



M-IEM | HCTM

MASTER THESIS

Breast Cancer Screening in the Netherlands: a Health Economic Evaluation of Screening Strategies

G.J.W. VAN WEEZEL | S1802763

1 MAY 2023

UNIVERSITY OF TWENTE

DR. IR. A.A.M. SPIL

PROF. DR. IR. E.W. HANS

ACCENTURE THE NETHERLANDS

V. VAN DE BURGT, MSC

**accenture**

**UNIVERSITY
OF TWENTE.**

COLOPHON

MANAGEMENT

Behavioural, Management and Social Sciences
Industrial Engineering and Management

DATE

May 1, 2023

REFERENCE

Master Thesis IE&M

VERSION

Final

STATUS

Final

PROJECT

Master Thesis IE&M

PROJECT NUMBERS

1914100060

AUTHOR(S)

G.J.W. van Weezel

TELEPHONE

+31 6 1049 5861

EMAIL

gielvanweezel@gmail.com

POSTAL ADDRESS

P.O. Box 217
7500 AE Enschede

WEBSITE

www.utwente.nl

FILENAME

Thesis Final 2023_05_01.docx

COPYRIGHT

© University of Twente, The Netherlands, 2023

© Accenture B.V., The Netherlands, 2023

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, be it electronic, mechanical, by photocopies, or recordings in any other way, without the prior written permission of the University of Twente or Accenture B.V..

Management Summary

Introduction

This study assesses the cost-effectiveness and the budget impact of different screening strategies for the breast cancer screening program in the Netherlands through a health economic analysis. Currently, all women between 50 and 75 are screened bi-annually to increase early detection and treatment and thus improve the survivability of breast cancer. It is currently unclear from the literature what the health-economic impact of this screening program is for society and women of different age ranges. The goal of this research is the following:

Assess the health-economic performance of breast cancer screening programs in the Netherlands by comparing populations that undergo different screening strategies through a newly developed population screening simulation tool.

The analysis is performed through discrete event simulation. This study provides a new tool with which screening strategies for breast cancer and other diseases can be tested and their health-economic performance analyzed. The health economic analysis of these interventions is important on a societal scale, as limited public health care resources should be spent optimally. This study has developed a new tool which Accenture, the company where the research takes place, can use to help global healthcare providers with cost-utility analyses.

Problem description

Public healthcare costs and healthcare personnel deficiencies are rising. Limited funds and personnel should therefore aim to improve public health optimally. Health economic analyses provide a way to gain insight into the cost-effectiveness of specific interventions to achieve optimal healthcare expenditure. The literature does currently not provide a consensus on the cost-effectiveness of the Dutch breast cancer screening program. The last Dutch evaluation of the breast cancer screening program took place in 2014. This study concluded that breast cancer mortality was reduced by 16% due to screening, but they did recommend additional research for cost-effectiveness due to changing circumstances in treatment and detection. Since then, new guidelines for health economic analyses in the Dutch context have been published. These included more cost parameters, like a societal burden from missed work and extra medical costs for longer survival. The guidelines also include more ways to calculate patient utility to get a better overview of a patient's quality of life during procedures.

Approach

A literature study was performed on the background of breast cancer, population screening programs, and health economic analyses. A systematic review of breast cancer screening program evaluations made it clear that research on the subject in the Netherlands was outdated and not up to current analysis standards. Studies that were analyzed in the systematic review showed uncorrelated outcomes, with costs per additional quality-adjusted life year (QALY) ranging between €5,000 and €54,000 per QALY. This caused the research focus to shift towards creating a new health-economic analysis of the Dutch breast cancer screening program.

A discrete event model was created that simulates women from birth to death to assess the cost-effectiveness of the breast screening program and various screening strategies. Women are individually

simulated with unique parameters like healthy life expectancy, the probability of getting cancer at various points in their lives, and tumor growth rates. The model allows for different screening strategies: the normal range of bi-annual screening between 50 and 75, no screening at all, or screening of specific age intervals. The model was validated to reflect Dutch figures on mortality, cancer incidence, and diagnoses through screening. Various submodules have been created to replicate tumor growth, screening and clinical diagnosis, and cancer staging based on primary tumor size to achieve this validated model.

All models were designed based on data gathered from the literature and reliable Dutch government sources. A new tumor growth model was constructed that follows the natural progression of a tumor, including the possibility for growth stagnation and spontaneous regression. The tumor detection model's screening and clinical sensitivity replicated actual findings of tumor sizes according to data of the Dutch Integral Cancer Center (IKNL). The staging model also reflects the relation between primary tumor size and cancer stage, as found in the literature. The patient-level model, which contains the patient's life trajectory, was modeled to include cancer screening, diagnosis, and treatment guidelines in line with the Dutch guidelines. The theoretical model framework, computerized model implementation, and validation in line with validation models make the model a reliable representation of the Dutch breast cancer screening program and its effects on population health. The model's outcome under the currently in-place screening strategy can replicate actual cancer data presented by the IKNL.

Conclusion

The status of the breast cancer screening program according to the literature is overwhelmingly positive, as determined by a systematic review performed for this research. The literature does advise further research and health economic analyses into the breast cancer screening program than has been done, and advise the inclusion of more costs and utility measures in further research.

The screening program provides benefits for individual women. The probability of being diagnosed with severe cancer is reduced by more than half, and there is no significant change in average quality-adjusted life years experienced during their lives due to the possibility of false diagnoses and screening.

A tool has been created that allows comparison of screening strategies for breast cancer. The model has been built in compliance with the latest guidelines for health economic analyses. Compared to older models, this model considers more factors for utility and disutility, and includes costs related to medical, societal and diseases in won life years. The model is validated in accordance with the AdViSHE tool recommended by health economic evaluation guidelines. Cost, utility, treatment, survival data and various other input parameters for the model were based on the literature. This model can now serve as a base for further health economic evaluations on screening programs and could, with small modifications, be used for other diseases and proposed screening programs like bowel cancer, ovarian cancer, prostate cancer, and lung cancer.

The results show that there are no significant differences in average quality of life experienced between tested strategies, with the average QALYs of all strategies between 72.2 and 72.4. At the same time, all costs related to breast cancer per individual are higher as more screening takes place. Screening clearly reduces the average severity of breast cancer in diagnosed patients, but also lead to more diagnoses in general. Of the screening strategies compared, the current strategy provides a great tradeoff between total costs, overdiagnoses, quality of life, and reduction of late-stage cancers.

The budget impact analysis shows significant costs associated with the screening program. A new ethical discussion is recommended to decide what these costs and changes to the program would implicate. Comparing the differences in total costs to the current baseline of bi-annual screening between 50 and 75, leads to decreases of €387M, €100M, and €137M annually when moving to no screening, tri-annual screening, or reduced screening between 55 and 70, respectively. Increasing screening to annual screening or extending the age ranges to include all women between 40 and 80 would increase annual costs with €309M and €258M, respectively, compared to the baseline.

What's next

This study provides an extensive evaluation of the current breast cancer screening program in the Netherlands and provides a model that allows for future exploration of screening strategies, also for other types of cancer. Two main recommendations arose from the results of this thesis.

First, there is the recommendation for governmental bodies and cancer institutes to review the experiments and, where possible, replicate them with their input data. All data in this study was found in the literature, as primary sources were unavailable. Cancer institutes could redo the experiments using more recent and real data. This would provide an even more realistic view of the status of the screening program and the comparison between potential strategies. This study evaluated the effects of changing input data through a probabilistic sensitivity analysis. Effects of changing variables is expected to be limited in terms of life gain and quality adjusted life years, but changing treatment costs could significantly affect the outcomes.

Second, there is a need to start an ethical discussion about the screening program, healthcare costs in general, and the outcomes of this study. The cost calculations that have now been performed, result in significantly higher figures than previous studies. This, in turn, raises the question about the validity of current willingness-to-pay thresholds in future analyses. This research shows that clinical benefits of the screening program include a stage shift to lower stages but come at a high cost to society and more cancer diagnoses in general. The results of this study are a starting point for this discussion, where the value of early diagnoses and prevention of severe disease needs to be weighed against the costs and burden to the rest of society.

This research provides a new, validated model for health economic analyses for cancer screening programs, with this version focused on breast cancer. The results of the study on breast cancer screening strategies aim to provide a basis for an ethical discussion on how benefits and drawbacks should be valued. The current screening program in the Netherlands is one of the best options. Severe cancer incidence is significantly reduced due to the screening program, but that does come with an increase in overall diagnoses and at a high cost. There are two explicit recommendations: an ethical discussion should be started on the results of this study and the screening program, and cancer institutes and governing bodies should research this model further, where they should populate the model with their own governed data, which was unavailable for this research. This research led to many new insights into the breast cancer screening program, which could prove valuable for society and science.

Preface

This is the final thesis of my master's at the University of Twente. It is the culmination of the wonderful past 6.5 years I've enjoyed at the university. Over these years, I've grown fond of the university and explored many of its offerings in studies and extracurricular activities.

In 2016, I started my Bachelor's in Technical Medicine. Friends I've met through study and with whom I joined a student association are still dear to me, and now that almost all of us live in Utrecht, we still meet up almost weekly. My Bachelor's went well, and apart from one module I missed I managed to finish it in three years and a bit. However, this study wasn't what I had hoped, and to better fit my ambitions, I decided to go for a master's in Industrial Engineering and Management.

This master started fast. I interrupted my studies for a year on the board of the Student Union, which reinvigorated my love for the university. It fits right in with what I wanted to do with my life, and I gained invaluable experience this year. During this year, I acquired many new skills and interests, so after returning to my studies, I decided to spice up my master's by adding an extra major. I was now an IEM student majoring in Health Care Technology Management and Financial Engineering and Management.

My final year of classes combined 70EC with a chair position in the University Council. I was doing everything I wanted, and the world lay at my feet. I thoroughly enjoyed this challenge of studying and developing myself through my position at the UC. It also gave me the inspiration for my master's thesis. I wanted to do something that could better society and would combine things I enjoyed. I wanted to combine my passion for healthcare and finance in a project that could invoke discussion about public policy.

I managed to create an idea for a thesis, which eventually became the one in front of you right now. I was able to realize this at Accenture, where people were fond of this idea. The past six months there have been amazing. The great coworkers, fantastic working atmosphere, and being able to help others there with various projects gave me a good insight into what working for a company like Accenture would be like. During that time, I also managed to write this thesis, which had become quite the project in itself.

I wouldn't have been able to do this without the support of everyone around me. Of course, I want to thank my supervisors at the university, Ton Spil and Erwin Hans, for the interesting discussions and insights. They often came with questions from angles I didn't even want to look at yet, luckily this turned out for the better. I want to thank Verona van de Burgt for her support within Accenture, who even got me to join the annual H&PS ski trip. I want to thank my friends and roommates for their support and discussions over some aspects of my research, my brother Thijs for the technical discussions, and my sister Manon for the legal background. Above all, I want to thank my parents, Koen and Helma, and my girlfriend, Fleur for their undeniable support. They've been there for me many late nights for a discussion, advice, or a nice cup of tea when it was getting late.

Finally, I want to thank the University for the amazing 6.5 years I've been there. From the education in two different studies to the amazing positions on the SU board and University Council: it's been a blast!

Giel van Weezel,
10 April 2023

List of Tables

Table 1: BI-RADS scores and parameters (Ghaemian et al., 2021).....	20
Table 2: Main search terms and synonyms for search query.....	29
Table 3: General Characteristics of included papers.....	31
Table 4: Type of Health Economic Analysis performed per paper.....	31
Table 5: Costs measures included in calculation of outcome.....	32
Table 6: Utility measures used in calculation of outcome.....	32
Table 7: Tumor growth model used in simulation.....	33
Table 8: PICOT scope for this research.....	39
Table 9: Overview of costs to be included in a HEA from a societal perspective.....	40
Table 10: Initial patient attributes.....	49
Table 11: Base utility values per age used in the model (Wright et al., 2022).....	55
Table 12: Percentage at which the base utility is lowered per stage (Arrospide et al., 2016).....	55
Table 13: Costs for screening and diagnostics.....	56
Table 14: Treatment costs per stage and time point (van Luijt et al., 2017).....	56
Table 15: Example of utility discounting.....	57
Table 16: KPIs used for calibration of the model.....	58
Table 17: Comparing IKNL figures to results from the model calibration.....	59
Table 18: PSA input variables.....	63
Table 19: Experimental screening strategies.....	63
Table 20: 2K factorial experiment design.....	64
Table 21: Outcome KPIs of baseline simulation with 1M patients.....	65
Table 22: Confidence intervals of mean total QALY at various simulation sizes.....	67
Table 23: PSA Experiments and associated outcomes.....	68
Table 24: Results of different screening strategies.....	70
Table 25: Budget impact of various screening strategies compared to the baseline.....	76
Table 26: Overall survival and disease-specific survival from risk profile and stage.....	91
Table 27: Risk factors for breast cancer development.....	91

List of Figures

Figure 1: Stages of breast cancer development from ductal carcinoma	19
Figure 2: Costs per (QA)LY gained as per the Mühlberger et al. systematic review.....	26
Figure 3: Paper down-select according to PRISMA method	29
Figure 4: Costs per QALY and LYG outcomes of the different papers	33
Figure 5: Example of a Cost-Effectiveness plane showing when outcomes would be acceptable.....	42
Figure 6: Example of a Cost-effectiveness Acceptability Curve	42
Figure 7: Patient level model overview	44
Figure 8: Tumor Growth Model combining Gompertz Growth, Regression and Stagnation	45
Figure 9: Tumor growth quantiles from the Wisconsin breast cancer model.....	46
Figure 10: Tumor growth quantiles from the model used in this research.....	46
Figure 11: Tumor detection through screening and clinical diagnosis.	47
Figure 12: Relation between cancer stage and primary tumor size	48
Figure 13: Copula of cancer incidence age and life expectancy	51
Figure 14: First set of graphs from simulation output.....	59
Figure 15: Percentage of tumors discovered through the screening program	60
Figure 16: Age at diagnosis	60
Figure 17: Stage Distribution per modality.....	60
Figure 18: Relative survival per stage	60
Figure 19: ICER plane showing reproducibility of baseline KPIs in smaller simulations	66
Figure 20: Running mean of simulated populations final total QALY.....	67
Figure 21: Tornado graph showing the effect of the PSA on utility compared to the baseline	69
Figure 22: Tornado graph showing the effect of the PSA on costs compared to the baseline	69
Figure 23: ICER plane comparing cost and utility outcomes of the strategies to the baseline	72
Figure 24: Cost-effectiveness acceptability curve of the screening strategies compared to the baseline	73
Figure 25: Cancer stage distribution for the strategies. Incidence on y-axis and percentages in bars.	74
Figure 26: Cost-effectiveness plane showing 255 screening strategies compared to the baseline	75
Figure 27: OncoGuide's Breast Cancer care pathway Flowchart.....	89
Figure 28: Breast cancer incidence 1990-2021.....	90
Figure 29: Discovery through screening program per stage.....	90
Figure 30: Stage prevalence.....	90
Figure 31: Age at time of diagnosis.....	90

List of Formulas

Equation 1: Incremental Cost-Effectiveness Ratio.....	24
Equation 2: Discount rate	57

List of Abbreviations

Abbreviation	Definition
ACS	American Cancer Society
ADH	Abnormal Ductal Hyperplasia
AdViSHE	Assessment of the Validation Status of Health-Economic decision models
AJCC	American Joint Committee on Cancer
AUS	Australia
BIA	Budget Impact Analysis
BIRADS	Breast Imaging Reporting and Data System
BMI	Body Mass Index
BVO	Bevolkingsonderzoek – Population screening program
CBA	Cost Benefit Analysis
CBS	Centraal Bureau voor de Statistiek – Dutch bureau for statistics
CDC	Center for Disease Control
CEA	Cost-effectiveness Analysis
CEAC	Cost-Effectiveness Acceptability Curve
CE-plane	Cost-Effectiveness plane
CISNET	Cancer Intervention and Surveillance Modelling Network
CMA	Cost Minimization Analysis
COI	Cost of Illness Study
COVID-19	Coronavirus Disease 2019
CPU	Central Processing Unit
CUA	Cost Utility Analysis
DCIS	Ductal Carcinoma in Situ
DES	Discrete Event Simulation
EQ-5D-5L	European Quality of Life - 5 Dimensions - 5 Levels
ER	Estrogen Receptor
EUnetHTA	European Network for Health Technology Assessment
EU-TOPIA	European Union - Towards Improved Cancer Screening
GNP	Gross National Product
GP	General Practitioner
HEA	Health Economic Analysis
HER2	Human Epidermal Growth Factor receptor 2
ICER	Incremental Cost-Effectiveness Ratio
ICUR	Incremental Cost-Utility Ratio
IKNL	Integraal Kankercentrum Nederland – Dutch cancer center
ISPOR	The Professional Society for Health Economics and Outcomes Research, formerly International Society for Pharmacoeconomics and Outcomes Research
IT	Information Technology

KPI	Key Performance Indicator
LYG	Life Years Gained
MANC-RISK-SCREEN	Manchester Breast Cancer Model
MDACC	MD Anderson Cancer Center
MISCAN	Micro Simulation Screening Analysis
NA	Not Applicable
NL	Nederland - Netherlands
PAID	Practical Application to Include future Disease costs
PICOT	Patient, Intervention, Comparison, Outcome, Time
PR	Progesterone Receptor
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-Study	Personalized RiSk-based Mamma screening
PSA	Probabilistic Sensitivity Analysis
PYPL	Popularity of Programming Languages
QALY(s)	Quality-adjusted Life Year(s)
QoL	Quality of Life
R	Programming language for statistical computing and graphics
RCT	Randomized Control Trial
RIVM	Rijksinstituut voor Volksgezondheid en Milieu – Dutch Center for Disease Control
SEC	Social Economic Council
SiMRiSc	Simulation Model for breast cancer Risk
SMDM	Society for Medical Decision Making
TNM	Tumor, Nodes, Metastases
UK	United Kingdom
US	United States
USA	United States of America
UT	University of Twente
WBO	Wet op het Bevolkingsonderzoek – Law on population screening
WHO	World Health Organization
WTP	Willingness to Pay threshold

Table of Contents

Management Summary	3
Preface	6
List of Tables	7
List of Figures	8
List of Formulas	8
List of Abbreviations	9
Table of Contents	11
1. Introduction	13
1.1. Context	13
1.2. Research Motivation	13
1.3. Research Objective and Questions	16
2. Background	18
2.1. Breast Cancer	18
2.2. Population Screening Programs	21
2.3. Health Economic Evaluation	23
3. Systematic Review of Breast Cancer Screening Program Evaluations	28
3.1. Introduction	28
3.2. Methods	28
3.3. Results	30
3.4. Analysis of the results	34
3.5. Discussion	35
3.6. Conclusion	36
4. Methods	38
4.1. HEA Guidelines	38
4.2. Outcome Metrics & KPIs	41
4.3. Discrete Event Simulation Model	43
4.4. Model Calibration and Validation	58
4.5. Experiment Design	62
5. Results	65
5.1. Baseline Results	65
5.2. PSA Results	67
5.3. Experiment Results	70

5.4. Strategy Test Results.....	74
5.5. Budget Impact Analysis.....	75
6. Conclusion.....	77
7. Discussion.....	79
Bibliography	81
Appendix	89
Appendix I – Guidelines and Treatment	89
Appendix II – Breast Cancer	90
Appendix III - Screening Principles & legal grounds and Evaluations	93
Appendix IV - Screening Legal Grounds and Governmental Policies.....	95
Appendix V - Drawbacks and Bias	97
Appendix VI - Systematic Review of Breast Cancer Screening Program Evaluations	101
Appendix VII – Full Model Overview.....	110
Appendix VIII – Full code explanation.....	113
Appendix IX – AdViSHE questionnaire	162
Appendix X – 2K Factorial Experiment Design	169

1. Introduction

This thesis assesses the cost-effectiveness and the budget impact of different screening strategies for the breast cancer screening program in the Netherlands through a health economic analysis. This chapter will give context on the screening program and the need for a new evaluation. The motivation for this research will be given from three perspectives: a societal perspective, a scientific perspective, and the perspective of the company Accenture. The research problem, objective, and questions will be explored at the end of Chapter 1.

