis were studied through sensitivity analyses. These analyses studied both the statistical and parametric uncertainty. The methods used for this purpose in this thesis have been widely accepted for several years but needs to be continuously improved in order to help guide decision makers to an optimal resource allocation. [80]

Interpretation of the optimisation analysis

In addition to studying the statistical and parametric uncertainty, the major contributions of our studies were in regard to structural uncertainty, such as analysing the uncertainty in the continuously changing screening setting, especially providing methods on how to handle a large number of screening designs.

Attempting to address this analytical issue and structural uncertainty of how an optimal design of a screening program could be identified with a systematic approach was one of the important aims of this thesis. Analytically, this was implemented by integrating previously used disease progression modelling techniques with an Action variable (A) constructed as a matrix (see section 3.2). Similar variables have been previously described and applied in screening studies as it provides a flexible and rather simple way of evaluating multiple screening designs based on decision-analytic models. [13] The present analysis showed that seven different program designs could be considered optimal in the long-run with the ultimate choice depending on the threshold value of cost-effectiveness in the health care. Some kind of screening for AF appears to be cost-effective if the threshold is higher than €4100-€4800. Although we recognise that thresholds represent different things and take different magnitudes in different jurisdictions, applying a range of values (€4800-€31 000) that includes many commonly used threshold values, one-off screening at the age of 75 years appears to be the cost-effective design of a screening program. [151] With

an increasing threshold, the results indicate that additional screening occasions may be cost-effective. Further, as can be seen in Figure 18, the optimal screening designs for older individuals (>65 years) in the current population differ from long-term optimal screening designs (long-term optimal imply designs for all younger than the age of the first screening occasion and individuals born in the future).

Hence, based on our results, the optimal designs of the screening program depend on both the threshold value and the age of the investigated cohort. Different health care systems (the UK, Netherlands and Sweden) may implement different designs of a screening program. As shown in Table 15, the optimal strategy would vary, even though the three compared countries are relatively similar with regard to per capita income and funding of health care compared to rest of the world.

Table 15. Ages when screening for previously unknown should be conducted depending on country and cohort.

		Countries/ Thresholds		
		The United Kingdom £20 000 (€24 000)	Sweden 500 000kr (€54 000)	Netherlands €80 000
Co- horts	Long-term optimal	75	65, 75, 80	65, 75, 78, 82
	75-years-olds	75	75, 80	75, 78, 83
	80-years-olds	80	80	80

Thus, the threshold value can, as in our analysis, be decisive in determining both the age of screening initiation and the screening intervals. This highlights the need of knowing the threshold value for a QALY in the health care sector, as it is important not only for deciding if we should implement a technology but also for how the technology should be used.

In general, our optimisation analysis also highlights the need for a more systematic approach to the economic evaluation of screening programs. As indicated, it is inherently difficult to identify optimal screening designs relying solely on comparative clinical studies as every potential design that should be studied needs many subjects and a long follow-up. As disease progression with and without a screening program is often evaluated using decision-analytic modelling, we believe that the proposed framework is a logical extension to illuminate the thorny questions regarding different designs of the screening program. It has been illustrated that a reasonably simple and quick programming approach can provide a lot more information to decision makers regarding the starting age of a screening program as well as the marginal effects of repeated screening occasions.

Screening setting and interpretation of the ICER

In the previous section, the identification of optimal screening designs was highlighted. However, the optimal screening program design for various

countries is not always permanent because, as shown in this thesis, the setting surrounding the screening is of great importance for the cost-effectiveness results, and it often changes over time.

For instance, adding endovascular thrombectomy to standard care for acute stroke is dominant compared to standard care alone. This implies that the intervention does not only provide clinical benefits but also saves cost when using a lifelong perspective. Inclusion of these results had a negative impact on the cost-effectiveness of screening for AF. This effect was minor due to the low number of patients receiving the treatment, but this also shows that screening effectiveness has a strong connection with the screening setting. Hence, when the ICER is closer to the threshold value than in the case of screening for AF, potential changes in the surrounding world are of great importance and need to be analysed.

The second example of changes in the screening setting showed that changes in treatment for those patients with screening detected AF from NOAC to warfarin lowered the ICER for the screening. This may seem surprising as NOAC has been shown to be cost-effective compared to warfarin in patients with AF in Sweden and other countries. [152] This was a result of the fact that the ICER of NOAC compared to warfarin was €5371 per QALY gained in the screening model. The ICER is higher than the ICER for the AF-screening program (that included warfarin treatment). When the ICER of a new treatment in a screening program is higher than the current ICER of the screening, the cost-effectiveness will be negatively affected. This highlights the risk of blindly relying on the ICER in decision-making situations without considering the effect on the public health. The ICER of NOAC treatment compared to warfarin (€5371 per QALY) is still considered relatively low in most health care sectors, and hence, it provides QALYs at a reasonable cost for the society. Choosing a screening program design that uses warfarin instead of NOAC in Sweden would be a poor decision, even though the ICER of this program would seem more favourable at first glimpse.

Section 4.1-2 highlights a similar problem; fewer screening occasions result in lower ICERs compared to screening with more screening occasions. Screening once at the age of 75 years implied the lowest cost per QALY, but as shown in the expanded analysis, adding additional screening occasions could result in larger gains of QALYs and to what in Sweden are considered reasonable costs. Hence, choosing solely based on the lowest ICER without studying the marginal ICER against all alternatives can lead to sub-optimal resource allocation also in this case.

Further expansion of the analysis

In addition to updating the models to the continuously evolving world around the screening, several methodological issues need a further investi-

gation. For instance, in our optimisation analysis, we focused solely on different designs in terms of when the screening was scheduled. In practice, the general methodological approach can be applied to other types of screening optimisations, for instance, when several different screening tests are available. Further, the framework could also be applied in other areas where there is a large amount of possible decision points, such as how to use MedTec products, diagnostics, cut-off levels for surgery, optimal dosing of pharmaceuticals and other subsequent management decisions.

Another area that requires further research is how statistical uncertainty should be handled within this framework. For instance, PSAs of a large number of alternative screening designs can be challenging not only due to limited computer capacity but also in terms of interpretation of the results. When there are a large number of different alternatives, the probability that each of the designs is the most cost-effective will be very low, thus making it hard to interpret the results. In such situations, it is important to fully understand the consequences and the uncertainty surrounding the different decisions. The expected value of perfect information (EVPI) is a framework that, at least partly, can handle this problem. [79] However, while faster methods to perform EVPIs is currently being developed, [153] the large number of alternative screening designs remains a problem. A variety of ways are available to handle this, but methods that do not deviate from the systematic approach are highly recommended. Tailored approaches based on medical characteristics or other parameters of the screening program can be of interest for advancing the study.

For instance, in a scenario where the underlying disease has slow progression or if the screening procedure is physically not possible to carry out within a particular time window, frequent screening occasions may not be relevant. In such situation, A can be used to systematically limit the number of studied designs. This is achieved by excluding designs that are evidently irrelevant from an economical, clinical or otherwise plausible perspective. Technically, this can be handled by systematically putting restrictions to A and creating a new limited action variable, for instance, only including a_x that uses a minimum of two decision epochs (two years) between the screening occasions. This is implemented by not allowing 1s (screening) in adjacent columns in any of the row vectors. From a programming point of view, the limited action variable is then employed in the model using the same approach as with the original A . Incorporating a restriction of at least two years between the screening occasions would, in our example, limit the number of possible designs that need to be studied from 2 147 483 648 to 3 524 578 (a 99.84% reduction of available designs).

In other screening programs, it may, for various reasons, not be possible to perform more than a certain number of screening occasions. For instance, program designs with more than six screening occasions may not

be relevant. We could then create a new limited action variable that contains only row vectors that meet this restriction, resulting in a 99.95% reduction of the available designs.

Another approach to drastically reduce the number of potential programs in the case of screening for AF is to use the fact that this screening implies diminishing returns of the number of screening occasions; screening once will always produce a lower ICER than to screen twice given screening at the same age(s). The same is true for two versus three screening occasions and so on. Using a step-by-step approach, where all designs that use one screening occasion are first tested, then two, three, four... occasions, we can stop adding additional screening occasions when the marginal ICER of the most cost-effective of the last added designs is higher than, for instance, €100 000. In our case, that would imply that we only tested programs with up to six screening occasions which would reduce the number of potential programs to 942 642 (a 99.95% reduction of available designs).

Future studies should explore additional methods concerning how to limit the number of irrelevant designs effectively or provide other approaches that can decrease the computer burden of PSAs and EVPIs.

Methodological alternatives

A closely related method for analysing decision problems of the kind evaluated in this work is MDP models (see section 2.4). It should be noted that our evaluation approach (Markov Model including an action variable) carries the same properties as an MDP. MDPs are often more efficient than Markov models in terms of shorter calculation times as they include several decision nodes embedded in the model structure. Rather than analysing all possible management options one at the time, formulas with reward functions are employed and the optimal management strategy is identified where the cumulative function of all rewards is maximised. There are several situations where MDPs are well suited to handle optimisation problems, for instance, non-reversible decisions such as when surgery for liver transplantation should be performed, [154] when screening with a specific interval should start or at what clinical threshold the treatment should start. [13] In our case, a POMDP could have been used as the model is only partly observable when we do not know which individuals that have unknown AF, only the probability that they have the disease. However, the combination of model states regarding disease progression and possible events implies that the number of reward functions would, for our decision problem, increase exponentially and would become computationally intractable to identify and solve exactly. [155] The solution would be to use approximation or sampling techniques. [81, 155-157] However, an MDP analysis solved with sampling would resemble the analysis performed with

a Markov model in the current work. An example of such solution to a similar decision problem regarding screening for breast cancer has been published by Maillart et al. [156] Therefore, one may argue that the advantage of a shorter calculation time using POMDPs is counterbalanced by an approach where developed disease progression models do not require substantial restructuring to be used for optimisation analyses. Furthermore, the widespread use of Markov models in medical decision making could potentially imply better understanding and acceptance of the analytic results as compared to other approaches.

Despite the potential limitations, we believe that the advantages of a systematic and explicit approach to economic evaluations of screening programs as proposed in this thesis may contribute to better-informed decision-making. There are currently many screening programs undertaken in routine clinical practice, and we suspect that the way they have been implemented (both in terms of age of initiation and screening intervals) may often be a result of rather arbitrary sequential decision-making. If anything, subjecting some of the more established screening programs to the analytic methods applied in this work would most likely provide some interesting food for thought.