1.1. Context

The Dutch Healthcare system is under pressure. It suffers from staff shortages and increasing costs. The Social Economic Council concludes that around 1 in 7 people are currently working in healthcare in the Netherlands and expects this to grow to 1 in 4 by 2040 (Sociaal-Economische Raad, 2020). Furthermore, healthcare-related costs amounted to 12,7% of the Dutch GNP in 2015 and are expected to rise to 18% in 2060 (Medisch Contact, 2022). There is a need to evaluate the cost-effectiveness and budget impact of parts of the care chain, to make our care system future-proof.

Breast cancer is the most prevalent type of cancer in the Netherlands. Annually, more than 18,000 people are diagnosed with this type of cancer in the Netherlands. This means that 14.2% of all women will get this disease in their lifetimes, and 3.7% of all women will die from it. Breast cancer screening programs have been set up to diagnose the disease at an earlier stage. Breast screening is widely implemented in many healthcare systems to speed up the diagnosis of smaller, asymptomatic breast cancers, reduce cancer mortality, and increase the overall utility of the population. Mammography is used for most women starting in middle age, while screening recommendations and practices vary by area (Clift et al., 2022; Integraal Kankercentrum Nederland, 2022; van Haperen, 2018).

The breast cancer screening program was first implemented in 1989. Since then, better treatment options have become available for later stages, improving the disease's overall survivability. The procedure is often considered unpleasant, and there is a risk of overdiagnosis and overtreatment. Concerns about the program's effectiveness have become more apparent over recent years. (Winter & Velden, 2022)

The Dutch breast cancer screening program will be the subject of the health economic analysis performed in this research. The latest evaluation commissioned by the House of Representatives (Tweede Kamer) stems from 2014, and the recommendation was made to include more cost and utility measures in new analyses. This research will focus on performing a health economic analysis of the breast cancer screening program by following the Dutch guidelines to assess the budget impact and cost-effectiveness of the program. (Health Council of the Netherlands, 2014a)

1.2. Research Motivation

The importance of performing a health economic analysis for the breast cancer screening program can be explained through three relevant stakeholders. First is the citizen, payer, and possible patient, who should be able to make an informed decision on partaking in the screening program, and whose taxes should be used efficiently to improve public health. Second is the company, Accenture, whose clients that perform these screenings need to know whether they are effective. Accenture is incentivized to keep growing and innovating in healthcare intelligence, and this research will help with that. Finally, there is the scientific

relevance of the research. The knowledge of effective breast cancer treatment strategies must evolve to formulate the best strategies to treat the disease.

1.2.1. Societal Relevance

The Dutch Federation of Medical Professionals summarizes the status of the problems of the Dutch Healthcare sector in September 2022 as follows: "Our sector has been under great pressure for too long due to the shortage of healthcare staff, an aging population, increasing numbers of patients with more complex care needs, ever-increasing healthcare costs and the still ongoing impact of the corona pandemic." (Federatie Medisch Specialisten, 2022).

Dutch healthcare suffers from staff shortages and increasing costs. The Social Economic Council concludes that around 1 in 7 people are working in healthcare in the Netherlands. SEC expects this to grow to 1 in 4 by 2040 (Sociaal-Economische Raad, 2020). Furthermore, healthcare-related costs amounted to 12,7% of the Dutch GNP in 2015 and are expected to rise to 18% in 2060 (Medisch Contact, 2022).

The Dutch Population Health Screening employs around 950 professionals. The total costs for the screening program are unclear but estimated to be €125M out of €648M spent on preventive care in total, with around €80M directly allocated to the breast cancer screening program through the budget of Bevolkingsonderzoek Nederland (Nederlandse Rijksoverheid, 2021; PriceWaterhouseCoopers Accountants N.V., 2022). These employee and cost estimations exclude in-hospital care like that used in additional diagnostic procedures. Finding the current effects of screening and optimizing this process could significantly impact healthcare expenditure and staff shortages.

According to the literature, women experienced diagnostic breast procedures in various ways. Five often-reported motifs that affected the experience were found through thematic analysis of the literature, including fear, pain and discomfort, waiting, the physical surroundings, and staff interactions. These themes proved that women's experiences with diagnostic procedures went beyond the actual examination (Clark & Reeves, 2015). Screening was considered a preventative medicine and vital for a healthy lifestyle. Participants in distinct experiments showed pragmatism and comfort with the prospect of early disease detection. (Poulos & Llewellyn, 2005; Solbjør et al., 2011). However, pain and discomfort were common during the overall mammography examination, with studies reporting that over 70% of women experienced pain and discomfort (Drossaert et al., 2002; Engelman et al., 2008; Van Goethem et al., 2003). Alleviating the perception of women with regards to screening and its effectiveness could allow women to make a more informed decision on whether they should join the screening program and allows them to compare their own experience to the potential benefits.

1.2.2. Accenture Relevance

Accenture is a professional services company specializing in IT services and consulting. It is the most admired IT services company according to Fortune and is the global leader in many of its service offerings. Accenture is active in the healthcare sector and focuses on addressing and solving current issues. Four trends have emerged from Accenture's Digital Health Technology Vision 2022: WebMe, The Programmable World, The Unreal, and Computing the Impossible. The latter three of these directly relate to this research:

"The Programmable World tracks how technology is being threaded through our physical environments in three layers: Connected, experiential, and material. The Unreal explores the "unreal" qualities that are becoming intrinsic to artificial intelligence, and even data, making the synthetic seem passably authentic.

Computing the Impossible outlines the outer limit of what is computationally possible as a new class of computing machines emerges with a new curve of compute capability to tackle grand challenges.”(Accenture, 2022a)

This research aims to optimize the screening strategy for breast cancer in the Netherlands by replicating a cohort of persons in the Netherlands in a simulation environment. This simulation environment will then be used for experimentation through artificial intelligence to derive an optimal screening strategy. Accenture's quantum computing ventures aim to solve highly complex business problems much more rapidly and efficiently than classic computers (Accenture, 2022b). An extensive simulation like the one proposed in this research can be searched for optimal paths more efficiently using quantum computing in the future. Quantum algorithms can reduce Monte-Carlo simulations to Bernoulli trials, considering all possible outcomes (Blank et al., 2021; Heinrich, 2003).

Accenture is also involved in Healthcare on Azure in a joint venture with Avanade and Microsoft. This is the foundation for new ways to deliver healthcare, where they hold a leading global position in health care optimization. It is intended to aid professionals in making better clinical decisions and extend insights on direct care, administration, and finance in healthcare. It is an essential step towards commercial precision medicine and efficient care. This research will fit in perfectly with the goals of Healthcare on Azure, as the research aims to improve population health through a more thorough screening. Leveraging the Health on Azure platform and data could lead to more relevant and precise outcomes and is a perfect step for further research if this research provides a proof of concept for a working optimization system (Avanade et al., 2021).

The research has been conducted at Accenture in the Netherlands, under the client group Health and Public Service, in the Technology department. The initial results of this study are directly relevant to some of Accenture's clients, like government health departments and disease control centers. These parties are tasked with promoting public health in their respective constituencies and should therefore be concerned with the results of this study. Other clients from Accenture, like health insurers, might also benefit from the framework. The goal of the simulation framework is to be adaptable to other scenarios. Therefore, the tool could become helpful in determining the screening efficacy of other diseases. It could be used to act preventively on risk factors and even gauge the potential harm and benefit of other types of intervention.

Contributing to society through science with this research is one of the ways Accenture helps accelerate public healthcare. Besides their traditional consulting activities, they can show policymakers that they can aid in the optimization of the current healthcare landscape through research like this. Accenture can contribute to society by allocating company research resources to goals that align with their expertise and societal benefits.

Accenture is one of the leading companies offering future-proof IT solutions. They drive innovation in multiple ways, in healthcare and many more industries. This research could prove the benefits of combining scientific research and consultancy by utilizing and showing the potential of novel techniques Accenture possesses to deliver a quality product and report that is of use to many of Accenture's clients in healthcare.

1.2.3. Scientific Relevance

The literature does not provide consensus on the effectiveness of screening programs for breast cancer. A systematic review of previous studies can be found in Chapter 3. It is unclear whether all benefits weigh up to the potential harms. This research aims to give a better insight into the cost-effectiveness of the screening program.

Multiple systematic reviews recommend reevaluating the effectiveness of current screening programs. Expanding the current selection of models for evaluating screening with a model that includes utility and costs is necessary. Health economic analyses are only complete with the inclusion of costs and Quality Adjusted Life Years (QALYs). The systematic reviews, the Dutch Government, and the World Health Organisation (WHO) mention this absence.

Discrete event simulation models have evolved. The analysis model from the Erasmus University, MISCAN, is written in Delphi (Tan et al., 2006). Before 2006, this was a relatively well-known programming language. However, it has lost over 95% of its market share to a share of 0.12% currently, making it the least popular programming language currently in use. Usage of R and Python has increased 10-fold, making them some of the most popular languages around currently (*PYPL Popularity of Programming Language Index*, 2022). To keep models usable for the future, it is vital to use a future-proof language that can be written and interpreted efficiently. Delphi is objectively faster in a single-threaded experiment, but the increase in computing power and the rise of quantum computing makes up for that. The fact that R and Python are more popular, easier to write and understand, and are already used as interpreter languages for quantum computing, makes these languages better suited for this project. A new simulation model set up in these languages will engage more scientific users due to the ease of accessibility.

The scientific relevance of this research is due to the expansion of the knowledge base on breast cancer screening and relevant models associated with the analysis. It will advance our understanding of the viability of screening techniques by implementing recommendations from multiple reviews. It will quantify these results through models to reach a conclusion on the topic that needs clarification.

1.3. Research Objective and Questions

The goal for this research is defined as follows:

Assess the health-economic performance of breast cancer screening programs in the Netherlands by comparing populations that undergo different screening strategies through a newly developed population screening simulation tool.

Research questions need to be answered to achieve the research objective. The research questions provide a stepwise guide to solving the research problem. The questions provide a way to clarify the stages of research and provide answers needed to achieve the main research goal.

- 1) What is the current status of the breast cancer screening program?
 - a) What is the health impact?
 - b) What is the economic impact?
 - c) What is the best method to evaluate screening policies?
- 2) How does the population health with the currently implemented breast cancer screening policies compare to the health of a similar population with different screening strategies?
 - a) How is a successful screening policy defined?

- b) What are the relevant outcome metrics of a health economic analysis?
- c) Can a simulation model be created to compare the different strategies?
- 3) Are different screening strategies health economically viable?
 - a) What are the health outcomes of simulations that include patient utility and costs?
 - b) What are the total societal costs of these screening strategies?
 - c) Is the combination of these costs and health gains acceptable according to health economic evaluation metrics?

Each question will be answered in a separate chapter. Before answering the research questions, chapter 2 will give background information on breast cancer, population screening programs, and health economic analyses. In Chapter 3, research question 1 will be answered through a systematic literature review. In Chapter 4, questions 2a and 2b will be answered through a literature review. Question 2c will be answered by creating a simulation model. Question 3 will be answered in Chapter 5, where the results of the research will be presented. In Chapter 6, a conclusion on the main goal will follow. In Chapter 7, the study limitations will be discussed. This chapter will also introduce the recommendations for further research and ethical discussions on the screening program, health economic evaluations, and the outcomes of this research.

2. Background

This chapter will provide more information on the characteristics and treatment of breast cancer. Basic knowledge, current statistics, and literary conclusions on screening programs are also covered. Furthermore, principles and methods for health economic evaluations are explored through the literature.

2.1. Breast Cancer

The lifetime risk of breast cancer for women in the US is one in nine, and one in five breast cancer patients will pass away because of the disease. After age 40, age-specific incidence rates increase significantly. Breast cancer incidence rises with age in industrialized nations with high disease rates until plateauing at 75 to 80 years of age. Breast cancer is rare in all populations before age 35 (Strayer et al., 2014).

The proportion of diagnoses of non-invasive breast lesions increased significantly because of the widespread adoption of screening mammography in the 1980s (i.e., Ductal Carcinoma in Situ (DCIS)). Small invasive malignancies have become more common as well. Although widespread mammography screening has significantly boosted the diagnosis of early breast cancers, the prevalence of late-stage breast cancers has not significantly decreased. Stage-specific mortality has improved, and overall mortality has decreased from 30% to 20%. The drop in breast cancer mortality is primarily attributable to improved therapy (Strayer et al., 2014).

2.1.1. Anatomy

The breasts (mammas) comprise glandular and supporting fibrous tissue embedded in a fatty matrix and blood arteries, lymphatics, and nerves. The subcutaneous tissue covering the major and minor pectoralis muscles contains the mammary glands. The breasts have an extensive lymphatic system, and knowledge about lymphatic drainage is crucial for predicting metastases of cancer cells. The subareolar lymphatic plexus receives lymph from the nipple, areola, and lobules of the mammary glands. More than 75% of lymph drains from the lateral breast quadrants from the axillary lymph node to the anterior or pectoral nodes. Some lymph may flow directly to the other axillary nodes or the interpectoral, deltopectoral, supraclavicular, or inferior deep cervical nodes. While lymph from the inferior quadrants may travel deeply to abdominal lymph nodes, most of the remaining lymph, notably from the medial breast quadrants, drains to the parasternal lymph nodes or the opposite breast (Moore et al, 2018).

2.1.2. Pathology

The two major groups of breast cancers are carcinomas and sarcomas. Carcinomas develop from the breast's epithelial tissue. The cells that line the lobules and terminal ducts make up the epithelial component responsible for producing milk in normal circumstances. Most breast cancer cases are carcinomas. Sarcomas are uncommon malignancies that develop from the breast's connective tissue. Myofibroblasts and blood vessel cells are among the stromal component cells; cancers emerging from these supporting cells include phyllodes tumors and angiosarcoma. Less than 1% of initial breast cancers are sarcomas (Department Of Pathology, 2023).

Within carcinomas, there are more divisions in cancer types to be made. The first distinction is between in situ and invasive carcinoma. In situ, meaning in place is a pre-invasive carcinoma that has not yet spread to neighboring breast tissue. These cancer cells are growing in a confined environment of their places of origination, like the normal lobules or ducts. In situ carcinoma has the potential to spread and become

invasive cancer. Invasive cancer spreads outside the original environment and grows into the surrounding tissue, like connective tissue. Invasive carcinomas can spread further to other body sites, mainly through lymph nodes. These metastases can infect lymph nodes and other organs (Department Of Pathology, 2023).

Around 80% of carcinomas are invasive ductal carcinomas. Invasive lobular carcinomas make up approximately 10-15% of all cases. The remainder is made up of other special types of breast cancer. These include colloid, medullary, micropapillary, papillary, and tubular carcinomas. All these carcinomas have distinct features, varied prognoses, and treatment options. Therefore, differentiating between them during diagnosis is important (Department Of Pathology, 2023).

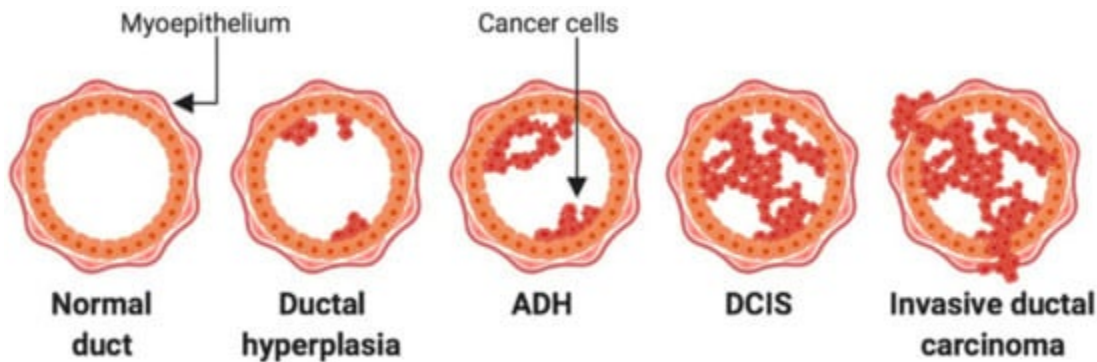


Figure 1: Stages of breast cancer development from ductal carcinoma (Tower et al., 2019)

Ductal carcinoma starts in one of the ducts in the breast. The myoepithelium of the normal duct can experience mutations, which can cause ductal hyperplasia. There is a potential for this hyperplasia to evolve into abnormal ductal hyperplasia (ADH). A proportion of the cells causing this hyperplasia are cancer cells, but it is too soon to tell if ADH will become a dangerous carcinoma. If the ADH cancer cells keep growing, Ductal Carcinoma in Situ (DCIS) is formed. DCIS is the most diagnosed type of breast cancer. At this point, it is not dangerous in all cases, but the potential to evolve into invasive ductal carcinoma is greater. Invasive ductal carcinoma is a tumor that should be treated, as it can endanger the host. The process of breast cancer development is shown in Figure 1 (Tower et al., 2019).

2.1.3. Disease Diagnosis and Staging

The largest part of breast cancer diagnoses in the Netherlands come through the breast cancer screening program. During the screening, a mammogram is made. This is an x-ray scan of the breast. The result of this mammography is defined in a BI-RADS classification (Breast Imaging Reporting and Data System-classification). The scores, definitions, first follow-up steps, and the corresponding probability of cancer are shown in Table 1. After a cancer diagnosis through biopsy, the cancer stage needs to be determined (Integraal Kankercentrum Nederland, 2022).

Table 1: BI-RADS scores and parameters (Ghaemian et al., 2021)

BI-RADS Score	Meaning	Follow-up	Likelihood of Cancer
0	Need additional imaging	New imaging	10%
1	Negative	Routine screening	~0%
2	Benign	Routine screening	~0%
3	Probably Benign	Biopsy or checkup within 6 months	0%<x<2%
4	Suspicious	Biopsy	>2%
5	Highly suggestive of malignancy	Biopsy	>95%

The American Joint Committee on Cancer (AJCC) publishes the cancer staging manual. This manual is vital in defining the diagnosis of various cancer types, including breast cancer. The staging of a tumor provides a route to various treatment options and allows for international comparisons of cancer statistics, outcomes of research, and clinical trials. The AJCC stages invasive carcinomas and DCIS according to the breast staging system; other breast cancer types are staged slightly differently. Sarcomas and phyllodes tumors are staged according to the soft tissue sarcoma system, and breast lymphomas similar to Hematologic malignancies (Amin et al., 2017).

The staging tables use T, N, and M tumor classification, sometimes combined with HER2, ER, and PR status. This means that the extent of the tumor (T), regional lymph node status (N), and distant metastases (M) are considered, combined with human epidermal growth factor receptor 2 (HER2), Estrogen Receptor (ER), and Progesterone Receptor (PR) status. A combination of these factors will lead to a clinical prognostic stage group. These stage groups range from 0-IV, with IV defining cancer with distant metastases and 0 defining DCIS (Amin et al., 2017).