6. CONCLUSIONS

- This work provided a simple and systematic method for the economic evaluation of multiple screening designs to identify an optimal design.
- In the application of screening for AF, it was demonstrated that the starting age of the screening program, as well as the number of potential repeated screening occasions, depend on the threshold for cost-effectiveness and the age of the investigated cohort.
- Methods to make the number of possible screening designs more manageable are available but need to be further studied.
- Changes in the screening setting are important as they affect the cost-effectiveness of the screening. This requires flexible modelling with continuously updated models.

Case-specific conclusions:

- All strategies that included screening for CRC were cost-effective compared to no screening. Screening with OC once implied the lowest cost for gaining QALYs while screening every tenth year gained additional QALYs to a reasonable cost. However, other factors such as patient acceptability of the test and availability of human resources also must be considered.
- Our model-based analysis showed that the screening for unknown AF in 75/76-year-old individuals, performed in the STROKESTOP-study, is cost-effective.
- By examining the results of more than two billion simulated population screening program designs for unknown AF, seven designs were deemed cost-effective depending on how much we are prepared to pay for gaining QALYs. Our results showed that repeated screening implied additional health benefits to a reasonable cost compared to one-off screening.

7. REFERENCES

- 1.** Maxim, L.D., R. Niebo, and M.J. Utell, *Screening tests: a review with examples*. Inhal Toxicol, 2014. 26(13): p. 811-28.
- 2.** Ilic, D., et al., *Screening for prostate cancer: an updated Cochrane systematic review*. BJU Int, 2011. 107(6): p. 882-91.
- 3.** Gotzsche, P.C. and M. Nielsen, *Screening for breast cancer with mammography*. Cochrane Database Syst Rev, 2011(1): p. CD001877.
- 4.** Socialstyrelsen (The National Board of Health and Welfare), *Screening programs*. <http://www.socialstyrelsen.se/riktlinjer/nationellascreeningprogram>, Accessed August 21, 2017.
- 5.** Svennberg, E., et al., *Mass Screening for Untreated Atrial Fibrillation: The STROKESTOP Study*. Circulation, 2015. 131(25): p. 2176-84.
- 6.** Wong, W.W., et al., *Cost-effectiveness of screening for hepatitis C in Canada*. CMAJ, 2015.
- 7.** Aronsson, M., et al., *Cost-effectiveness of mass screening for untreated atrial fibrillation using intermittent ECG recording*. Europace, 2015. 17(7): p. 1023-9.
- 8.** Lansdorp-Vogelaar, I., A.B. Knudsen, and H. Brenner, *Cost-effectiveness of colorectal cancer screening - an overview*. Best Pract Res Clin Gastroenterol, 2010. 24(4): p. 439-49.
- 9.** Fenwick, E., K. Claxton, and M. Sculpher, *Representing uncertainty: the role of cost-effectiveness acceptability curves*. Health Econ, 2001. 10(8): p. 779-87.
- 10.** Zelen, M., *Optimal Scheduling of Examinations for the Early Detection of Disease*. Biometrika, 1993. 80(2): p. 279-293.
- 11.** Kirkizlar, E., et al., *Timing of testing and treatment for asymptomatic diseases*. Math Biosci, 2010. 226(1): p. 28-37.
- 12.** Wilschut, J.A., et al., *How much colonoscopy screening should be recommended to individuals with various degrees of family history of colorectal cancer?* Cancer, 2011. 117(18): p. 4166-74.
- 13.** Zhang, J., et al., *Optimization of PSA screening policies: a comparison of the patient and societal perspectives*. Med Decis Making, 2012. 32(2): p. 337-49.
- 14.** Kahn, R., et al., *Age at initiation and frequency of screening to detect type 2 diabetes: a cost-effectiveness analysis*. Lancet, 2010. 375(9723): p. 1365-74.

- 15.** Altman, D.G. and J.M. Bland, *Statistics Notes: Diagnostic tests 1: sensitivity and specificity*. BMJ, 1994. 308(6943): p. 1552.
- 16.** Bharti, B. and S. Bharti, *Receiver-operating characteristic curve analysis in diagnostic, prognostic and predictive biomarker research: trade-off between sensitivity and specificity with change of test cut-offs*. J Clin Pathol, 2009. 62(11): p. 1051.
- 17.** Wilson, J.M. and Y.G. Jungner, *Principles and practice of mass screening for disease*. Bol Oficina Sanit Panam, 1968. 65(4): p. 281-393.
- 18.** Andermann, A., et al., *Revisiting Wilson and Jungner in the genomic age: a review of screening criteria over the past 40 years*. Bull World Health Organ, 2008. 86(4): p. 317-9.
- 19.** Tabar, L., et al., *The Swedish Two-County Trial twenty years later. Updated mortality results and new insights from long-term follow-up*. Radiol Clin North Am, 2000. 38(4): p. 625-51.
- 20.** Feig, S., *Cost-effectiveness of mammography, MRI, and ultrasonography for breast cancer screening*. Radiol Clin North Am, 2010. 48(5): p. 879-91.
- 21.** Hager, J., et al., *Revisiting the cost-effectiveness of screening 65-year-old men for abdominal aortic aneurysm based on data from an implemented screening programme*. Int Angiol, 2016.
- 22.** Ferlay, J., et al., *Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012*. Int J Cancer, 2015. 136(5): p. E359-86.
- 23.** U. S. Preventive Services Task Force., *Screening for colorectal cancer: recommendation and rationale*. Am Fam Physician, 2002. 66(12): p. 2287-90.
- 24.** U. S. Preventive Services Task Force., *Screening for colorectal cancer: U.S. Preventive Services Task Force recommendation statement*. Ann Intern Med, 2008. 149(9): p. 627-37.
- 25.** Winawer, S.J., et al., *Guidelines for colonoscopy surveillance after polypectomy: a consensus update by the US Multi-Society Task Force on Colorectal Cancer and the American Cancer Society*. CA Cancer J Clin, 2006. 56(3): p. 143-59; quiz 184-5.
- 26.** Winawer, S., et al., *Colorectal cancer screening and surveillance: clinical guidelines and rationale-Update based on new evidence*. Gastroenterology, 2003. 124(2): p. 544-60.
- 27.** Smith, R.A., V. Cokkinides, and H.J. Eyre, *American Cancer Society guidelines for the early detection of cancer, 2006*. CA Cancer J Clin, 2006. 56(1): p. 11-25; quiz 49-50.
- 28.** Jackman, R.J. and C.W. Mayo, *The adenoma-carcinoma sequence in cancer of the colon*. Surg Gynecol Obstet, 1951. 93(3): p. 327-30.

- 29.** Muto, T., H.J. Bussey, and B.C. Morson, *The evolution of cancer of the colon and rectum*. Cancer, 1975. 36(6): p. 2251-70.
- 30.** Schreuders, E.H., et al., *Colorectal cancer screening: a global overview of existing programmes*. Gut, 2015. 64(10): p. 1637-49.
- 31.** Hardcastle, J.D., et al., *Randomised controlled trial of faecal-occult-blood screening for colorectal cancer*. Lancet, 1996. 348(9040): p. 1472-7.
- 32.** Kronborg, O., et al., *Randomised study of screening for colorectal cancer with faecal-occult-blood test*. Lancet, 1996. 348(9040): p. 1467-71.
- 33.** Schoen, R.E., et al., *Colorectal-cancer incidence and mortality with screening flexible sigmoidoscopy*. N Engl J Med, 2012. 366(25): p. 2345-57.
- 34.** Lansdorp-Vogelaar, I., A.B. Knudsen, and H. Brenner, *Cost-effectiveness of colorectal cancer screening*. Epidemiol Rev, 2011. 33: p. 88-100.
- 35.** Sekiguchi, M., et al., *Optimal use of colonoscopy and fecal immunochemical test for population-based colorectal cancer screening: a cost-effectiveness analysis using Japanese data*. Jpn J Clin Oncol, 2015.
- 36.** Wilschut, J.A., et al., *Cost-effectiveness analysis of a quantitative immunochemical test for colorectal cancer screening*. Gastroenterology, 2011. 141(5): p. 1648-55 e1.
- 37.** de Wijkerslooth, T.R., et al., *Immunochemical fecal occult blood testing is equally sensitive for proximal and distal advanced neoplasia*. Am J Gastroenterol, 2012. 107(10): p. 1570-8.
- 38.** Brenner, H., C. Stock, and M. Hoffmeister, *Effect of screening sigmoidoscopy and screening colonoscopy on colorectal cancer incidence and mortality: systematic review and meta-analysis of randomised controlled trials and observational studies*. BMJ, 2014. 348: p. g2467.
- 39.** Friberg, L. and L. Bergfeldt, *Atrial fibrillation prevalence revisited*. J Intern Med, 2013. 274(5): p. 461-8.
- 40.** Heeringa, J., et al., *Prevalence, incidence and lifetime risk of atrial fibrillation: the Rotterdam study*. Eur Heart J, 2006. 27(8): p. 949-53.
- 41.** Stefansdottir, H., et al., *Trends in the incidence and prevalence of atrial fibrillation in Iceland and future projections*. Europace, 2011. 13(8): p. 1110-7.
- 42.** Wolf, P.A., R.D. Abbott, and W.B. Kannel, *Atrial fibrillation as an independent risk factor for stroke: the Framingham Study*. Stroke, 1991. 22(8): p. 983-8.
- 43.** Lee, H.Y., et al., *Quality-adjusted life expectancy (QALE) and loss of QALE for patients with ischemic stroke and intracerebral hemorrhage: a 13-year follow-up*. Stroke, 2010. 41(4): p. 739-44.