The Dutch clinical guideline database, 'Richtlijndatabase,' uses the combination of T, N, and M status to define a clinical disease stage and a treatment plan. The risk profile and the cancer stage have been analyzed to calculate disease-specific 5-year survival rates (Amin et al., 2017; Nationaal Borstkanker Overleg Nederland & Nederlandse Internisten Vereniging, 2020).

2.1.4. Guidelines and Treatment

The Dutch clinical guideline database has published the complete guideline regarding breast cancer. For this research, it is crucial to consider the guidelines for diagnosis and treatment, as this will partly be implemented in the simulation model. The Integral Cancer Centre Netherlands (IKNL, Integraal Kankercentrum Nederland) has transformed the information from the guidelines into a care pathway flowchart. A copy of this flowchart can be found in Appendix I – Guidelines and Treatment

(Integraal Kankercentrum Nederland, 2020; Nationaal Borstkanker Overleg Nederland & Nederlandse Internisten Vereniging, 2020).

This care path includes all possible treatment options for breast cancer in the Netherlands. People with a screening result of BI-RADS 0,3,4, or 5 will be referred to the hospital through their General Practitioner (GP). After a positive diagnosis, a broad spectrum of treatment options is available. Various surgical procedures and chemo-, radio-, hormone-, and (neo)adjuvant systemic therapies can be carried out. The

proposed type(s) of therapy is based on the patient's tumor stage, hormone response, age, and clinical fitness (Integraal Kankercentrum Nederland, 2020)

As a vast range of treatments and diagnoses exists, this research will be constrained to averages and generally expected outcomes, with variable numbers drawn from representative probability functions for the patient's expected utility, life expectancy post-treatment, and chances of the disease worsening. The topic of interest for this research is the (pre-)diagnostic stage of the disease. It will be assumed that average survival rates from the cancer stage at diagnosis will be consistent regardless of the means of diagnosis. Survival rates and treatment costs will be based on averages per cancer stage.

Breast cancer is a severe disease with varying pathological pathways. Carcinomas develop from the breast's epithelial tissue, and Sarcomas are malignancies that develop from the breast's connective tissue. 56% of breast cancer is diagnosed through screening, and early diagnosis can drastically improve an individual's chance of survival. Stage I DCIS has an average 5-year relative survival of 99%, whereas late-stage III cancer survival can be as poor as 33% in 5 years. Many risk factors contribute to breast cancer development, both modifiable and non-modifiable. Additional information on breast cancer statistics, staging and risk factors is given in Appendix II

2.2. Population Screening Programs

To offer an early treatment or intervention and thereby lower the incidence and mortality of the health problem or condition within the population, screening is used to identify individuals in an apparently healthy population who are at higher risk of developing a health problem or condition. The WHO European Region is moving toward increased health monitoring and screening for noncommunicable illnesses. In some cases, there needs to be more proof of effectiveness. The potential harm of screening, its cost and burden on the health system, and the necessity of rigorous quality assurance are issues which policymakers, healthcare professionals, and the public frequently appear oblivious to (World Health Organization, 2020).

Screening is different from early diagnosis. Screening works like a sieve and is never 100% accurate. It separates people who likely have a condition from those who do not. The original aim of screening is not to be diagnostic but to refer individuals with positive or suspicious findings to other professionals for diagnosis. The aims of screening, according to the WHO, are the following (World Health Organization, 2020):

- Reduce mortality by early detection and early treatment (most clear in breast cancer screening)
- Reduce incidence of a disease by identifying and treating its precursors (most clear in cervical cancer screening)
- Reduce the severity of the condition by identifying patients and offering treatment (most clear in diabetic eye disease screening)
- Increase choice for individuals by identifying conditions at an earlier stage (most clear in prenatal screening)

2.2.1. Introduction

The Dutch National Population Screening Program consists of eight screening programs. Three screening programs are for cancer; the others are aimed at pregnant women and newborns. The three types of cancer screening are colorectal, cervical, and Breast cancer (RIVM, 2022).

The screening program in the Netherlands is part of the care chain, whereby timely transition from the screening program to hospital diagnostics and treatment through referral is key for improving population health. Eligible citizens are first invited for a screening round. They are expected to make an informed decision on joining the program through information facilitated by the 'Rijksinstituut voor Volksgezondheid en Milieu' (RIVM), the Dutch Center for Disease Control. After the screening, a person is referred to care if the results of the screening warrant further diagnostic tests or treatment (Population Screening Programs | RIVM, n.d.).

2.2.2. Screening Principles & legal grounds

In 1968, Wilson and Jungner published ten fundamental principles for screening in an article for the World Health Organization. The National Health Council ('Gezondheidsraad') condensed these principles into five vital points for Dutch healthcare. More information on screening principles and evaluations is given in Appendix III (Gezondheidsraad., 2008; Wilson & Jungner, 1968).

- Screening should be focused on an important health problem.
- Screening should be proven effective with benefits outweighing harm.
- Screening should be reliable and valid.
- Respect for autonomy should be central.
- Resources should be used efficiently, and explicit accountability regarding cost-effectiveness and equity is required.

The Dutch House of Representatives (Tweede Kamer) requested research on the effectiveness of screening in the Netherlands was published in 2014. It was stated that screening was effective and that around 17 percent, or 31 prevented deaths per 100,000 women were attributable to population screening for breast cancer. It is stated that the effectiveness met the initial expectations, but whether the full benefits outweigh the harms of screening still needs to be answered (Health Council of the Netherlands, 2014b).

The Netherlands has codified regulations on population health screening into law, into the 'wet op het bevolkingsonderzoek,' or law on population health screening (WBO). This law describes rules that screening programs should follow, how and when they should be evaluated, and how an acceptable proposition for a screening program can receive a permit for implementation.

Cost-effectiveness for healthcare in the Netherlands is an important metric to gain approval for interventions. The importance of cost-effectiveness, combined with the WHO's advice on reevaluating current screening methods, means that the breast cancer screening program in the Netherlands should be health-economically reevaluated. The best way to evaluate this would be through a cost-effectiveness analysis, emphasizing how these interventions impact current healthcare budgets. Further information on the legal grounds for screening and government policies for implementation is given in Appendix IV.

2.2.3. Drawbacks and Bias

Some biases and drawbacks make it difficult to assess the effectiveness of screening programs. The principles of Wilson and Jungner and the Health Council state that costs should be balanced compared to other healthcare spending per health gain. However, the Health Council also stated that this aspect had not been adequately assessed for breast cancer screening (Health Council of the Netherlands, 2014b). Therefore, the following drawbacks and biases are important to recognize to create a model that allows for accurate assessment.

- Data quality: data regarding incidence and mortality must be accurate.
- Data size: the often limited factual data can result in too large confidence intervals.
- Large incidence at start of screening: If a screening is first implemented, the first rounds will see an above average incidence.
- Mortality comparison over time: improvement in mortality over time can be due to screening, but also due to better treatment.
- Lead time bias: Earlier detection and longer perceived survival is not necessarily proof of the benefit of early detection.
- Length time bias: slow-growing tumors are more likely to be found through screening, aggressive tumors due to symptoms.

An explanation of these problems, an idea of how to deal with this, and how they are tackled in this research are given in Appendix V - Drawbacks and Bias.

Screening programs have the noble goal of improving population health, and if properly carried out, this can also be the result. For example, breast cancer screening in the Netherlands is assumed to prevent around 850 deaths annually. To create a successful screening program, Wilson and Jungner's ten principles of screening must be adhered to, as has also been recommended by the WHO. Screening programs should be evaluated periodically, for which there are multiple ways. These analyses also need to consider a range of commonly seen biases to prevent overvaluing the screening program.

2.2.4. Breast Cancer Screening

Bevolkingsonderzoek Nederland manages the current breast cancer screening program in the Netherlands. They organize the screening program on behalf of the RIVM. More than 1.2 million breast cancer screening program invites were sent out in 2021. Around 72.5% or 886.000 women participated in the population screening of these invites. This eventually resulted in the detection of 6362 breast cancer cases after around 22850 referrals for further examination (Bevolkingsonderzoek Nederland, 2022). RIVM states that the health gain of the breast cancer screening program is estimated at 850-1000 prevented deaths per year (RIVM, 2022).

Statisticians, physicians, and the public have questioned cancer screening programs and the added clinical value since at least the early 2000s (Lerner, 2016). Recent studies offer insights into the actual clinical benefit of the screening program, and the results of these studies differ wildly. The American Cancer Society (ACS) is an advocate of the screening program and recommends that women over 44 get mammograms every year (American Cancer Society, 2020). Other studies result in an opposite conclusion. Authors of the Danish Cochrane Collaboration concluded that current screening policies need re-evaluation, as universal screening may no longer be reasonable. This conclusion was aided by evidence of high false-positive rates, low decreased mortality in a screened versus a control population, and increased regard for emotional and physical distress for the screened population (Gøtzsche & Jørgensen, 2013).

2.3. Health Economic Evaluation

Health economics aims to provide analytical tools to support decision-making to promote efficiency and equity. It also provides a way to assess resource use in health care. Scarcity is an important topic: funds that are spent in one way cannot be spent elsewhere. Therefore, it is essential to decide where the

benefits are maximized. Health economic evaluation is about analyzing the costs, harms, and benefits associated with healthcare processes, intending to find the maximum benefits obtained with resource constraints (Shiell et al., 2002).

2.3.1. Types of Evaluation

There are multiple types of health economic evaluations. Some are explained in more detail below (*Economic Evaluation Overview | POLARIS | Policy and Strategy | CDC, n.d.;* McNamee et al., 2016):

- Cost Effectiveness Analysis (CEA), which measures any point of interest against money spent. This could be the expected reduction in cases per euro spent on COVID-19 vaccines or the gain in life years per euro spent on a screening program.
- Cost Benefit Analysis (CBA). A CBA weighs the total costs of an intervention against the monetary value of the benefits, as the valuation of life years and medical costs averted through this intervention. It results in a monetary net benefit/loss value.
- Cost-Utility Analysis (CUA). CUA combines the quantity of life gained with the quality at which this is experienced, resulting in Quality Adjusted Life Years (QALYs). The cost per QALY can then be compared to government standards.
- Cost Minimization Analysis (CMA). This is used to assess the best option if the outcomes in health are similar. It analyzes different treatment paths that lead to a similar outcome and advises choosing the most inexpensive option. This method is often unrealistic.
- Cost of Illness study (COI). A COI is often used to assist the other methods of analysis. It provides a set way to calculate the total costs of illness. This includes everything on a population level: from medical costs of diagnostic tests and treatment to travel costs and productivity losses. These calculated costs can then be used in another analysis.
- Budget impact analysis (BIA). To assess the total financial consequences of a new intervention, a BIA is done. Even if a new intervention's incremental cost-effectiveness ratio (ICER) is acceptable, the total budget impact might be too high for society.

Furthermore, there are some critical outcome metrics to consider. These are the ICER and the Incremental Cost-Utility Ratio (ICUR). ICUR is a version of ICER, where the effect is the utility. They are calculated by dividing the difference in costs by the difference in effects.

$$ICER = \frac{C_{Baseline} - C_{Experiment}}{E_{Baseline} - E_{Experiment}}$$

Equation 1

Acceptable ICER values are usually determined per country in a willingness to pay (WTP) threshold. For example, the Netherlands has a variable WTP threshold between 20.000 and 80.000 euros per QALY, depending on the severity of the disease. The ICER outcomes of a new intervention can be plotted on a CE-Plane. The outcomes of a health economic evaluation can be presented as an ICER. If the ICER from the CEA and the total costs of the corresponding BIA are acceptable, there is a good case for adopting the new intervention (Dowie, 2004).

2.3.2. Methods for Evaluation

It is essential to follow a set of guidelines to gain insightful results from a health economic analysis that are scientifically accurate and societally accepted. The European Network for Health Technology

Assessment (EUnetHTA), supported by the European Commission has compiled a guideline on various methods based on current practices in Europe. This document is based on many local guidelines, including the Dutch guideline for economic evaluations in healthcare. These documents will be used to evaluate multiple methods of health economic analysis and select one for this research (Swedish Council on Health Technology Assessment et al., 2015; Zorginstituut Nederland, 2016).

The guidelines on modeling good research practices, referenced by many European health economic analysis guidelines, define the following three types of models: State-transition models, discrete event simulation (DES) models, and dynamic transmission models. Dynamic transmission models will not be considered as these are meant for evaluating interventions against infectious disease. State transition models suffer from the Markovian assumption that transition probabilities do not depend on historical states. (Caro et al., 2012)

For this research, a discrete event simulation is preferred with a simulated cohort that changes over time. Patients' transition probabilities to subsequent states depend on previous states of specific attributes, which cannot be generalized but need to be patient specific. DES also allows for easier parameter customization and can be used to create visual representations of the model that could appeal to clients and the company. Individual-level DES models allow for patient-specific parameters that together resemble the total population characteristics.

2.3.3. Other Evaluations of Screening Programs

In 2020, a systematic review from the EU-TOPIA (European Union – Towards Improved Cancer Screening) concluded that there is evidence that screening programs cause a reduction in breast cancer-related mortality in Europe. The estimates range between 33% and 44% for the screening-related reduction in Western Europe. This review was based on 38 cohort studies, 17 case-control studies, and seven randomized control trials (RCT). It does not mention costs, quality of life, or utility (Zielonke et al., 2020).

A 2016 two-part systematic review and meta-analysis on the effectiveness and harms of breast cancer screening to provide an update to the screening recommendations in the USA was performed. The results were split into two papers: one on the effectiveness and one on the harms. The study on the harms concluded that false positives and overdiagnosis are common, up to 54%. Women with the wrong diagnosis experienced more anxiety, distress, and breast cancer-specific worry. Women also reported pain during mammography, with ranges varying wildly between studies (1%-77%). This results in women declining further screening (11%-46%).

Furthermore, there are estimates that between 2 and 11 deaths per 100,000 women can be attributed to radiation-induced cancer from mammography. The study on the effectiveness concluded that mortality is generally reduced due to screening. This reduction was insignificant in all age groups, and the effects were minor. The incidence of advanced cancer was reduced, but all-cause mortality was not. The research was based on 38 studies, including five other systematic reviews. Some of the studies were based on outdated imaging modalities. An overview of the quality of life and cost-effectiveness is not reported (Nelson, Fu, et al., 2016; Nelson, Pappas, et al., 2016).

A systematic review of the cost-effectiveness of risk-based screening from April 2021 concludes that risk-based screening could be an economically efficient alternative to the current screening situation. However, the review reported that none of the studies evaluated reported the adverse effects on the utility of unnecessary interventions. Productivity changes due to screenings are also ignored in almost all

studies, even though this contributes to a significant economic burden. Whilst the review's main focus was on the effectiveness of risk-based screening, it does provide insight into the current cost-effectiveness of screening and the limitations of studies regarding the effectiveness of screening (Khan et al., 2021).

Thirty-two economic evaluations of breast cancer screening and seven primary breast cancer prevention evaluations were analyzed for another systematic review. This review concluded that only half of these studies modeled harms related to overdiagnosis, often indirectly and without reporting the magnitude. These results still led to gains in life expectancy and QALYs. The costs associated with these increases vary per study, as seen in the graph below. Recent studies in the UK and Netherlands show that costs per life year gained may be higher than presented by evaluations from the Dutch government. The review also concludes that the harms and losses of quality partly offset the potential gains in a lifetime and quality of life due to overtreatment and overdiagnosis. This has not been adequately accounted for in the evaluated studies, and studies that try to account for it do this in an unclear, ambiguous way. The results from the systematic review were summarized in Figure 2 (Mühlberger et al., 2021).

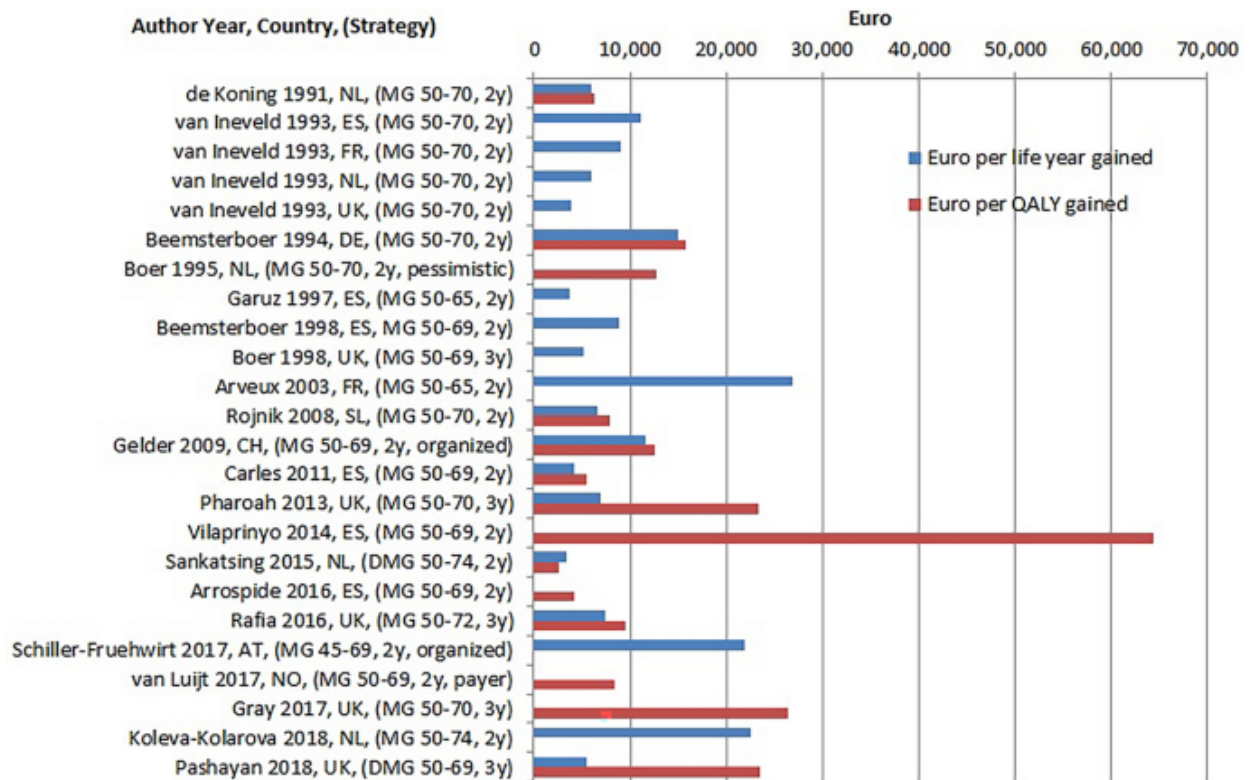


Figure 2: Costs per (QA)LY gained as per the Mühlberger et al. systematic review (Mühlberger et al., 2021)

The conclusions from another systematic review are similar: Quality of Life (QoL) values currently used in studies do not fully capture benefit and harm, undervaluation of these effects might lead to inappropriate advice and decisions on screening, the methodological challenges for valuing QoL in mammography need to be addressed, and the risks of screening should be incorporated in the valuation process. These are the highlights from a 2019 systematic review of the economic measures on breast cancer screening programs performed by the Universities of Melbourne (AUS) and Birmingham (UK). The review also recommends

further research that does include all relevant utilities and disutilities for the economic evaluation of breast cancer screening (Bromley et al., 2019).