44. Thrall, G., et al., *Quality of life in patients with atrial fibrillation: a systematic review*. Am J Med, 2006. 119(5): p. 448 e1-19.
45. Henriksson, K.M., et al., *Survival after stroke--the impact of CHADS2 score and atrial fibrillation*. Int J Cardiol, 2010. 141(1): p. 18-23.
46. Friberg, L., et al., *Increased mortality in paroxysmal atrial fibrillation: report from the Stockholm Cohort-Study of Atrial Fibrillation (SCAF)*. Eur Heart J, 2007. 28(19): p. 2346-53.
47. Ghatnekar, O., et al., *Cost of stroke in Sweden: an incidence estimate*. Int J Technol Assess Health Care, 2004. 20(3): p. 375-80.
48. Hart, R.G., L.A. Pearce, and M.I. Aguilar, *Meta-analysis: antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation*. Ann Intern Med, 2007. 146(12): p. 857-67.
49. Camm, A.J., et al., *2012 focused update of the ESC Guidelines for the management of atrial fibrillation: an update of the 2010 ESC Guidelines for the management of atrial fibrillation. Developed with the special contribution of the European Heart Rhythm Association*. Eur Heart J, 2012. 33(21): p. 2719-47.
50. Fuster, V., et al., *2011 ACCF/AHA/HRS focused updates incorporated into the ACC/AHA/ESC 2006 guidelines for the management of patients with atrial fibrillation: a report of the American College of Cardiology Foundation/American Heart Association Task Force on practice guidelines*. Circulation, 2011. 123(10): p. e269-367.
51. Roche, F., et al., *Frequent and prolonged asymptomatic episodes of paroxysmal atrial fibrillation revealed by automatic long-term event recorders in patients with a negative 24-hour Holter*. Pacing Clin Electrophysiol, 2002. 25(11): p. 1587-93.
52. Doliwa Sobocinski, P., et al., *Improved screening for silent atrial fibrillation after ischaemic stroke*. Europace, 2012. 14(8): p. 1112-6.
53. Camm, A.J., G. Corbucci, and L. Padeletti, *Usefulness of continuous electrocardiographic monitoring for atrial fibrillation*. Am J Cardiol, 2012. 110(2): p. 270-6.
54. Hendrikx, T., et al., *Screening for atrial fibrillation with baseline and intermittent ECG recording in an out-of-hospital population*. BMC Cardiovasc Disord, 2013. 13(1): p. 41.
55. Hobbs, F.D., et al., *A randomised controlled trial and cost-effectiveness study of systematic screening (targeted and total population screening) versus routine practice for the detection of atrial fibrillation in people aged 65 and over. The SAFE study*. Health Technol Assess, 2005. 9(40): p. iii-iv, ix-x, 1-74.
56. Camm, A.J., et al., *2012 focused update of the ESC Guidelines for the management of atrial fibrillation: an update of the 2010 ESC Guidelines*

for the management of atrial fibrillation--developed with the special contribution of the European Heart Rhythm Association. Europace, 2012. 14(10): p. 1385-413.

57. Fitzmaurice, D.A., et al., *Screening versus routine practice in detection of atrial fibrillation in patients aged 65 or over: cluster randomised controlled trial. BMJ, 2007. 335(7616): p. 383.*

58. Wheeldon, N.M., et al., *Screening for atrial fibrillation in primary care. Heart, 1998. 79(1): p. 50-5.*

59. Doliwa, P.S., V. Frykman, and M. Rosenqvist, *Short-term ECG for out of hospital detection of silent atrial fibrillation episodes. Scand Cardiovasc J, 2009. 43(3): p. 163-8.*

60. Doliwa, P.S., M. Rosenqvist, and V. Frykman, *Paroxysmal atrial fibrillation with silent episodes: intermittent versus continuous monitoring. Scand Cardiovasc J, 2012. 46(3): p. 144-8.*

61. Doliwa Sobocinski, P., et al., *Improved screening for silent atrial fibrillation after ischaemic stroke. Europace, 2012. 14(8): p. 1112-1116.*

62. European Heart Rhythm Association, et al., *Guidelines for the management of atrial fibrillation: the Task Force for the Management of Atrial Fibrillation of the European Society of Cardiology (ESC). Eur Heart J, 2010. 31(19): p. 2369-429.*

63. Mozaffarian, D., et al., *Heart disease and stroke statistics-2015 update: a report from the american heart association. Circulation, 2015. 131(4): p. e29-e322.*

64. World health organization (WHO), *The WHO stroke surveillance. World Health Organization. 29-7-2004. 2004.*

65. Mukherjee, D. and C.G. Patil, *Epidemiology and the global burden of stroke. World Neurosurg, 2011. 76(6 Suppl): p. S85-90.*

66. Dewilde, S., *Quality of life decrements after stroke. Abstract., in ISPOR. 2014.*

67. van Exel, N.J., et al., *Burden of informal caregiving for stroke patients. Identification of caregivers at risk of adverse health effects. Cerebrovasc Dis, 2005. 19(1): p. 11-7.*

68. Dawson, J., et al., *Association between disability measures and healthcare costs after initial treatment for acute stroke. Stroke, 2007. 38(6): p. 1893-8.*

69. Emberson, J., et al., *Effect of treatment delay, age, and stroke severity on the effects of intravenous thrombolysis with alteplase for acute ischaemic stroke: a meta-analysis of individual patient data from randomised trials. Lancet, 2014. 384(9958): p. 1929-35.*

70. Jauch, E.C., et al., *Guidelines for the early management of patients with acute ischemic stroke: a guideline for healthcare professionals from*

the American Heart Association/American Stroke Association. *Stroke*, 2013. 44(3): p. 870-947.

71. Berkhemer, O.A., et al., *A randomized trial of intraarterial treatment for acute ischemic stroke*. *N Engl J Med*, 2015. 372(1): p. 11-20.

72. Broderick, J.P., et al., *Endovascular therapy after intravenous t-PA versus t-PA alone for stroke*. *N Engl J Med*, 2013. 368(10): p. 893-903.

73. Kidwell, C.S., et al., *A trial of imaging selection and endovascular treatment for ischemic stroke*. *N Engl J Med*, 2013. 368(10): p. 914-23.