From these systematic reviews and the underlying papers, it can be concluded that there is still much uncertainty regarding the economic viability of breast cancer screening. There is a good chance that the screening program does reduce cancer-related mortality, but the cost at which this happens is unclear. Assessing quality-adjusted life years gained or lost through screening has often been too difficult. More research is needed to determine if the ICER of screening for breast cancer is within acceptable WTP thresholds. First, a new systematic review will be carried out that evaluates more papers on health economic analyses of the breast cancer screening program. After that, a new health economic evaluation should be performed that accounts for utility values in the screening process and include various recommendations for further research that were found through the new systematic review.

3. Systematic Review of Breast Cancer Screening Program Evaluations

A Systematic Literature Review was performed to assess the processes and outcomes of previous health economic evaluations on breast cancer screening programs. The systematic review in article form can be found in Appendix VI. The highlights of the findings can be found here.

3.1. Introduction

For this systematic review, the following research question was formulated:

- What is the cost-effectiveness of the breast cancer screening program?

Three sub-questions were formulated to answer the main question and to provide context on the methods for performing a health economic analysis on the breast cancer screening program.

- What are health economic analysis methods used to evaluate the cost-effectiveness of breast cancer screening programs?
- What are the parameters used in health economic analysis on the cost-effectiveness of breast cancer screening programs?
- What are the reported outcomes of health economic evaluations for breast cancer screening programs?

The systematic review aims to answer these questions to provide an overview of and comparison between HEA methods, parameters, and outcomes. These outcomes can be used for a further, independent HEA on the Breast Cancer Screening Program.

3.2. Methods

This section presents the data processing method and selection of the found material used for this systematic literature review. This systematic review was performed according to preferred reporting items for systematic reviews and meta-analyses guidelines (PRISMA) (Moher et al., 2009).

3.2.1. Data Sources and Search Method

A systematic search is done in the databases of Scopus and PubMed. The search aims to find health economic analyses on breast cancer screening programs using simulations or models. The articles covering this subject are selected between the years 2016 and 2022. The search was based on the following topics:

- Health economic analysis and/or cost-effectiveness
- Breast cancer screening and/or mammography screening
- Simulation and/or model

Combining these topics resulted in the following search query shown in Table 2. The query is used for the systematic literature review and given as input for the selected databases. The search is applied for the title, keywords, and abstract in these databases. The resulting list of articles is sorted by relevance. The choice was made to include 'Simulation and/or model' in the search query, as leaving this out would result in a significant multitude of results, most of which were focused on health economic outcomes of novel breast cancer screening methodologies. In contrast, this study focuses on the HEA of the current methods.

3.2.2. Inclusion and exclusion criteria

In this systematic review paper, only papers including the keywords Cost-effectiveness, breast cancer screening, and simulation are included. However, some synonyms of these keywords were also added to prevent missing out on relevant literature. These can be found in Table 2. Articles before 2016 were excluded from this search because of the time relevance. Papers not describing specific methods and outcome measures relevant to this research were excluded. Lastly, case studies and systemic review papers were also excluded due to their irrelevance in writing systematic reviews. This selection procedure is shown in Figure 3.

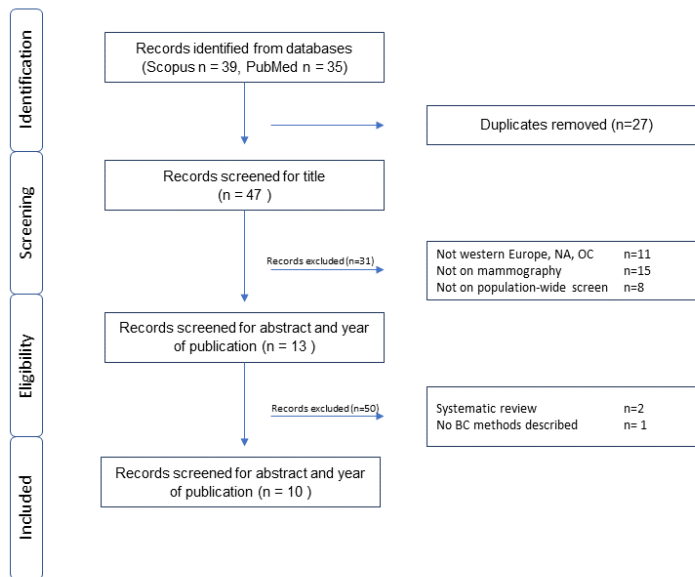


Figure 3: Paper down-select according to PRISMA method

Table 2: Main search terms and synonyms for search query

Cost-effectiveness	Breast Cancer Screening	Discrete event Simulation
Cost benefit	Mammography screening	Simulation
Cost utility	Breast screening	Markov model
Cost of illness		Population model
Budget impact		Cohort model
Health economic		
Value for money		

3.2.3. Data Extraction and Synthesis

The results gained from the databases, based on the in- and exclusion criteria, were checked by the author by reading the abstract of each article. Then the author decided which articles were relevant. This resulted in ten articles that provided relevant information for this systematic review. These articles were read, and

data were extracted according to a form. One person reviewed the publications. The following data were extracted from all publications:

- Basic information
 - Title
 - Authors
 - Year of publication
 - Country of research
 - Country of data
- Research question/objective
- Main screening strategy
- Target population
- Type of HEA
- Type of simulation
- Outcome measures
- Cost outcome metric
- Costs used
- Utilities used
- Tumor growth model

Extracting this data allows the separation of the findings into categories, e.g., cost estimations from papers that use Quality Adjusted Life Years (QALYs) can be compared to those that use Life Years Gained (LYG).

3.2.4. Risk of Bias assessment

The databases used for this systematic literature review are PubMed and Scopus. However, many more databases could have been included in the research. Leaving these databases out makes the risk of leaving out relevant sources more significant. Next, there is the possibility of publication bias, where studies showing adverse outcomes are not published. Another bias is the researcher's personal bias, which is a relevant bias as the research was performed alone. The use of PRISMA guidelines for performing this systematic review is a personal choice, as the author has worked with this method before. This method might introduce a bias, resulting in different paper selections and results than when another systematic review methodology was followed.

3.3. Results

3.3.1. General Characteristics

In Table 3, the characteristics of the ten selected articles are displayed. The articles were all published in 2016 or later. The locations were split into European Union and North America (the paper from the UK was from 2016 and is therefore classified as EU). These papers' research methods and outcomes differ in the type of HEA, the primary reported outcome, types of costs, utilities, and tumor growth model used. Most articles researched the costs per non-quality-adjusted or quality-adjusted life year, whereas one article only compared the harms and benefits, not the costs. The table displays the exact number of publications per category. The most used HEA is a cost-effectiveness analysis resulting in costs per QALY.

Table 3: General Characteristics of included papers

Characteristics	Value	N
<i>Year of publication</i>	2016	2
	2017	2
	2018	3
	2019	2
	2021	1
<i>Location of research</i>	European Union*	6
	North America	4
<i>Location of Data</i>	European Union	5
	North America	4
	Other (Norway)	1
<i>Type of HEA</i>	Cost-Effectiveness	9
	Budget Impact	2
	Harm-benefit	1
<i>Main reported outcome</i>	cost/LYG	4
	cost/QALY	6
	other	1
<i>Costs considered</i>	Screening	9
	Diagnosis	9
	Treatment	9
	Societal	1
<i>Utilities considered</i>	Per disease stage	6
	Per event	2
	per adverse effect	1
<i>Tumor Growth Model</i>	Not defined	2
	Markov Model	1
	Growth equation	7

Table 4: Type of Health Economic Analysis performed per paper

Main outcome measure	Type of Health Economic Analysis					
	Paper	Harm-benefit	Cost-Effectiveness	Cost-Utility	Budget Impact	Comparative Effectiveness
<i>Harm/benefit €/QALY</i>	Zielonke 2021	X				
	Rim 2019		X	X		
	Shih 2019		X	X		
	Mittmann 2018		X	X	X	
	Rafia 2016		X	X		
	v.Luijt 2017		X	X		

€/LYG	Arrospide 2016		X	X	X	
	Carter 2018		X			X
	Schiller 2017		X			
	Koleva 2018		X			

Table 5: Costs measures included in calculation of outcome

		Costs Used			
Main outcome measure Harm/benefit €/QALY	Paper	Screening	Diagnostics	Treatment	Social Burden
		€/LYG	Zielonke 2021		
	Rim 2019	X	X	X	
	Shih 2019	X	X	X	X
	Mittmann 2018	X	X	X	
	Rafia 2016	X	X	X	
	v.Luijt 2017	X	X	X	
	Arrospide 2016	X	X	X	
	Carter 2018	X	X	X	
	Schiller 2017	X	X	X	
	Koleva 2018	X	X	X	

Table 6: Utility measures used in calculation of outcome

		Utilities Used			
Main outcome measure Harm/benefit €/QALY	Paper	Different per Cancer stage	Screening	Diag-nostics	Notes
		€/LYG	Zielonke 2021		
	Rim 2019				Decremental after detection
	Shih 2019	X			
	Mittmann 2018	X			
	Rafia 2016	X	X		Slight decrease for stress and anxiety
	v.Luijt 2017	X	X	X	
	Arrospide 2016	X			Different Age ranges
	Carter 2018				NA
	Schiller 2017				NA
	Koleva 2018				NA

Table 7: Tumor growth model used in simulation

Main outcome measure	Paper	Euro / outcome	Type of simulation	Tumor growth model used
Harm/benefit €/QALY	Zielonke 2021		Microsim (MISCAN)	Exponential
	Rim 2019	53918	Microsim (Stanford)	Constant
	Shih 2019	43087	Microsim (MDACC)	Exponential
	Mittmann 2018	32134	Microsim (Wisconsin)	Gompertz
	Rafia 2016	21266	Microsim	Exponential
	v.Luijt 2017	12445	Microsim (MISCAN)	Exponential
	Arrospide 2016	5117	Microsim	Progressive stochastic
€/LYG	Carter 2018	28415	Monte Carlo computer simulation (Treeage)	Gompertz
	Schiller 2017	23918	Microsim	Markov state transition
	Koleva 2018	6582	Microsim (SiMRiSc)	Exponential

3.3.2. Study results

The results of the studies are analyzed and summarized in four tables. The first table, Table 4, discusses the type of health-economic analysis used in the research. The second table, Table 5, discusses the costs used, and the third table, Table 6, discusses the utilities used in the different papers. The final table, Table 7, summarizes the results of the papers in the cost per outcome, the type of simulation, and the tumor growth model used.

Between the articles, there are many similarities. Nine out of ten articles choose a cost-effectiveness analysis as the primary health economic analysis. Nine out of ten articles choose a discrete-event microsimulation as the simulation type. Even though research often starts in similar ways, the outcomes vary greatly. The results can be seen in Figure 4. All values are corrected for inflation and converted to Euros.

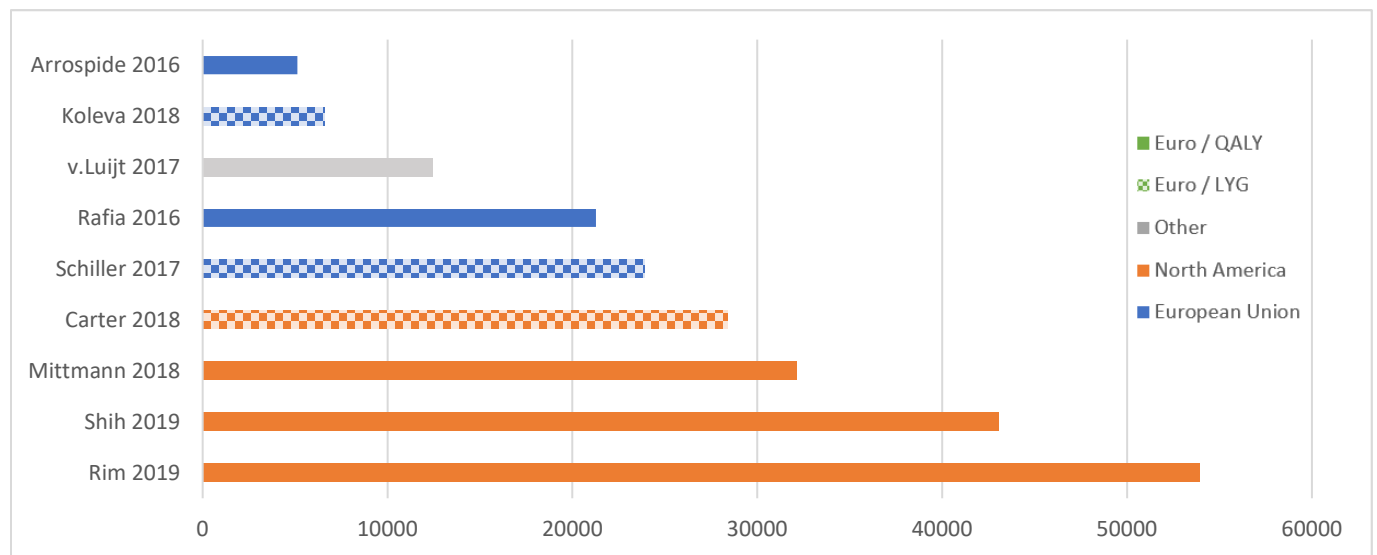


Figure 4: Costs per QALY and LYG outcomes of the different papers

3.4. Analysis of the results

3.4.1. Cost versus benefit

There is no clear consensus visible on the costs per QALY or LYG. There is also no statistical significance in whether the calculations for cost per life year gained are higher than the costs per quality-adjusted life year gained. The results vary too much to state that screening programs fall above the willingness-to-pay threshold. The willingness to pay threshold in the Netherlands varies between €20,000 and €80,000 per QALY, depending on the severity of the disease (Nederland, 2018). None of the studies included all cost criteria from the Dutch guidelines for health economic evaluations (Versteegh et al., 2016).

3.4.2. Cost versus region

Screening programs in North America cost more per gained (quality-adjusted) life year than in the European Union. This could be due to higher healthcare expenditures in general, qualitatively differing screening techniques, or evaluations that took more costs into account.

3.4.3. Use of cost measures

Most papers include the costs of screening, diagnosis, and treatment as inputs to calculate the total cost. Only the paper from Shih et al. (Tina Shih et al., 2019) additionally used the productivity loss due to early death in their cost calculation. None of the papers provides societal cost estimates for productivity loss for visits to screening facilities. The paper from Schiller-Fruehwirth et al. (Schiller-Fruehwirth et al., 2017) notes that the cost estimate is from a societal perspective and, therefore, only uses healthcare-related costs by society. The paper from Arrospide et al. makes an interesting note on implementation costs. It states that the costs per gained life year are a difficult indicator, as you cannot implement this intervention for just one patient: to create an effective screening program, it needs to be implemented for the entire population. This will result in a significant annual budget impact on the healthcare system. Other interventions for disease often only incur costs for people that are ill, resulting in lower total costs, even at higher costs/QALY rates (Arrospide et al., 2016).

3.4.4. Use of utility measures

The six papers calculating the added QALYs did this differently. In two papers, minor reductions to the current quality of life (QoL) were given during screening and diagnosis (Tina Shih et al., 2019; van Luijt et al., 2017).

One of the papers updates the QoL for a fixed duration after receiving a certain treatment, before returning to a baseline (Tina Shih et al., 2019).

One of the papers uses an initial decrease after diagnosis and then a continuous decrease per year for the duration of the disease (Rim et al., 2019).

The paper from Rafia et al. updates patients' QoL values at the time the disease progresses to the next stage, rather than after diagnosis. Furthermore, it adds temporary disutility during the screening progress and during an 'anxiety phase' when a patient is awaiting the diagnosis results (Rafia et al., 2016).

Four out of six papers updated the QoL of the patients at the time that the disease clinically progressed to the next stage before the patient is notified or diagnosed (Arrospide et al., 2016; Mittmann et al., 2018; Rafia et al., 2016; van Luijt et al., 2017).

3.4.5. Tumor growth models

Very little information can be found on the initiation of the tumor growth models. Modeling primary tumor growth is essential for determining screening diagnosis parameters. As for the tumor growth models, there is consensus on an initial exponential growth phase. Some papers choose to flatten the curve at a certain tumor size through a Gompertz equation, representing empirical evidence about slowing growth after reaching a certain size. Few papers consider the possibility of naturally occurring regression (Schiller-Fruehwirth et al., 2017; van Luijt et al., 2017). The other papers opt for a constant growth rate at which a tumor keeps growing, albeit constant, or entered in an equation determining the size according to a function.

3.4.6. HEA Methods

Nine out of ten studies analyzed in this systematic review covered a cost-effectiveness-focused health economic analysis. Only one study performed a harm-benefit analysis (Zielonke et al., 2021). Of these nine studies, all nine performed a Cost-Effectiveness analysis with the outcome focus on the cost per gain in life years or quality-adjusted life years. These CEAs were all performed through discrete event simulation (DES). Of the DES models, five papers used a model listed on the CISNET model registry (National Cancer Institute, 2022). One of the models was reused from another paper, one model used a ready-to-use computer program, and two models were developed specifically for the study. DES is the preferred modeling method for health economic analyses. Other modeling options were included in the search query and named in the resulting papers but were not used in any research analyzed.

3.4.7. Harm versus benefit

The 2021 paper from (Zielonke et al., 2021) is the only included paper that performed a HEA on the harms and benefits of breast cancer and did not include a cost calculation. The paper focused on the effects of changing the age ranges of the screening program. The increase in breast cancer deaths averted and life years gained was compared to the increase in overdiagnoses and false positives. The research concludes that the current age ranges of 50-75 are optimal for the Netherlands. However, further research to extend the harm-benefit analysis to a cost-effectiveness analysis is necessary to advise on health policymaking. These further studies should include additional screening and treatment-related effects in the calculation of quality of life.

3.5. Discussion

None of the papers calculated the societal burden associated with people visiting screening facilities, despite the possible substantial impacts on the total costs at this scale. This constitutes a significant gap in our understanding, which requires further consideration in implementation and research efforts. In this regard, the Dutch health economic analysis guidelines emphasize the importance of determining societal costs and short-term burdens in such evaluations. This view is reinforced by the suggestions of Carter et al. and Van Luijt et al. (K. J. Carter et al., 2018; van Luijt et al., 2017), who call for a more comprehensive analysis of the negative effects of screening and societal burdens in future research.

The way the QALYs are calculated varies between the papers, and it seems difficult to find a consensus on how to calculate this. Papers including more disutility for calculating the QALY, end up on the higher ends of the cost per QALY for that region. Rafia et al. conclude the highest cost/QALY calculation for the EU, and Rim et al. and Shih et al. the highest in NA, both using more QoL adjustments than their counterparts (Rafia et al., 2016; Rim et al., 2019; Tina Shih et al., 2019). However, there is too little evidence to conclude

a causal correlation at this point due to fractured data from different regions that make it difficult to draw meaningful comparisons, and it remains unclear why there is so much variation in the calculation of the costs per (quality-adjusted) life year between the various papers.

Four out of six papers updated the patients' QoL when the disease clinically progressed to the next stage before the patient was notified or diagnosed. Dutch guidelines suggest using QoL values of patients at different stages, and these QoL values are only measured after diagnosis. It is unclear if having received the diagnosis lowers a patient's perception of his quality of life. This effect is unaccounted for in research but presumed to exist in practice.

As for the tumor growth models, the Gompertz functions best replicate the natural history model of observed tumor sizes but lack a regressive component. Further research could include adding the possibility for random regression and growth stagnation at times, which would make for more realistic growth patterns. This is suggested by the paper from Carter et al. regarding overdiagnosis and false positives. A better understanding of tumor growth and nonprogressive tumors would lead to better implementation of overdiagnosis in models.