74. Ciccone, A., et al., *Endovascular treatment for acute ischemic stroke*. *N Engl J Med*, 2013. 368(10): p. 904-13.

75. Goyal, M., et al., *Randomized assessment of rapid endovascular treatment of ischemic stroke*. *N Engl J Med*, 2015. 372(11): p. 1019-30.

76. Campbell, B.C., et al., *Endovascular therapy for ischemic stroke with perfusion-imaging selection*. *N Engl J Med*, 2015. 372(11): p. 1009-18.

77. Jovin, T.G., et al., *Thrombectomy within 8 hours after symptom onset in ischemic stroke*. *N Engl J Med*, 2015. 372(24): p. 2296-306.

78. Saver, J.L., et al., *Stent-retriever thrombectomy after intravenous t-PA vs. t-PA alone in stroke*. *N Engl J Med*, 2015. 372(24): p. 2285-95.

79. Drummond, M.F., et al., *Methods for the economic evaluation of health care programmes*. 3rd ed. Oxford medical publications. 2005, Oxford ; New York: Oxford University Press. xv, 379 p.

80. Briggs A, S.M., Claxton K., *Decision modeling for health economic evaluation*. 2006: Oxford press.

81. Lovejoy, W.S., *Computationally Feasible Bounds for Partially Observed Markov Decision-Processes*. *Operations Research*, 1991. 39(1): p. 162-175.

82. Karlsson, G. and M. Johannesson, *The decision rules of cost-effectiveness analysis*. *Pharmacoeconomics*, 1996. 9(2): p. 113-20.

83. Johannesson, M. and M.C. Weinstein, *On the decision rules of cost-effectiveness analysis*. *J Health Econ*, 1993. 12(4): p. 459-67.

84. Weinstein, M.C., *Principles of cost-effective resource allocation in health care organizations*. *Int J Technol Assess Health Care*, 1990. 6(1): p. 93-103.

85. Belgian Health Care Knowledge Centre (KCE), *Threshold values for cost-effectiveness in health care. KCE reports 100 C*. 2008.

86. National Institute for Health and Clinical Excellence (NICE), *Guide to the Methods of Technology Appraisal*. 2004.

87. Socialstyrelsen (The National Board of Health and Welfare). *Nationella riktlinjer för diabetesvården. National guidelines for diabetes care*. 2010.

- 88.** Gasmelseid, T.M., *Pharmacoinformatics and drug discovery technologies : theories and applications*. 2012, Hershey, Pa.: Medical Information Science Reference. xxii, 418 p.
- 89.** Claxton, K., et al., *Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold*. Health Technology Assessment, 2015. 19(14): p. 1-+.
- 90.** Svensson, M., F.O. Nilsson, and K. Arnberg, *Reimbursement Decisions for Pharmaceuticals in Sweden: The Impact of Disease Severity and Cost Effectiveness*. Pharmacoeconomics, 2015. 33(11): p. 1229-36.
- 91.** The Organisation for Economic Co-operation and Development (OECD). *Health at a Glance: Europe.*; Available from: <http://www.oecd-ilibrary.org/docserver/download/8114211e.pdf?expires=1479303854&id=id&accname=guest&checksum=606ED6717BDE57AE3D49A243286DA7BC>. Accessed August 21, 2017.
- 92.** Gopalappa, C., et al., *Probability model for estimating colorectal polyp progression rates*. Health Care Manag Sci, 2011. 14(1): p. 1-21.
- 93.** Regionalt cancercentrum., *Swedish colorectal cancer registry 2014. Svenska Kolorektalcancerregistret 2014*.
- 94.** Luning, T.H., et al., *Colonoscopic perforations: a review of 30,366 patients*. Surg Endosc, 2007. 21(6): p. 994-7.
- 95.** Statistics Sweden (SCB). *Life Tables 2014*. <http://www.scb.se/>. Accessed July 1, 2016.
- 96.** Lo, S.H., et al., *Colorectal cancer screening uptake over three biennial invitation rounds in the English bowel cancer screening programme*. Gut, 2015. 64(2): p. 282-91.
- 97.** Kapidzic, A., et al., *Attendance and diagnostic yield of repeated two-sample faecal immunochemical test screening for colorectal cancer*. Gut, 2015.
- 98.** Telford, J.J., et al., *The cost-effectiveness of screening for colorectal cancer*. CMAJ, 2010. 182(12): p. 1307-13.
- 99.** Burstrom, K., M. Johannesson, and F. Diderichsen, *Increasing socioeconomic inequalities in life expectancy and QALYs in Sweden 1980-1997*. Health Econ, 2005. 14(8): p. 831-50.
- 100.** Ness, R.M., et al., *Utility valuations for outcome states of colorectal cancer*. Am J Gastroenterol, 1999. 94(6): p. 1650-7.
- 101.** Jonas, D.E., et al., *Willingness-to-pay to avoid the time spent and discomfort associated with screening colonoscopy*. Health Econ, 2010. 19(10): p. 1193-211.
- 102.** Tilson, L., et al., *Cost of care for colorectal cancer in Ireland: a health care payer perspective*. Eur J Health Econ, 2012. 13(4): p. 511-24.