The studies agree on using discrete event simulation for health economic analysis and focus on cost-effectiveness in general. The adaptability and expandability of these DES models make them versatile, and often models can be reused and adapted for other research, as shown by studies adapting CISNET models.

No papers from later than 2021 were included, and no papers later than 2019 showed quantitative cost data. Papers in this category could not be found with this search query, and it is assumed that research on the cost-effectiveness of breast cancer screening programs has not taken place or has not been published.

The data used in the various papers included in this review may not be directly comparable due to differences in the data sources, which can lead to difficulties in comparing the results. Costs of the screening program, disease treatment, normal life expectancy, cancer incidences, and screening effectiveness can still differ between these countries. Even though this review attempted to limit this by only including papers from Western Europe and North America, it cannot be said that the data used for these analyses is interchangeable, and conclusions on one country do not necessarily apply to others.

3.6. Conclusion

Screening has the potential to increase the well-being of a population. According to current literature, these health benefits fall within the conventional cost-benefit parameters. However, different research uses a lot of different methods for the approach and calculation of these harms and benefits, and none of the papers manages to catch all factors described in these papers in one research. Furthermore, research from different countries cannot be used to conclude other countries. Further research is necessary to create a sounding conclusion on the cost-effectiveness of breast cancer screening, factoring in different possibilities of tumor growth, different QoL measures at various points in a patient's life, and summing up all costs incurred over a patient's life that are attributable to the screening program. It is challenging to decide what factors should be counted, how they should be counted, and if they can be counted at all. However, guidelines on these types of analyses provide more information on the calculation of societal burden, short-time decreased utility, and integration of post-care healthcare-related costs. Progress in determining the effectiveness of screening programs can be made by following (inter)national guidelines more explicitly.

All papers, except for one, use a microsimulation as the method through which the analysis is performed. A microsimulation is a computer model that simulates individuals or smaller entities. A Discrete Event Simulation is a form of microsimulation that allows more interdependencies between timings and patient attributes than the broader term microsimulation. Some of the used microsimulations are also discrete event simulations. The papers included in the systematic review consider discrete event simulation and microsimulation to be the best methods to perform health economic analyses on the breast cancer screening program.

The first sub-question, 'What is the current status of the screening program according to the literature', can now be answered. The health impact is perceived to be positive but is challenging to quantify. The paper from Zielonke notes that the benefits of the screening program outweigh the potential harms, which is in line with public perception and portrayed messages from screening initiators. The economic impact varies greatly over the different papers and ranges between €5000 per gained life year and €54000 per gained quality-adjusted life year. Some of the calculated costs are above the €20K willingness-to-pay threshold, but all fall within the €80K WTP threshold. There is consensus on the method for health economic evaluation of screening programs, which is through microsimulation or the more detailed discrete event simulation. It is important to note that none of the papers included all costs that should be considered according to the Dutch guidelines for health economic analyses (Zorginstituut Nederland, 2016). However, some papers call for the inclusion of more cost, harm, and benefit parameters in future research.

The systematic review shows that breast cancer programs are evaluated positively. Papers hold a critical note where they recommend the inclusion of more costs, harms, and benefits in future analyses, but in general, papers are in favor of the current screening programs. Although difficult to calculate, the health impact is significant, and the economic impact falls within acceptable margins. Therefore, the status of screening programs according to the literature is positive.

4. Methods

The Methods section will answer question 2. The definition of a successful screening policy will be given in 4.1, where the Dutch guidelines on health economic analyses will be discussed. In 4.2, relevant outcome metrics of health economic analyses are presented and explained. In 4.3, the idea behind a discrete event simulation is explained before explaining the model built for this research. In 4.4, the desired output of the model is discussed, as well as methods for calibration and validation. In 4.5, the experimental setup is given that will eventually be used to get to the results for Chapter 5.

4.1. HEA Guidelines

The Dutch Health Institute (Zorginstituut Nederland) published guidelines for performing economic evaluations in healthcare. It contains a step-by-step guide for HEAs, including best practices for costs and quality of life calculations. This subchapter will summarize these steps so that they may be implemented in the model for this research (Zorginstituut Nederland, 2016). It will go over the standard analysis in a health economic evaluation, the topic of evaluation will be formulated, the means of analysis will be explained, the cost input data as prescribed in the guidelines will be discussed, and the definition of a successful screening program will be given.

4.1.1. Standard Analysis

The guidelines describe a standard analysis, which should be used as the base case for most analyses. It contains the following important criteria:

- The evaluation should be focused on a societal perspective,
- The Patient, Intervention, Comparison, Outcome, and Time (PICOT) should encompass the Dutch population, compared to standard care, and evaluated over a lifelong timeline,
- The default preferred economic analysis is a Cost-Utility Analysis,
- The costs should be discounted by 4%, health effects by 1.5% annually,
- Uncertainty and sensitivity analyses should be univariate, and probabilistic sensitivity analyses and scenarios should be shown,
- The effectiveness of interventions should be based on systematic reviews,
- All costs relevant to healthcare, patient, and family should be considered, including a friction cost method for productivity losses. Future costs of care due to longer patient survival should also be included,
- Effects should be expressed in QALYs, according to the EQ-5D-5L Dutch valuation. QALYs are a generic instrument to measure the quality of life in 5 levels. If relevant, gained life years should also be calculated,
- Reports and results should include the total and incremental costs and effects and the ICER. A univariate sensitivity analysis should be expressed in a Tornado-diagram and table. Scenario analysis through a table and probabilistic sensitivity analysis through a CE-plane and Cost-Effectiveness Acceptability Curve (CEAC).

4.1.1. Analysis Framework

The topic of a health economic evaluation should be clearly defined before starting the analysis. The guidelines suggest defining the goals, user, perspective, and PICOT or research question. The goals and relevance of the research have been defined in Chapter 2. The perspective of the study has been described in earlier chapters and through the conclusion of the systemic review. The perspective of this study will

be broader than earlier studies through the standard analysis suggestions of the guidelines. The evaluation will be performed from a societal perspective, meaning all costs, benefits, and harms will be incorporated. The PICOT elements for this research can be seen in Table 8.

Table 8: PICOT scope for this research

PICOT	Research scope
Patient	The entire female population of the Netherlands
Intervention	Bi-annual screening for breast cancer
Comparison	No screening
Outcome	ICUR/ICER, LYG, CE-plane, CEAC
Time	Lifetime

The guidelines suggest further refinement for incorporating a patient cohort, interventions, controls, and outcomes. For this research, refinements will be implemented through an iterative approach. The goal is to create a working model first and update that model according to, for example, relevant demographic data for the population. The model will be validated at different points in the iterations to enable more precise finetuning to create a realistic-as-possible comparison to the real world.

4.1.2. Analytical approach

Multiple techniques can be used to perform a health economic analysis. The guidelines describe multiple evaluations and approaches. The most common evaluation types are CEA, CUA, and BIA. The preferred analytical methods are through decision models, either as an empirical or modeled approach. Common types of models are described in 2.3.2, with decision trees, Markov models, discrete event simulation models, and dynamic transmission models as most often referenced. The systematic review in 2.3.3 found an overwhelming preference for cost-effectiveness analyses through simulation. The ISPOR/SMDM guidelines also suggest that DES models are best suited for the analysis in this research. This research will therefore focus on a CEA in the form of a CUA with an additional BIA.

Discounting effects and costs become increasingly important if the analysis timeline extends far into the future. Discounting is the process of converting a value received in the future to an equivalent value if it were received today. The Dutch guidelines suggest a cost discount rate of 4% and an effect discount rate of 1.5% annually.

In order to increase the reliability of the results, an uncertainty analysis should be performed. Uncertainty analyses determine the variability of average costs, effects, and ICER and quantify the effects of this uncertainty. In patient-level models, like the simulation for this research, variability is essential to remember. The population size should be large enough to create a consistent result with each performed simulation. Used distributions and their parameters should be reported and explained, and the influence of parameter uncertainty should be evaluated with probabilistic sensitivity analyses (PSA). The resulting ICER should be calculated by combining the results of multiple PSA iterations. Univariate and multivariate sensitivity analyses can also be used to determine the relative influence of input parameters on the model's outcomes.

It is highly likely that not all data for the model can be found through databases or literature. This missing data could still be essential for the research. The guidelines therefore define ways to deal with missing data. Methods range from asking expert panels for their opinion to extrapolation techniques.

Studies also need to be validated in order to create results that are relevant for the Netherlands. The model itself, the input data, the code, and the outcomes need to be validated. The tool AdVisHE provides a checklist to test and validate all relevant parts of the research.

4.1.3. Input data

Data on effectiveness in a model-based study should be based on relevant literature. These studies and their data can be gathered through a literature search by performing a systemic review, using other systematic reviews, or combining other studies. The guidelines advise combining multiple randomized control trials into useful data, even if the comparison between intervention and effect is not direct.

Costs can be approached from distinct perspectives: the societal, healthcare, or patient perspective. As this research focuses on the societal implications, the societal perspective is most appropriate. The guidelines prescribe the use of the cost categories seen in Table 9 in a study from a societal perspective.

Table 9: Overview of costs to be included in a HEA from a societal perspective

Perspective	Cost Categories	Specification
<i>Societal</i>	Healthcare	<ul style="list-style-type: none"> • All healthcare costs related to intervention • All healthcare costs in gained life years
	Patient & Family	<ul style="list-style-type: none"> • Travel costs • Own payments • Time costs (e.g. missed work) • Home care
	Other sectors	<ul style="list-style-type: none"> • Costs outside of healthcare • Costs of productivity loss

For all these costs, relevant values will be derived from the literature and the cost guidelines recommended by the Dutch HEA guidelines.

Values for quality of life (QoL or Utility) are used to measure QALYs. These need to be based on validated questionnaires or data from the literature. The preference for the standard analysis is to use data from the EuroQol 5 dimensions 5 levels (EQ-5D-5L) questionnaire, which can be used to define a patient’s utility at a certain disease stage. For this model, relevant utility values will be derived from the literature, preferably from Dutch research that defined the QoL at different intervention and disease stages using the EQ-5D-5L questionnaire. If EQ-5D-5L values cannot be found, other values can be used if they are validated, and the use is motivated well.

For all these costs, relevant values will be derived from the literature and the cost guidelines recommended by the Dutch HEA guidelines.

Values for quality of life (QoL or Utility) are used to measure QALYs. These need to be based on validated questionnaires or data from the literature. The preference for the standard analysis is to use data from the EuroQol 5 dimensions 5 levels (EQ-5D-5L) questionnaire, which can be used to define a patient's utility at a certain disease stage. For this model, relevant utility values will be derived from the literature, preferably from Dutch research that defined the QoL at distinct intervention and disease

stages using the EQ-5D-5L questionnaire. If EQ-5D-5L values cannot be found, other values can be used if validated and the use is motivated well.

4.1.4. Screening policy

This leads to the following outcomes that Combining the information from chapters 2.2, 2.3, and 4.1 allows us to formulate an answer to sub-question 2.1, the definition of a successful screening policy. There are a couple of criteria for a screening program to be successful.

The first is the susceptibility to the Wilson & Jungner criteria, as defined by them in 1968 and reiterated by the WHO in 2020. They were refined more by the Dutch RIVM in 2008 into five key points: focusing on an important health problem, proven effectiveness, reliability, and validity, respecting autonomy, and efficient resource use. A screening policy can be implemented for four main reasons. Reducing mortality by early detection and treatment is most important for the breast cancer screening program (RIVM, 2022; Wilson & Jungner, 1968; World Health Organization, 2020).

A health economic evaluation can be performed to evaluate criteria 2, 5, and the reason for implementation. The Dutch guidelines advise performing these analyses through discrete event simulation, in line with the papers in the systematic review of Chapter 3. Chapter 2.3 describes different types of evaluations. Chapter 4.1 describes the method of how this should be performed in a Dutch context. The other criteria for a successful screening policy are considered met. Breast cancer is an important health problem, mammography is reliable and valid, and individuals can make an autonomous choice on whether they want screening.

4.2. Outcome Metrics & KPIs

Answering the rest of the research questions requires information on what is needed from the health economic analysis regarding outcomes. The Dutch policies ask for certain outcomes in the standard analysis, as seen in 4.1. The outcomes that need to be achieved are the total and incremental costs and effects of implementing a screening program, discounted with their respective rates. These costs should also be evaluated through a CE-plane and CEAC. The important cost outcomes are derived through a Cost-Utility analysis, including a Budget Impact Analysis.

Knowing if the found costs are acceptable to Dutch standards is essential. The CE-plane and the CEAC can help with that. The CE-plane provides a graphical way to show if the cost increase or savings combined with the gain or loss in QALYs is acceptable. The resulting cost from a simulation can be plotted on the base graphic, which can be seen in Figure 5. Many different cost and utility values will come forth from the various simulations performed during analysis. All these costs and utilities will be plotted on the plane. Figure 5 shows the various areas of the curve to explain whether an outcome in that plane is acceptable. Two Dutch willingness-to-pay lines are added in: one at €80,000 per QALY, which is often acceptable for severe diseases, and one at €20,000 per QALY, which is acceptable for most diseases in the Netherlands (Zorginstituut Nederland, 2015). More information about the willingness to pay thresholds can be found in Appendix IV. In general, new therapies are accepted if they fall into the areas marked by the green colors. If the disease or disutility experienced by the patient is severe, the blue category can also be acceptable. The orange categories are rarely acceptable: they provide clinical benefits at too high costs or cost savings at too much quality loss. The red categories are unacceptable: the cost savings are paired with too much loss of quality of life. The yellow category is debatable: patients will have slightly less quality of life, but this goes paired with enormous cost savings. In some cases, this can be acceptable.

The CEAC, or Cost Effectiveness Acceptability Curve, shows the likelihood that a new intervention is cost-effective when plotted against various willingness-to-pay thresholds. The analyses will show varying values on the CE plane. The CEAC will translate these values to a probability of cost-effectiveness at various WTP thresholds by determining which share of the outcomes is within the acceptable margins of a certain willingness to pay. An example of a CEAC can be seen in Figure 6.

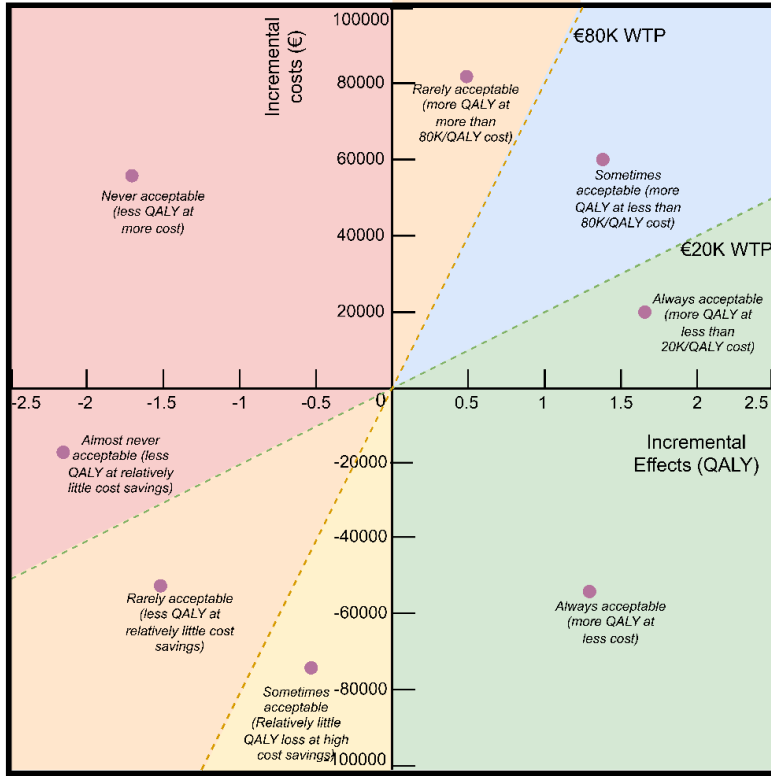


Figure 5: Example of a Cost-Effectiveness plane showing when outcomes would be acceptable

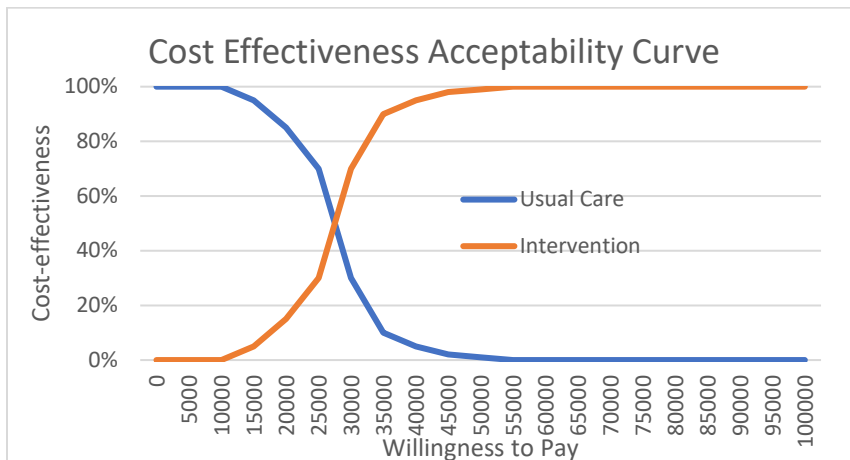


Figure 6: Example of a Cost-effectiveness Acceptability Curve

4.3. Discrete Event Simulation Model

A discrete event simulation model will be built to perform the health economic analysis and determine the desired outcome variables. Discrete event simulation models a system as a sequence of events in time, where each event changes something in the system or simulated individual. A DES model can have a set of trajectories where individual events can influence the followed trajectory of an entity in the system. These possibilities make discrete event simulation great for performing health economic analyses. Patients can be individually modeled with unique traits and attributes that influence the probability of following a specific trajectory. Individual patient modeling is important for the distinction between a DES and a Monte Carlo model. A discrete model has the advantage that states only update at given points in the system, giving a computational advantage over a continuous system, where states are constantly updated.

Furthermore, the advantage of the dynamic aspect in a DES model allows for time-dependent changes in the system, where changes happen based on the system's time. Time dependence is not found in static models. In a Discrete Event Simulation model, the transition into a new stage in the simulation can be based on previous events in the patient's life. Every patient will thus have a unique history that influences the following decisions in their simulated life. The combination with sub-modules that model cancer growth, diagnosis, and staging allow for a simulation that could resemble the actual outcomes of the Dutch breast cancer screening program.

The guidelines for performing a standard health economic analysis will be combined with the guidelines from the RIVM on screening and Oncoguide and the Guideline Database on diagnosis and treatment of breast cancer. The DES model will be built in RStudio using the package 'Simmer' (R Core Team, 2017; RStudio Team, 2020; Ucar et al., 2018). It will generate patients with different cancer risks, dying from cancer, and healthy life expectancies. Screening programs may be introduced at times to observe the effects on total costs accrued, life years gained, and quality of life differences over time.

The model was built in phases that incrementally added detail. The first step was creating a simple model containing patients that die at a certain age. The final step was a complex model with a working screening program and implemented death rates per type of cancer incidence, diagnosis, and received treatments. To calibrate this model, outcomes from the model were compared to the Breast Cancer Figures presented by the IKNL based on Dutch data and other important metrics from the CBS (Centraal Bureau voor de Statistiek, 2022; Integraal Kankercentrum Nederland, 2022). The final model was validated according to the AdViSHE tool for validation of health economic models and by comparing key performance indicators and outcome graphs to the real-world scenario (Vemer et al., 2016).

4.3.1. Model Overview

The discrete event simulation model will simulate individual women and follow them over their entire lifetimes. Patients will be created in the model at birth and will be given individual cancer-related attributes. These attributes will influence if, and when a patient will get cancer, how fast the cancer will grow and how severe it might get. At points in time, the patients will be checked either by themselves or through screening. This will influence the next steps in the patient's lives: should they have a positive indication after screening or self-diagnosis, the patient will move to diagnosis where the cancer's stage is determined and then to the hospital to receive treatment. All steps of the disease, screening, diagnosis, and treatment bring with them certain costs and disutilities to a patient's quality of life. At the end of the simulation, all these costs and utilities will be summed up per patient to arrive at a total cost figure, total

life years and total quality adjusted life years. This model includes the recommendations from the systematic review in 3.6, to include more costs and utility parameters and to follow guidelines on simulation modelling and health economic analyses.

The model in R is split up into different sections. The modeling structure itself is based on six individual building blocks called trajectories. These trajectories each contain one location, where the prescribed steps from breast cancer screening, diagnostics, and treatment guidelines come into play. The building blocks are Initialization, Home, Screening, Diagnostics, Hospital, and Dead.

This patient-level model forms the base of the discrete event simulation. Several cancer-related models have been implemented in the back end of the simulation. These models are the tumor growth model, the cancer detection model, and the cancer staging model. The growth model recreates the naturally occurring growth of a tumor so that each patient has an individual carcinoma. The detection model determines the possibility of clinical and screening detection by creating individual parameters for each patient. The staging model is derived from the literature and, by correlation of size and stages, will return a stage of a stage based on the size of the primary tumor.

The models, combined with various input parameters shown in 4.3.6, result in a discrete event simulation model that simulates patients with breast cancer throughout their lives, with various cancer detection and treatment methods. By running simulations with different screening strategies, the outcomes can be compared to the current baseline. A graphical overview of the entire model is presented in Appendix VII – Full Model Overview.

4.3.2. High level model

The broad overview of the patient-level model and possible patient flows is shown in Figure 7.

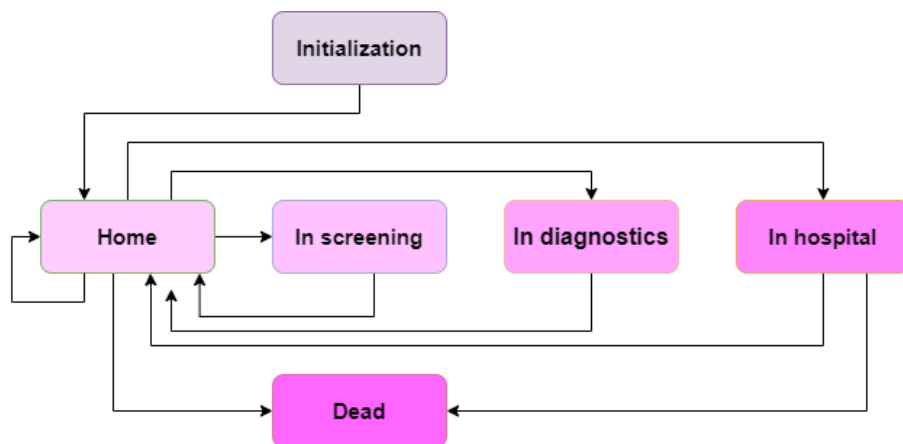


Figure 7: Patient level model overview

This model can be seen as the life of patients. Patients start in 'Initialization' but immediately move to 'Home'. Patients will remain at 'Home' most of their lives. Every month a check takes place to see what a patient's next step will be. When a patient is between 50 and 75 years old, she will get an invitation for screening once every two years. If a patient has been invited, the patient might go to screening for the next step. After screening, the patient will be given a BIRADS score and transferred back to home. At home, patients can also clinically detect a tumor themselves. Depending on this BIRADS score or personal observation, the patient might go to 'Diagnostics', where the clinical severity of possible breast cancer will

be determined. After a couple of hours in diagnostics, the patient will return home to await the results. If cancer is found, the patient must go to 'Hospital' at the next step. In the hospital trajectory, the patient will be treated as well as possible for their type of cancer. This hospital represents the entire treatment stage, including time spent at home between hospital visits to simplify the model. A patient can be cured, can be cured with a new, reduced life expectancy, or she can die in the hospital. If a patient survives, she will return home to continue her life, including other screenings and follow-ups. If a patient does not die due to breast cancer, she will live until her life expectancy runs out. The final 'Death' trajectory will record the patient's final attributes. In 4.3.6 and Appendix VIII – Full code explanation, more information on the specific actions per trajectory is given.

4.3.3. Tumor growth model

The tumor growth model is based on literature. As evaluated through the systematic review of Chapter 3, previous studies often used tumor growth models based on a simple formula, with little room for spontaneous regression or early growth stagnation. This model tries to combine multiple tumor growth trajectories into one formula, where elements of randomness cause patient-level variations of growth rate, maximum sizes, and regression probabilities. In Figure 8, the combination of the varying model elements can be seen. In these figures, the size of the tumor can be seen on the y-axis, with time passing on the x-axis.

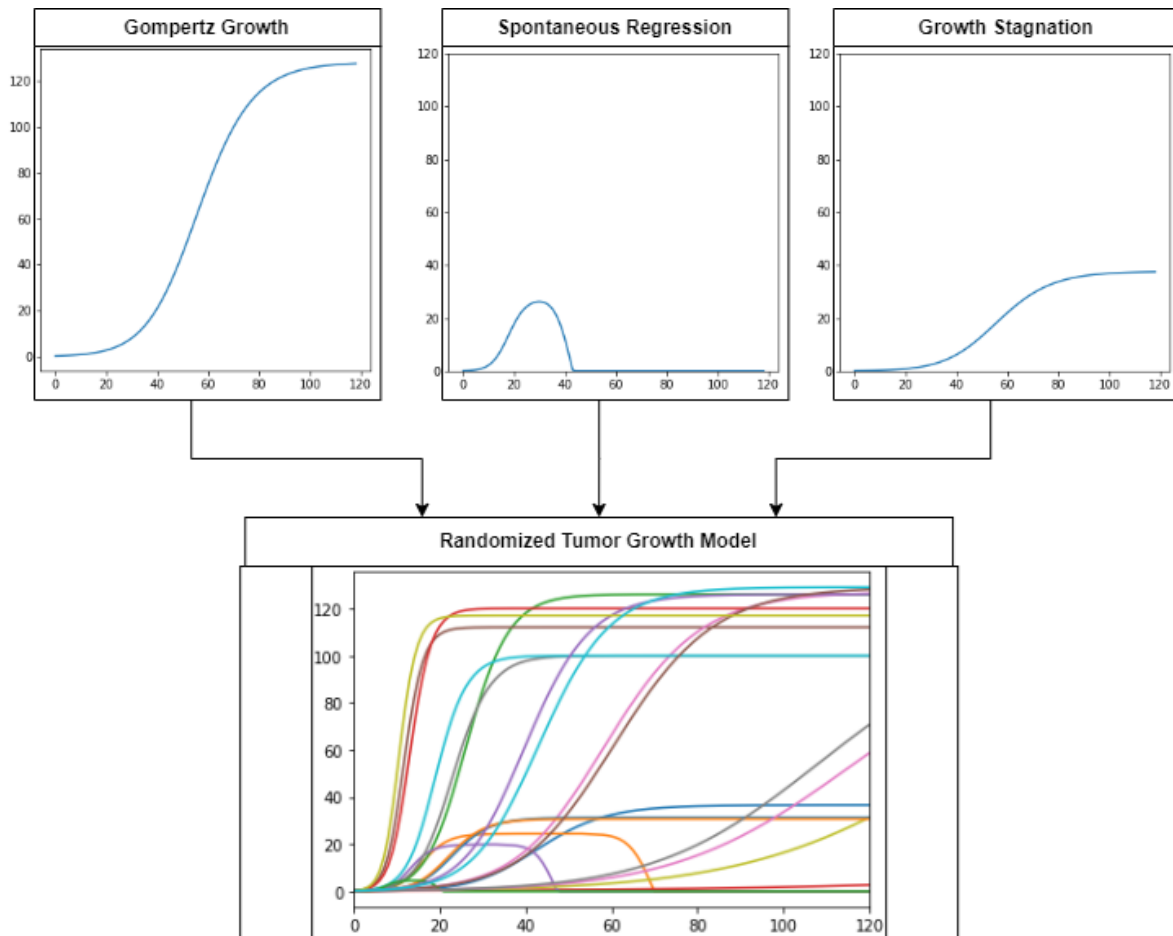
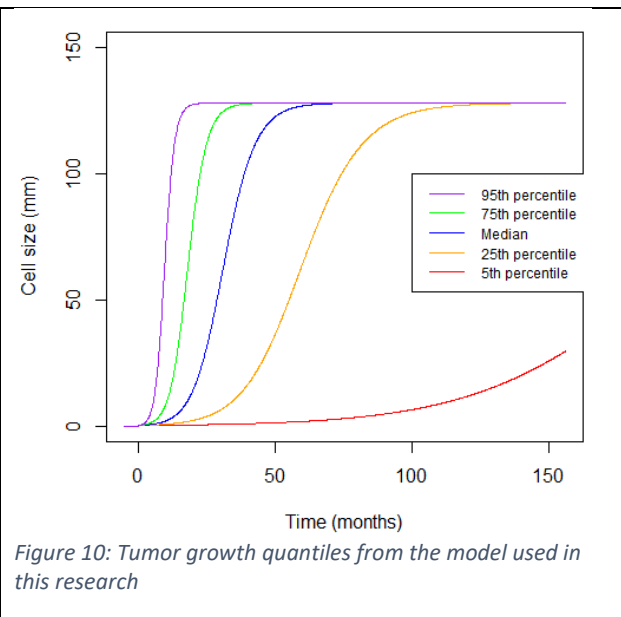
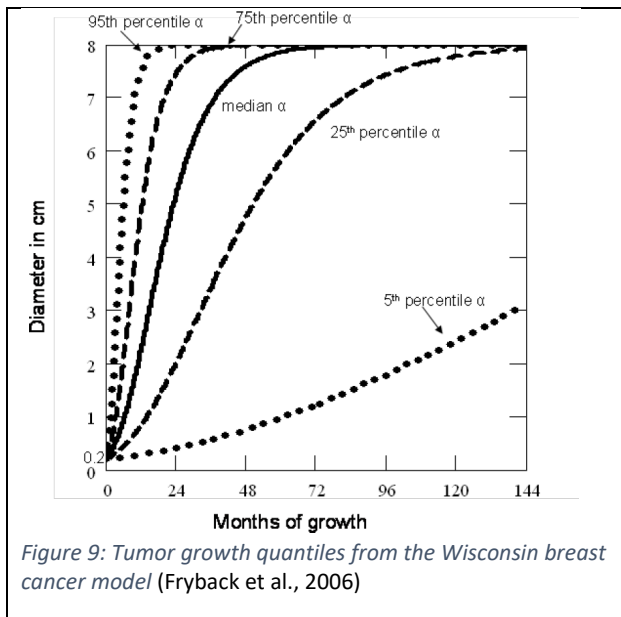


Figure 8: Tumor Growth Model combining Gompertz Growth, Regression and Stagnation

Various factors connect the patient-level model to the tumor growth model. The tumor growth model depends on the factors 'time,' 'growth rate,' and 'recession onset.' The model follows an initial Gompertz curve, allowing the tumor to grow to a maximum diameter. This maximum diameter is set per patient at around 128 mm. The randomly drawn regression or stagnation onset, which affects a portion of the patients, allows the tumor to go into regression or stagnate at a certain point in time.

This model allows for the implementation of distinct tumor types, combining aggressively growing, slow growing, docile, stagnating, and even regressive tumors into one model. The model only needs four patient attributes to retrieve a patient's tumor size at the current time in the patient-level model.

The tumor model was validated by comparing it with similarly defined models in the CISNET model repository. The Wisconsin model showed the graph in Figure 9 of the growth of tumors over time. Their model includes no natural regression or stagnation, but the growth rates of the model could be compared. In Figure 10 on the right, the growth quantiles of the tumor growth model of this research are compared to those of the Wisconsin model on the left (Fryback et al., 2006; National Cancer Institute, 2022).



4.3.4. Cancer detection model

Screening and clinical detection happen by the patient or physician noticing the primary tumor size. In general, screening detects a tumor at a smaller size than a patient would notice something is wrong herself. Detection sizes are not deterministic but vary per patient. These variations prompt the need for a cancer detection model. Inspired by the MiSCaN model (van den Broek et al., 2018), a method for cancer detection through screening or clinical diagnosis was devised. An illustration of this model can be seen in Figure 11. The model in this illustration is a theoretical explanation. Therefore, no axis labels are given, and the distributions differ from the actual model.

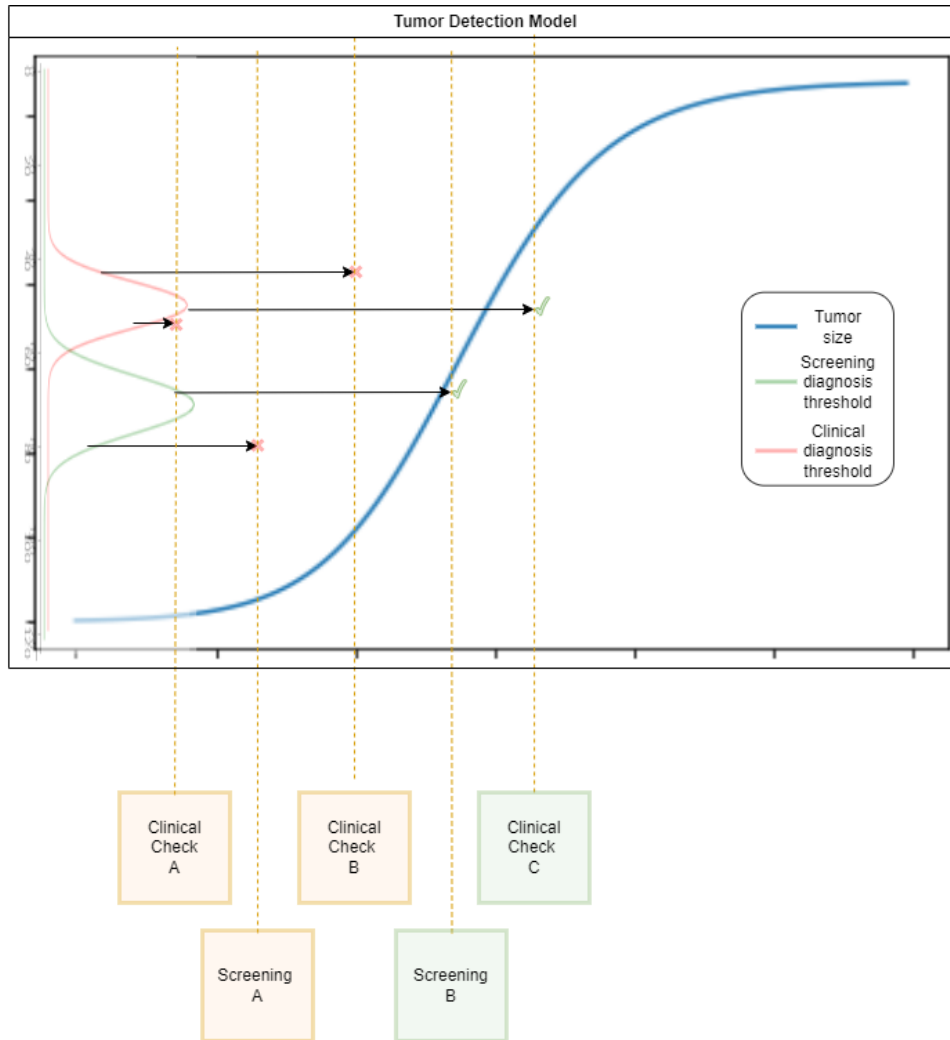


Figure 11: Tumor detection through screening and clinical diagnosis.

According to the randomized tumor growth model, the probability of detection increases as a tumor grows. There are two types of detection possible: clinical detection, where the patient detects a lump in her breast herself and is then sent to diagnostics, and screening detection, where a sizeable suspicious lump has been found through screening. The thresholds of these detections are drawn from statistical distributions: the mean of the screen-detection distribution is drawn from a Weibull distribution and is, on average, lower than that of the clinical detection threshold drawn from a lognormal distribution. In most cases, the threshold for clinical detection is higher than that of screening detection.

In Figure 11, this process is visually explained. The main graph in blue is a patient's unbridled tumor progression. At various points in time, clinical and screening checks have taken place. For each check, a random value is drawn from the distributions of the respective check. This value represents the detection threshold. If the tumor size is above that threshold at that time, the tumor is detected. If the threshold is too high, the tumor will continue growing according to the growth model until the next checkup.

If the cancer is detected, the patient will be referred to diagnosis, where the Cancer staging model will determine what stage the cancer is in.

4.3.5. Cancer staging model

Literature shows a link between the size of the primary tumor and the TNM stage. There are no direct thresholds at which stage cancer enters the next stage. However, based on the Synthetic dataset from the IKNL (Integraal Kankercentrum Nederland (IKNL), 2022), a visualization can be created to show the correlation between primary tumor size and tumor stage. The staging model is visualized in Figure 12.



Figure 12: Relation between cancer stage and primary tumor size (Integraal Kankercentrum Nederland (IKNL), 2022)

Based on this data, the patients in the simulation get assigned a tumor stage at the time of diagnosis. This tumor stage is a combination of the size of the primary tumor and a random variable. If the tumor size is 12mm, and the random variable from a $U[0,1]$ distribution is 0.9, this cancer will be classified as Stage II.

The choice to use the relation between size and stage is made because the size of the primary tumor is the only parameter initially measured when screening and is the leading indicator in clinical diagnoses. More factors affect a cancer's stage, including hormone status and breast density. Including these variables is deemed outside of the scope of this research on screening effectiveness.

4.3.6. Patient Level Model

The three cancer models are activated in different sub-trajectories in the patient-level model. The sub-trajectories are explained in more detail. For an explanation in great detail, Appendix VII contains the full explanation of the code of all models and trajectories. This chapter goes more in-depth on what happens in the various sub-trajectories and links the earlier described models in 4.3 to explain the simulation model.

Initialization

In initialization, the initial patient parameters are set before the patient is moved to Home. The parameters are derived from literature and open data. This data is transformed into probability distributions, from which every patient can draw a unique value. In Table 10, different patient-specific parameters are explained. In Appendix VII – Full code explanation, the parameters and functions that initialize the values can be found in more detail.