- 103.** Pharmaceutical Specialties in Sweden (FASS). <http://www.fass.se/>. Accessed July 1, 2016. 2013.
- 104.** Socialstyrelsen (The National Board of Health and Welfare). NordDRG. 2015; Available from: <http://www.socialstyrelsen.se/klassificeringochkoder/norddrg>. Accessed August 21, 2017.
- 105.** Tandvårds- och läkemedelförmånsverket (TLV). *Tandvårds- och läkemedelförmånsverkets allmänna råd; Allmänna råd för ansökan om pris och subvention för förbrukningsartiklar*. TLVAR, 2011. 1.
- 106.** Mandel, J.S., et al., *Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Control Study*. N Engl J Med, 1993. 328(19): p. 1365-71.
- 107.** Tappenden, P., et al., *Option appraisal of population-based colorectal cancer screening programmes in England*. Gut, 2007. 56(5): p. 677-84.
- 108.** Goede, S.L., et al., *Cost-effectiveness of one versus two sample faecal immunochemical testing for colorectal cancer screening*. Gut, 2013. 62(5): p. 727-34.
- 109.** Ko, C.W. and J.A. Dominitz, *Complications of colonoscopy: magnitude and management*. Gastrointest Endosc Clin N Am, 2010. 20(4): p. 659-71.
- 110.** Warren, J.L., et al., *Adverse events after outpatient colonoscopy in the Medicare population*. Ann Intern Med, 2009. 150(12): p. 849-57, W152.
- 111.** Lee, D., et al., *Cost effectiveness of CT colonography for UK NHS colorectal cancer screening of asymptomatic adults aged 60-69 years*. Appl Health Econ Health Policy, 2010. 8(3): p. 141-54.
- 112.** Heitman, S.J., et al., *Prevalence of adenomas and colorectal cancer in average risk individuals: a systematic review and meta-analysis*. Clin Gastroenterol Hepatol, 2009. 7(12): p. 1272-8.
- 113.** Tappenden, P., et al., *Colorectal cancer screening options appraisal. Cost-effectiveness, cost-utility and resource impact of alternative screening options for colorectal cancer. Report to the English Bowel Cancer Screening Working Group*. 2004, University of Sheffield.
- 114.** Winawer, S.J., et al., *Colorectal cancer screening: clinical guidelines and rationale*. Gastroenterology, 1997. 112(2): p. 594-642.
- 115.** Parekh, M., A.M. Fendrick, and U. Ladabaum, *As tests evolve and costs of cancer care rise: reappraising stool-based screening for colorectal neoplasia*. Aliment Pharmacol Ther, 2008. 27(8): p. 697-712.
- 116.** Friberg, L., et al., *Population screening of 75- and 76-year-old men and women for silent atrial fibrillation (STROKESTOP)*. Europace, 2013. 15(1): p. 135-40.

- 117.** Friberg, L., M. Rosenqvist, and G.Y. Lip, *Evaluation of risk stratification schemes for ischaemic stroke and bleeding in 182 678 patients with atrial fibrillation: the Swedish Atrial Fibrillation cohort study*. Eur Heart J, 2012. 33(12): p. 1500-10.
- 118.** Granger, C.B., et al., *Apixaban versus warfarin in patients with atrial fibrillation*. N Engl J Med, 2011. 365(11): p. 981-92.
- 119.** European Heart Rhythm, A., et al., *Guidelines for the management of atrial fibrillation: the Task Force for the Management of Atrial Fibrillation of the European Society of Cardiology (ESC)*. Europace, 2010. 12(10): p. 1360-420.
- 120.** Skeppholm, M. and L. Friberg, *Adherence to warfarin treatment among patients with atrial fibrillation*. Clin Res Cardiol, 2014. 103(12): p. 998-1005.
- 121.** Socialstyrelsen (The National Board of Health and Welfare), *Nationella riktlinjer för hjärtsjukvård. National cardiovascular guidelines*. 2013.
- 122.** Connolly, S.J., et al., *Dabigatran versus warfarin in patients with atrial fibrillation*. N Engl J Med, 2009. 361(12): p. 1139-51.
- 123.** Patel, M.R., et al., *Rivaroxaban versus warfarin in nonvalvular atrial fibrillation*. N Engl J Med, 2011. 365(10): p. 883-91.
- 124.** Healey, J.S., et al., *Subclinical atrial fibrillation and the risk of stroke*. N Engl J Med, 2012. 366(2): p. 120-9.
- 125.** Marfella, R., et al., *Brief episodes of silent atrial fibrillation predict clinical vascular brain disease in type 2 diabetic patients*. J Am Coll Cardiol, 2013. 62(6): p. 525-30.
- 126.** Wallentin, L., et al., *Efficacy and safety of apixaban compared with warfarin at different levels of predicted international normalized ratio control for stroke prevention in atrial fibrillation*. Circulation, 2013. 127(22): p. 2166-76.
- 127.** Lopes, R.D., et al., *Efficacy and safety of apixaban compared with warfarin according to patient risk of stroke and of bleeding in atrial fibrillation: a secondary analysis of a randomised controlled trial*. Lancet, 2012. 380(9855): p. 1749-58.
- 128.** Levin, L.A., et al., *A cost-effectiveness analysis of screening for silent atrial fibrillation after ischaemic stroke*. Europace, 2015. 17(2): p. 207-14.
- 129.** Ghatnekar, O., *The burden of stroke in Sweden: studies on costs and quality of life based on Riks-Stroke, the Swedish stroke register.*, in *Department of public health and clinical medicine*. 2013, Umeå University.
- 130.** Burstrom, K., M. Johannesson, and F. Diderichsen, *A comparison of individual and social time trade-off values for health states in the general population*. Health Policy, 2006. 76(3): p. 359-70.