Table 10: Initial patient attributes

Parameter	Explanation
<i>Alive</i>	Keeps track of whether a patient is alive. 1 for alive patients, 0 when patients die.
<i>PatientAge</i>	Keeps track of a patient's current age in hours. Initialized at 0 with a maximum value of 100 years, or 876,000 hours.
<i>HealthyLifeExpectancy</i>	Life Expectancy of a patient if she does not get cancer.
<i>CurrentUtility</i>	Stores patient's current utility. Lower during screening and diagnostic procedures and is kept a lower value after diagnosis and treatment.
<i>TumorSize</i>	Current size (diameter in mm) of tumor in the patient. Initialized at 0, with a maximum value of around 128.
<i>TumorStartAge</i>	Age when tumor starts growing, in hours
<i>TumorGrowthRate</i>	Patient-specific tumor growth rate to use in the tumor growth model
<i>Invite</i>	Notes if a patient has an active invitation for screening. 0 if there is no invitation, 1 if there is one.
<i>BIRADS</i>	Patient's BIRADS score after screening
<i>Referral</i>	If the patient has an active referral for diagnosis
<i>NextStep</i>	Next step for the patient to visit in simulation, 1 sends a patient to Home, 2 to screening, 3 to diagnostics, 4 to the hospital, and 5 to death.
<i>TotalCosts</i>	Records total patient-incurred costs, both medical and societal
<i>MedicalCosts</i>	Records patient's medical costs related to screening, diagnosis, and treatment
<i>SocietalCosts</i>	Records patient's societal costs like economic burden and driving costs
<i>TotalUtility</i>	Records patient total Quality Adjusted Life Years
<i>WorstCancer</i>	Stores the worst cancer size a patient has experienced
<i>WorstStage</i>	Stores the worst cancer stage a patient has experienced
<i>IsCured</i>	If and how often the patient has been cured of cancer
<i>Screened</i>	Counts how often the patient has been screened

For *PatientAge*, a maximum age of 100 years was used. This maximum is standard in health economic analyses and advised through the Dutch guideline (Zorginstituut Nederland, 2016). For *TumorSize*, a maximum value was drawn from a normal distribution with a mean of 128 and a standard deviation of 10. The MISCAN model and others in the CISNET repository use this maximum size. (National Cancer Institute, 2022; Tan et al., 2006; van den Broek et al., 2018). The *TumorStartAge* was based on data from the IKNL synthetic dataset (Integraal Kankercentrum Nederland (IKNL), 2022). The *TumorGrowthRate* was explained in the tumor growth model and calibrated based on data from the Wisconsin model (Fryback et al., 2006).

In order to generate a combination of a life expectancy and a cancer onset age, a copula was derived from the distributions of cancer incidence ages and life expectancies. A copula is a multivariate distribution that evaluates the correlation between different variables. Copulas allow for the combination of multiple distributions into a single distribution, from which two related values can be picked. It is used to model the dependence of two variables. Figure 13 shows the copula in a density plot. By drawing two variables, healthy life expectancy, and cancer onset age, from the copula, the cancer onset ages are adjusted for the life expectancy.

The book *Simulation Modeling and Analysis*, written by Averill Law, proposes multiple ideas for generating random variables to be used in simulations. Conditional distributions would have been possible but need a level of detail that is often practically unobtainable. Multivariate normal and lognormal distributions are possible but prove problematic if variables need to be drawn from different distributions. Correlated random variates through copulas are possible, but there is a limitation on the set of returning values that are possible in reality. Combining a copula with constraints is the best option for obtaining these random variables (Law et al., 2015).

In Figure 13, The copula of cancer incidence age and life expectancy can be seen. The main plot shows the density of the combined distributions, both in points and an overlaid density plot. The plots on the side show the distributions for life expectancy and cancer incidence age. The distribution of the cancer onset age on the y-axis looks like a normal curve. However, fewer people are alive at age 75 than 45, even though the normal curve shows no difference in the probability of cancer onset. The probability of having cancer at age 75 is higher than at age 45. This is compensated for by combining the life expectancy and cancer onset age distributions in a copula. The density plot shows a more intense area at the 75-year point than the 45-year point on the y-axis.



Figure 13: Copula of cancer incidence age and life expectancy

Besides the advantage of using a copula for these specific patient parameters, it shows the flexibility for extending the model in future iterations. This shows a means for drawing random correlated variables, which is vital if the model is upgraded to a version allowing risk stratification. In a possible future version, extra parameters for the probability of cancer incidence could be linked to breast density, gene expression, or other factors that could influence cancer risk. This model has not implemented those possibilities, as that is outside the scope of this study.

Home

After the initialization, patients are transferred to 'Home.' 'Home' is where patients spend most of their lives. Every month, a check is performed to update patients' attributes and see if patients need to go elsewhere. This update consists of the following:

1. Patient's EntryAge is recorded to calculate later how long a patient has spent in this state.
2. Determined how long the patient will remain at home this iteration.
3. Patient's age is updated by duration from step 2.
4. Patient's age is checked to assess eligibility for screening (if between 50-75, once every two years). Attribute 'Invite' is updated.
5. Patient checks herself, to see if she notices the (possible) tumor herself by means of the Tumor Detection Model.
6. The patient's next step is determined, based on the following attributes:
 - a. Invite & screened, to see if the patient goes to screening next.
 - b. BIRADS & ClinicalCheck, to see if the patient goes to diagnostics next.
 - c. Referral, to see if the patient is referred to the hospital.
 - d. PatientAge & HealthyLifeExpectancy, to see if the patient dies due to age.
7. The patient's TumorSize is updated according to the Tumor Growth Model.
8. The 'WorstCancer' attribute is checked and updated for monitoring purposes.
9. The Patient's CurrentUtility is determined based on the patient's age and disease status.
10. The current utility is multiplied with the duration of this step and added to the TotalUtility for monitoring of QALYs.

After this update, the patient leaves home to see where she should go at this step based on the NextStep attribute. The patient will often return to Home, but she can also go to any of the other trajectories from here.

Screening

A patient is invited to the screening program approximately 13 times. An invitation will be sent to the patient every two years between the ages of 50 and 75. Based on data from Bevolkingsonderzoek Nederland and the University of Manchester, not all patients follow up on this invitation (Bevolkingsonderzoek Nederland, 2022; Wright et al., 2022). These sources suggest the following parameters:

Screening Round	Uptake probability
First Screening	0.78
Any screening after the first completed	0.90
Any screening after first if never screened before	0.25

Whether a patient will go to their screening appointment is determined in step 6 at home. In the screening trajectory, the following steps happen:

1. Patient's EntryAge is recorded to calculate later how long a patient has spent in this state.
2. Determined how long the patient would remain at the screening station.
3. Patient's age is updated by duration from step 2.
4. The patient's 'TotalTimeAtScreening' attribute is updated by adding the current duration of this screening visit.

5. The patient's current utility is determined based on the previous utility and a decrement for the discomfort experienced during screening.
6. The current utility is multiplied with the duration of this step and added to the TotalUtility for monitoring of QALYs.
7. The SocietalCosts attribute is updated by societal costs incurred in this step.
8. The MedicalCosts attribute is updated by medical costs incurred in this step.
9. The tumor is detected based on the Tumor Detection Model. The detection likelihood is reflected in the BIRADS score attribute.
10. 1 is added to the 'Screened' attribute to reflect the total screening visits the patient has had.
11. The patient's NextStep attribute is set to 1, which means the patient will be sent home after the screening visit.

Diagnostics

The patient can end up in diagnosis in two different ways. One is through clinical diagnosis, where the cancer detection model used in 'Home' resulted in the detection of the tumor by the patient herself. The other is if screening returned a BIRADS of 3, 4, or 5, as discussed in 0. In Diagnostics, the actual stage of the tumor is determined based on the tumor staging model, and the patient is referred to the hospital. The following steps are taken in diagnostics:

1. Patient's EntryAge is recorded to calculate later how long a patient has spent in this state.
2. Determined how long the patient would remain at the Diagnostics station.
3. Patient's age is updated by duration from step 2.
4. The patients 'TotalTimeAtDiagnostics' attribute is updated by adding the current duration of this visit.
5. The patient's current utility is determined based on the previous utility and a decrement for the discomfort experienced during diagnostics.
6. The current utility is multiplied with the duration of this step and added to the TotalUtility for monitoring of QALYs.
7. The SocietalCosts attribute is updated by societal costs incurred in this step.
8. The MedicalCosts attribute is updated by medical costs incurred in this step.
9. The attribute 'DiagnosedThrough' is set to either 1 or 0, based on if the patient's tumor has been discovered through screening or clinically.
10. The tumor's stage is determined based on the tumor size and the tumor staging model.
11. The WorstStage is recorded for monitoring purposes.
12. The DiagnosticsVisits attribute is incremented by 1 to reflect the patient's total visits to this stage.
13. The patient's NextStep is set to send the patient Home or to the Hospital, depending on the need for immediate treatment.

Hospital

Patients with cancer need to be treated to guarantee their best possibility of survival. For this analysis, it is assumed that all patients undergo the best possible treatment they can get. Based on this treatment, a new life expectancy will be determined.

This tool has been translated to R and optimized for use with the available parameters to be used in the Hospital part of the model. This tool will output the life expectancy after treatment.

The following steps are taken in the 'Hospital' trajectory:

1. Patient's EntryAge is recorded to calculate later how long a patient has spent in this state.
2. Determined how long the patient would remain at the hospital.
3. Patient's age is updated by duration from step 2.
4. The patient's 'TotalTimeAtHospital' attribute is updated by adding the current duration of this screening visit.
5. The patient's current utility is determined based on the previous utility and a decrement for the discomfort experienced during screening.
6. The current utility is multiplied with the duration of this step and added to the TotalUtility for monitoring of QALYs.
7. The SocietalCosts attribute is updated by societal costs incurred in this step.
8. The MedicalCosts attribute is updated by medical costs incurred in this step.
9. The patient's prognosis is calculated according to PredictV2.0.
10. The patient's cancer-related attributes are adjusted: BIRADS is reset, IsCured is incremented, and TumorSize is updated.
11. HospitalVisits is incremented by one.
12. NextStep is set to either home or dead, depending on the result of step 9

Death

A patient is moved to 'Death' if she is going to die. In Dead, the patient is taken out of the simulation. In this step, the costs for screenings, diagnostics and invites per patient are added to the medical costs. The costs for unrelated diseases in gained life years are calculated and stored in the PAIDCosts variable, and the societal costs are also calculated based on the number of screenings, diagnostics and hospital visits a patient has had. Finally, the final state of all attributes is recorded in a large data frame that covers all patient's attributes. This data frame is later used to create insightful results information.

4.3.7. Utility and Costs

More input data is used in the model than shown under 4.3 so far. All variables can be found in Appendix VII, and important variables and calculations related to costs and utilities are shown here.

Utility

Patients' base utility, or Quality of Life, varies over time. A person living to 100 years old will not experience all years at 100% quality. Instead, the base utility is associated with the age of the patient and will become less as patients are older. The base values used are shown in Table 11. They are derived from the Manchester model (Wright et al., 2022).

Table 11: Base utility values per age used in the model (Wright et al., 2022)

Index	Age	Utility
0	<30	1
1	30	0.9383
2	35	0.9145
3	40	0.9069
4	45	0.8824
5	50	0.8639
6	55	0.8344
7	60	0.8222
8	65	0.8072
9	70	0.8041
10	75	0.7790
11	80	0.7533
12	85	0.6985
13	90	0.6497
14	95	0.6497
15	100	0.6497

If a person experiences cancer, the utility will become lower. Per stage, this utility is brought down by a certain percentage. These percentages are derived from the study by Arrospeide, which was included in the systematic review (Arrospeide et al., 2016). The utility decrements per stage are shown in Table 12

Table 12: Percentage at which the base utility is lowered per stage (Arrospeide et al., 2016)

Stage	Decrement percentage
Healthy	0%
DCIS	10%
I	15%
II	20%
III	25%
IV	40%

Costs

In this analysis, two types of costs are calculated throughout the simulation, medical and societal. Medical costs are incurred for disease treatment, during diagnosis and screening. Societal costs are costs related to costs of productivity loss, visiting the various locations, and future disease costs. The societal costs are taken from the Guidelines, and the medical costs are taken from the literature. All costs are discounted at 4% annually.

This simulation simulates all patients from birth to death. To accurately represent the costs, the year at which the tumor starts growing is taken as the base year. This differs for each patient but represents the

costs as if the tumor were found today. All other costs are discounted for this start date. The costs used in the model are given in Table 13.

Table 13: Costs for screening and diagnostics

Cost type	Cost	Source
<i>Screening invite</i>	Drawn from normal distribution with mean €5 and SD €2, with a minimum of €2.	BVO NL Annual Report, checked with Rafia (PriceWaterhouseCoopers Accountants N.V., 2022; Rafia et al., 2016)
<i>Screening procedure</i>	Drawn from normal distribution with mean €100 and SD €5	BVO NL Jaarverslag, checked with Rafia (PriceWaterhouseCoopers Accountants N.V., 2022; Rafia et al., 2016)
<i>Diagnostic procedure</i>	Drawn from normal distribution with mean €212 and SD €5	From one of the papers of the systematic review, converted to 2023 euros (Rafia et al., 2016)

The costs for screening are given in the table below. These costs are taken from the article from van Luijt, included in the systematic review and converted to rounded 2023 euros (van Luijt et al., 2017). Costs are based on the cancer stage and split up into three sections. First, the costs for the initial treatment in the first year are given per month. Then, the costs for the next months are given, and finally, the treatment costs for the final six months of a patient's year are given. The costs are drawn from a random normal distribution, with a mean and a standard deviation given in Table 14.

Table 14: Treatment costs per stage and time point (van Luijt et al., 2017)

Stage	Initial 12 months cost	Initial SD	Continuous Monthly Costs	Continuous SD	Terminal 6 months Costs	Terminal SD
0	0	0	0	0	0	0
DCIS	70643	3000	1035	102	174504	28000
1	106869	2300	1644	84	133712	14000
2	214542	4800	3125	133	768847	9000
3	263578	16000	4348	721	138925	21000
4	247895	26000	9112	1158	182511	21000

All these costs are discounted at 4% annually. Costs incurred a long time after the inception of the tumor, have less effect on the final medical costs than the first treatment costs. Discounting happens through the same formula as for utility.

Societal costs were newly introduced by the guideline in 2014. These costs consist of other burdens to the patient or society. For example, the guidelines give base values for calculating travel costs to and from various types of clinics. Furthermore, the loss of productivity should be calculated whenever someone has to leave work for a visit to a medical facility. The third value calculated in this research is the cost of extra

life. Whenever a patient is treated so that she may live longer, the additional medical costs incurred in these gained life years should be calculated and included in the final cost of the intervention. This is a highly debated point in health economics, but the Dutch guidelines now include these costs in their guide. In the simulation, these costs are split into societal and PAID costs. The PAID costs include the costs for living additional years after treatment and are derived from the PAID tool (*The IMTA PAID Tool Version 3.0*, n.d.; Van Baal et al., 2011).

Discounting

Utility and costs will need to be discounted with a discount rate. For patient utility, the discount rate is 1.5%, and for costs it is 4% per annum. This discounted utility value is calculated because there is an expectancy that future disease burden will be less than it is currently. This is due to the expectation of better treatment possibilities in the future and the expectation that disease burden will be experienced as less severe as time goes on. The utility discount will only be applied to the decrement percentage: this way the base utility remains unaffected and stable over the different experiments. The formula for discounting utility is shown in Equation 2 (Parkinson & De Abreu Lourenco, 2015).

$$\text{Discounted Utility} = \frac{U_t}{(1 + r)^t}$$

Equation 2

Where t is the year, U is the utility decrement percentage, and r is the discount rate.

To clarify, an example is given. Patient X is diagnosed with stage II cancer at age 44. Table 15 shows the steps taken to calculate the utility over the next ten years.

Table 15: Example of utility discounting

Age	Base Utility	Decrement percentage	Discounted decrement	Final utility
44	0.9069	20	20	0,7255
45	0.8824	20	19.7931	0,7077
46	0.8824	20	19.6731	0,7088
47	0.8824	20	19.5884	0,7096
48	0.8824	20	19.5230	0,7101
49	0.8824	20	19.4696	0,7106
50	0.8639	20	19.4247	0,6961
51	0.8639	20	19.3858	0,6964
52	0.8639	20	19.3516	0,6967
53	0.8639	20	19.3210	0,6970
Total	8.77	-	19.55%	7.06

A patient living under normal circumstances will experience around 8.77 quality-adjusted life years in the ten years from 44 to 53. A patient with stage II cancer will only experience this as 7.06 quality-adjusted life years, with an average of 19.55% less utility experienced in these years compared to the healthy person.

4.4. Model Calibration and Validation

After a simulation run, a final data frame that includes all patient's final attributes is returned. This data can be used to gain insights into the simulation's performance. There are two purposes for this data: the first is to generate graphs that can be used for finetuning the model. The second is to calculate outcome data as described in 4.2, which will be used in the results of this study.

4.4.1. Calibration

A program that uses the data outputted at the end of the simulation has been written to create a set of graphs and KPIs. These graphs and KPIs contain key information that is needed to determine if the simulation shows a realistic image of the real world. The graphs and KPIs allow quick interpretation of the results and show where finetuning is needed. As testing is done with limited simulation sizes, confidence intervals are added to the KPIs to determine if they could be accurate and if more testing is needed. The KPIs used for calibration can be seen in Table 16, including their actual reference values.

Table 16: KPIs used for calibration of the model

<i>KPI</i>	Reference	Simulation (1M patients)
<i>Mean age of death</i>	83.0 (Centraal Bureau voor de Statistiek, 2022)	82.0
<i>Breast cancer incidence percentage</i>	14.3% (Integraal Kankercentrum Nederland, 2022)	13.9%
<i>Percentage detection through screening</i>	56.0% (Integraal Kankercentrum Nederland, 2022)	55.1%

Two sets of four graphs are produced. The first show the final age distribution, a histogram of cancer incidence ages, the numerical stage distribution over the different methods of diagnoses and the various tumor growth rates. The life expectancy and cancer incidence ages were compared to actual figures from CBS. The findings per diagnosis were compared to the actual figures published by IKNL, and the tumor growth rates give insight into where finetuning in tumor growth might be needed. These graphs are shown in Figure 14.