- 131.** Dolan, P., *Modeling valuations for EuroQol health states*. Med Care, 1997. 35(11): p. 1095-108.
- 132.** Levin LA, et al., *The cost-effectiveness of screening for silent atrial fibrillation after ischaemic stroke*, in ISPOR 16th Annual European Congress. 2013: Dublin.
- 133.** O'Brien, C.L. and B.F. Gage, *Costs and effectiveness of ximelagatran for stroke prophylaxis in chronic atrial fibrillation*. JAMA, 2005. 293(6): p. 699-706.
- 134.** Pennlert, J., et al., *Long-term risk and predictors of recurrent stroke beyond the acute phase*. Stroke, 2014. 45(6): p. 1839-41.
- 135.** Fang, M.C., et al., *Long-term survival after ischemic stroke in patients with atrial fibrillation*. Neurology, 2014. 82(12): p. 1033-7.
- 136.** Samsa, G.P., et al., *Performing cost-effectiveness analysis by integrating randomized trial data with a comprehensive decision model: application to treatment of acute ischemic stroke*. J Clin Epidemiol, 1999. 52(3): p. 259-71.
- 137.** Earnshaw, S.R., et al., *Cost-effectiveness of patient selection using penumbral-based MRI for intravenous thrombolysis*. Stroke, 2009. 40(5): p. 1710-20.
- 138.** Tawil, S.E., et al., *How many stroke patients might be eligible for mechanical thrombectomy?* European Stroke Journal, 2016. 1(4): p. 264-271.
- 139.** Socialstyrelsen (The National Board of Health and Welfare), *Nationella strokeriktlinjer. National stroke guidelines*. . 2017.
- 140.** Lowres, N., et al., *Feasibility and cost-effectiveness of stroke prevention through community screening for atrial fibrillation using iPhone ECG in pharmacies. The SEARCH-AF study*. Thromb Haemost, 2014. 111(6): p. 1167-76.
- 141.** Maeda, K., T. Shimbo, and T. Fukui, *Cost-effectiveness of a community-based screening programme for chronic atrial fibrillation in Japan*. J Med Screen, 2004. 11(2): p. 97-102.
- 142.** The Health Information and Quality Authority (HIQA). *Health technology assessment (HTA) of a national screening programme for atrial fibrillation in primary care*. 2015.
- 143.** Sanna, T., et al., *Cryptogenic stroke and underlying atrial fibrillation*. N Engl J Med, 2014. 370(26): p. 2478-86.
- 144.** Barrett, P.M., et al., *Comparison of 24-hour Holter monitoring with 14-day novel adhesive patch electrocardiographic monitoring*. Am J Med, 2014. 127(1): p. 95 e11-7.
- 145.** Gladstone, D.J., et al., *Atrial fibrillation in patients with cryptogenic stroke*. N Engl J Med, 2014. 370(26): p. 2467-77.

- 146.** Vaes, B., et al., *The diagnostic accuracy of the MyDiagnostick to detect atrial fibrillation in primary care*. BMC Fam Pract, 2014. 15: p. 113.
- 147.** Svennberg, C.E., et al., *Mass screening for silent atrial fibrillation in high risk patients - preliminary results from the STROKESTOP trial*. 2013. 34(suppl 1): p. 809.
- 148.** Engdahl, J., et al., *Uptake of atrial fibrillation screening aiming at stroke prevention: geo-mapping of target population and non-participation*. BMC Public Health, 2013. 13: p. 715.
- 149.** Bolin, K., et al., *The cost-effectiveness of growth hormone replacement therapy (Genotropin(R)) in hypopituitary adults in Sweden*. Cost Eff Resour Alloc, 2013. 11(1): p. 24.
- 150.** Bengtsson, T., *Population ageing - a threat to the welfare state? : the case of Sweden*. Demographic research monographs. 2010: Berlin ; Heidelberg : Springer, c2010.
- 151.** National Institute for Health and Clinical Excellence (NICE), *Guide to the methods of technology appraisal*. London: . 2013: London.
- 152.** Lanitis, T., et al., *Cost-effectiveness of apixaban versus warfarin and aspirin in Sweden for stroke prevention in patients with atrial fibrillation*. Thromb Res, 2014. 134(2): p. 278-87.
- 153.** Strong, M., A. Brennan, and J. Oakley, *How to Calculate Value of Information in Seconds Using 'Savi', the Sheffield Accelerated Value of Information Web App*. Value Health, 2015. 18(7): p. A725-6.
- 154.** Alagoz, O., et al., *Determining the acceptance of cadaveric livers using an implicit model of the waiting list*. Operations Research, 2007. 55(1): p. 24-36.
- 155.** Hauskrecht, M., *Value-function approximations for partially observable Markov decision processes*. Journal of Artificial Intelligence Research, 2000. 13: p. 33-94.
- 156.** Maillart, L.M., et al., *Assessing Dynamic Breast Cancer Screening Policies*. Operations Research, 2008. 56(6): p. 1411-1427.
- 157.** Hoey, J., et al., *Automated handwashing assistance for persons with dementia using video and a partially observable Markov decision process*. Computer Vision and Image Understanding, 2010. 114(5): p. 503-519.

Papers

The papers associated with this thesis have been removed for copyright reasons. For more details about these see:

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-141556>