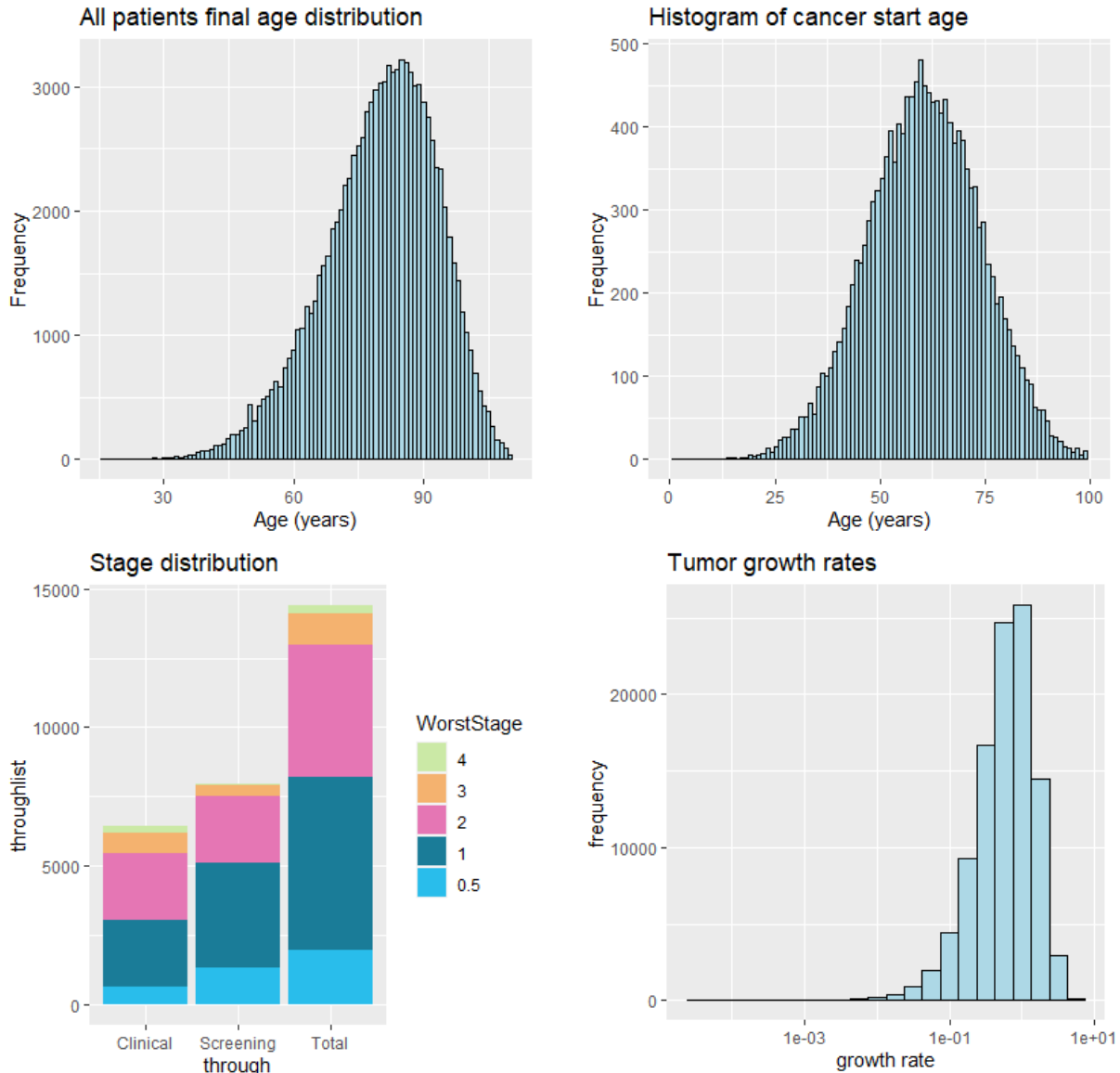


Figure 14: First set of graphs from simulation output

The second set of graphs (Table 17) shows the same statistics as the graphs IKNL presents on its website (Integraal Kankercentrum Nederland, 2022). There are four important graphs: the percentage of tumors discovered through the screening program split up per stage, the age at diagnosis per stage, the stage breakdown per means of diagnosis and the relative survival per diagnosed stage. The simulation model will be accurate if the graphs obtained through simulation are comparable to those from the IKNL. After calibration of the model, the model can be validated through the Assessment of the Validation Status of Health-Economic decision models tool. Table 17 shows figures 15, 16, 17, and 18. The graphs from the IKNL are on the left, and those from the model report. As all values resembled those from reality, the model is now considered calibrated and ready for validation.

Table 17: Comparing IKNL figures to results from the model calibration.

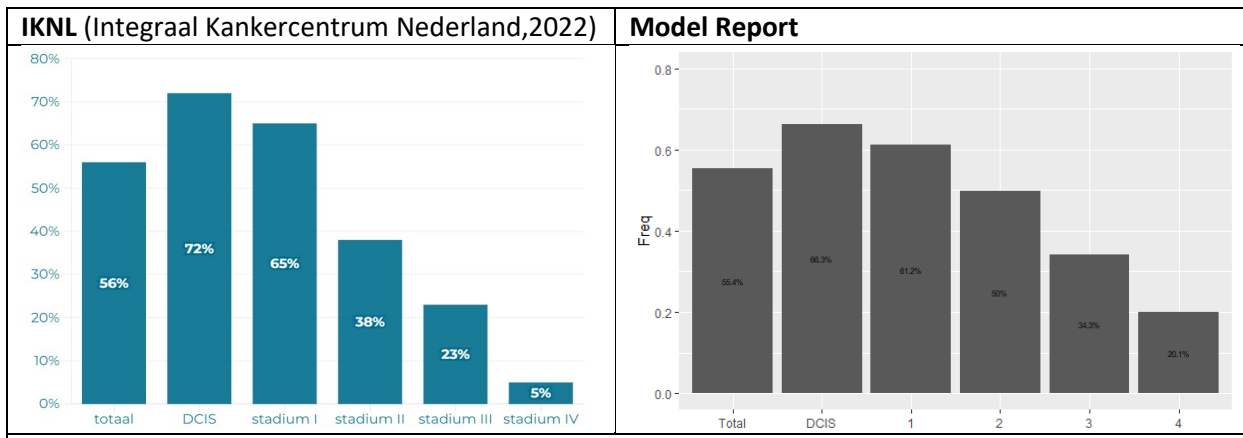


Figure 15: Percentage of tumors discovered through the screening program

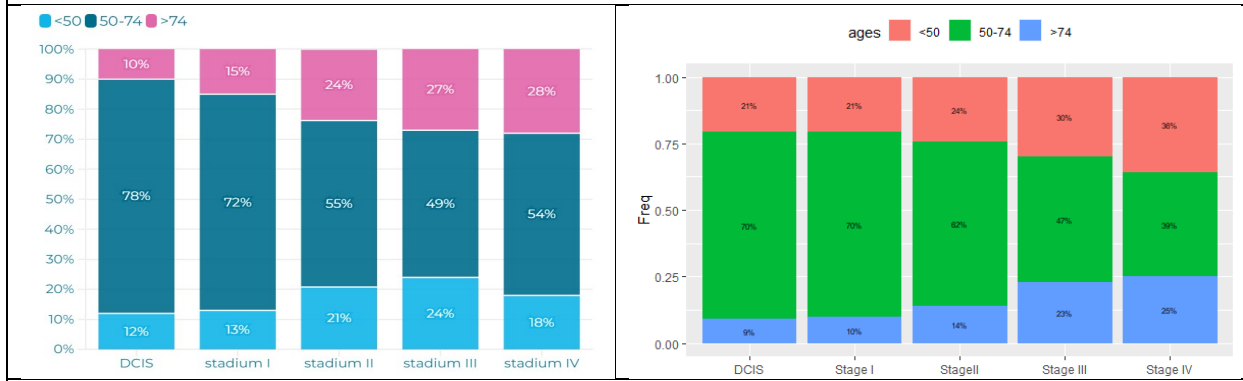


Figure 16: Age at diagnosis

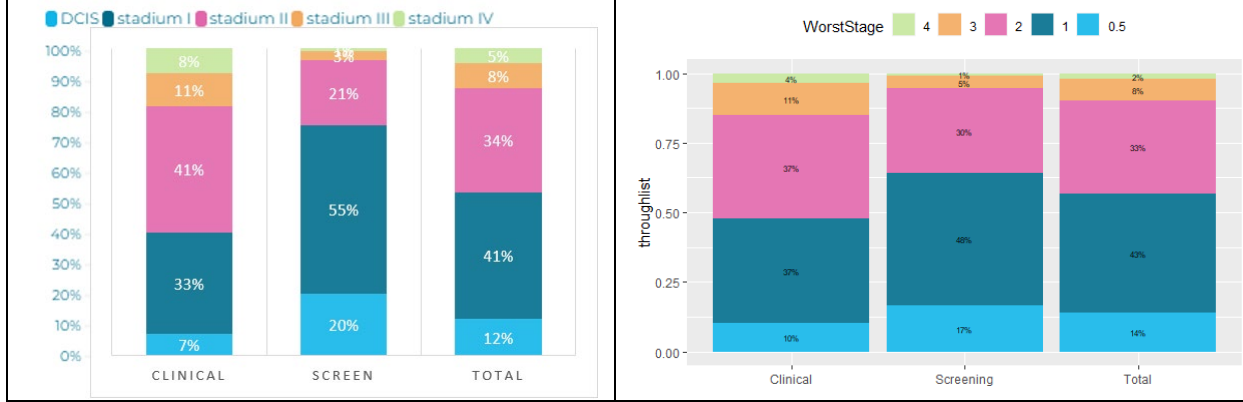


Figure 17: Stage Distribution per modality

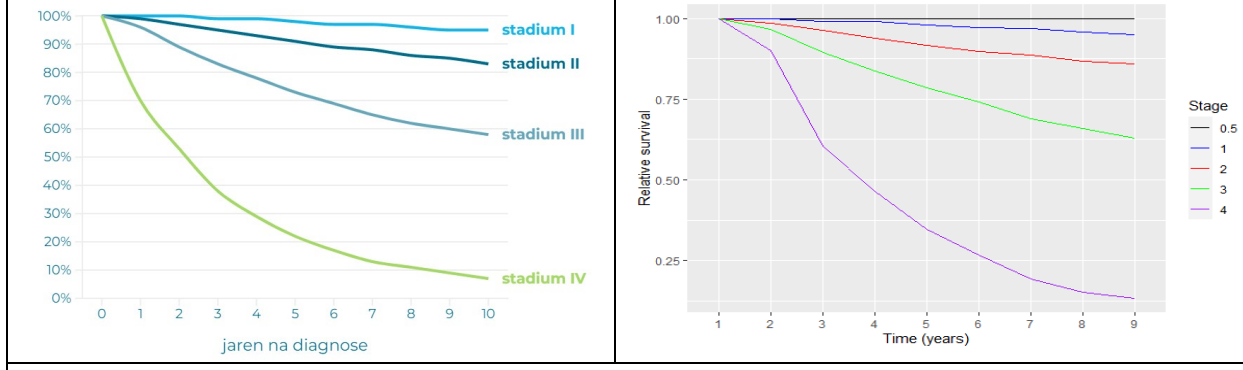


Figure 18: Relative survival per stage

4.4.2. Validation

The guidelines advise using the Assessment of the Validation Status of Health Economic decision models (AdViSHE) questionnaire to evaluate a health economic model. AdViSHE does not consider itself a validation tool, but combined with the resemblance of real-world data and the filled-in questionnaire, the model can be considered valid if all can be answered positively (Vemer et al., 2016; Zorginstituut Nederland, 2016).

AdViSHE is a questionnaire that can be used to improve the validation status of a model. It consists of five parts: the validation of the conceptual model, of input data, the computerized model, operational validation, and other validation techniques. The filled-in questionnaire can be found in Appendix VIII. AdViSHE is not a validation tool in itself, but if all parts of AdViSHE can be answered positively, and the output data from the model, as shown in Chapter 4.5, realistically resemble the actual data, the model can be considered valid.

Part A checks the validity of the conceptual model. The conceptual model describes the underlying system using logical, graphical, and verbal representation, which for this model is described in Chapter 4.3. Experts have been consulted to judge the appropriateness of this conceptual model and the concept has been compared to other conceptual models found in the literature.

The next part validates the input data. AdViSHE asks for justification and description of aspects of the input data. The search strategy of gathering data is questioned, which for this review was gathered through papers included in the systematic review. The data sources and their reason for inclusion is given, the assumptions on included or excluded data is given, and the use of distributions and parameters is explained.

The validation of the computerized model happens in part C of the AdViSHE checklist. It discusses the differences between the conceptual and computerized model, as well as the way the computerized model has been implemented in a software program. The conceptual model is an accurate representation of the actual computerized model, with no significant changes between the two. The model has been subject to extreme value testing, as seen in the PSA experiments. Unit tests have taken place to assess the validity of the submodules. Patient tracing has taken place to follow patients through the simulation.

Part D discusses techniques used to validate the model outcomes. The model outcomes have been judged and compared to the outcomes of other models. The medical outcomes concerning stage shifts are comparable; the average quality of life and total costs differed from other models as this model includes more data. The PSA also tested the validity with alternative input data, to see how much the model's performance would differ if actual data was different from the baseline data. The model has also been validated against empirical data, which it is able to accurately represent. This is described in Chapter 4.4.1.

The final part of the checklist, part E, checks if other validation techniques have been performed. During the implementation, calibration and validation phase, this model has been checked by performing patient tracing with individual walkthroughs. This was performed to check if the conceptual model was accurately represented in the computerized model. Many parts of the model were double programmed, with one version in python and one version in R. Double programming was performed to see if similar results could be achieved using different software packages and to test smaller parts of modules before

implementation in the full model. Naïve benchmarking was also performed at many times to create estimates of expected change before changing a parameter. This was performed through excel calculations to estimate the impact of changing a cost variable before PSA testing.

Validity is important for the interpretation of the results of the experiments. Various strategies can now be tested with the same input data used for validation. The only thing that will change is an external factor, the screening strategy. The logical deduction says that changing this will not change patient attributes, like the probability of getting cancer and the healthy life expectancy of patients. This means that results obtained through experimentation with a valid model can be seen as an accurate representation of reality with different screening strategies. This also answers research question 2.3: a simulation model has been created to compare screening strategies.

4.5. Experiment Design

Experimentation will consist of four distinct phases. The first will be creating a base result for various screening strategies. This base result will be on one large simulation, where all strategies use the same random numbers for patient parameters in order to arrive at a result. The second phase will be individual probabilistic sensitivity analyses of the base model to test the model's stability. The next phase will consist of experimentation with five strategies, which will be compared to the baseline and have their results analyzed through a CE-plane and CEAC. The final phase will consist of testing 256 screening strategies to see if there are strategies more advantageous than the one in use right now. Finally, a budget impact assessment of the screening program on the Dutch healthcare system will be made.

4.5.1. Base Model Testing

The first results will contain the information of a full model run under standard conditions. This will be a simulation run with 1M patients. This baseline can then be used as a baseline for the PSA and as a baseline for calculating the ICER of new strategies. The KPIs and output graphs of this baseline simulation are presented in the validation and the results sections.

4.5.2. PSA Experiments

The base-case analysis will represent patient-level variety by the implementation of different attributes per patient. This stochastic uncertainty is accounted for in the first base experiment. The Probabilistic Sensitivity Analysis experiments will test the same strategy as the baseline experiment, but in smaller simulation sizes and with more variation on the input parameters. PSA testing aims to test the effects of increased input variability, called second-order uncertainty. It is important to know the outcomes of the experiments if the input values would vary in reality from the projected theoretical values. If small deviations from the projected inputs would result in drastically different cost and effect outcomes, it might not be a good idea to implement policy changes based solely on one experiment. If results change as expected, the results are more acceptable and policy changes could be desirable (Koffijberg, 2021).

To address the second-order uncertainty, the same base experiment is run, but now with the following changes:

- The experiment size will be limited to 100k patients per run.
- All factors will be changed one at a time against the base case, leading to a total of ten experiments.

- An ICER compared to the baseline experiment and the difference in average costs and QALY will be calculated.
- The seed for the experiments will be the same for each experiment.

Table 18: PSA input variables

Variable	Base values	PSA Low Value	PSA High Value
Screening Disutility	0.15	0	0.5
Diagnosis Disutility	0.25	0	0.5
Screening Uptake	0.78	0.5	1
False Positive %	0.5%	0%	2.5%
Treatment Costs	1x	0.25x	3x
Societal Costs	1x	0x	3x
PAID Costs	1x	0x	3x

The results of these experiments will be listed in a tornado graph, with the variable on the y-axis and the effect on the costs and utility on the x-axis.

4.5.3. Strategy Experiments

The experiments will be conducted once with a large patient size in the base experimentation. All experiments will start with the same initial seed to ensure that the starting input values for patients in each experiment are the same. The experiments will provide a single output figure for the necessary KPIs. The experiments will each test a unique screening strategy. The strategies will be tested with a sample size of 100,000 patients divided into ten batches. The various experiments are listed in Table 19.

Table 19: Experimental screening strategies

Experiment	Screening ages	Strategy	Note
B0	50-75	Bi-annual	Actual situation
B1	None	-	No screening at all
B2	50-75	Annual	Increased screening
B3	50-75	Tri-annual	Decreased screening
B4	40-80	Bi-annual	ICER of adding age groups
B5	55-70	Bi-annual	ICER of less age groups

4.5.4. 2k factorial strategy testing

Even though a couple of strategies have been tested to see what effects changing the current strategy might have, it is hard to call the better-performing strategies optimal. They only test one age group at a time, which might not be practical to realize. To test larger and more strategies, a 2K factorial experiment can be designed. A 2K factorial experiment will test all combinations of possible strategies, given a set of conditions. The number of experiments to be tested will thus be 2^n , where n is the number of individual strategies. To limit resource use and create a manageable experiment, the number of strategies to be tested will be limited to eight, and the number of patients per simulation to 10,000. For reproducibility, all simulations will start at the same seed.

To create a 2K factorial experiment, first the various strategies to be tested must be defined. For this experiment, we will test various bi-annual screening intervals. There will be eight strategies of five-year intervals for patients between 40 and 80 years of age. Then, a table will be created listing all possible

strategy combinations. In this case, this will be $2^8 = 256$ different experiments. These experiments will be listed in a table, where at first every strategy is 'turned off' (indicated by a '-'). In the next experiment, one more strategy will be turned on (indicated by a '+') or off, until experiment 256 is reached, where all strategies are turned on. This will result in the experiments shown in Table 20 for a short overview, and Appendix IX for the full overview. Experiment 1 will be an experiment with no screening at all, experiment 125 will replicate the current screening strategy in place in the Netherlands (bi-annual between ages 50 and 75), and Experiment 256 will test bi-annual screening for everyone between 40 and 80.

Table 20: 2K factorial experiment design

Exp No	40-45	45-50	50-55	55-60	60-65	65-70	70-75	75-80
1	-	-	-	-	-	-	-	-
2	+	-	-	-	-	-	-	-
3	-	+	-	-	-	-	-	-
...
124	+	+	-	+	+	+	+	-
125	-	-	+	+	+	+	+	-
126	+	-	+	+	+	+	+	-
...
254	+	-	+	+	+	+	+	+
255	-	+	+	+	+	+	+	+
256	+	+	+	+	+	+	+	+

The outcomes of the 2k factorial experiment can be presented on an ICER plane to see if there are strategies that perform better than the current strategy in the Netherlands.

5. Results

In the results section, the model experiments are run and the results are presented. First, a baseline is set with a run of 1 million patients at the current strategy. Other experiments will have their ICER calculated to that baseline.

5.1. Baseline Results

The first experiment simulates a total of 1 million patients. As this is too resource intensive to run on a single processor on a laptop, a parallel computing version of the model was created and ran on the University of Twente's Jupyterlab, on a Dell Poweredge 760 server with 144 cores (*UT-JupyterLab Wiki [Jupyter Wiki]*, n.d.). This means that the simulation could be performed around 100 times faster in optimal settings than on a single notebook CPU. Even at these increased speeds, the simulation still takes around 2 hours for this number of patients. The main KPIs for the baseline test are presented in Table 21. All cost data was rounded to whole euros; utility data was rounded to two decimals. 95% Confidence intervals are given in the brackets.

Table 21: Outcome KPIs of baseline simulation with 1M patients

Variable	Value
Patients Simulated	1,000,000
Cancer Incidence Percentage	14.1%
Percentage of tumors found through screening	54.8
Mean Total Life Years	80.15 (95% CI: 80.13 – 18.18)
Mean Quality-adjusted Life Years	72.31 (95% CI: 72.29-72.33)
Mean Total Costs	20056 (95% CI: 19965 – 20147)
Mean Medical Costs	9259 (95% CI: 9221 – 9298)
Mean Societal Costs	641 (95% CI: 635 – 647)
Mean PAID Costs	10155 (95% CI: 10091 – 10220)

The following experiment was run to see if the baseline result was stable. The baseline test of 1M patients was repeated in smaller batches: tests of 1,000, 10,000, and 100,000 patients. The ICER of these tests with the same strategy was calculated compared to the 1M baseline. These ICER values were plotted on an ICER plane, netting the results shown in Figure 19.

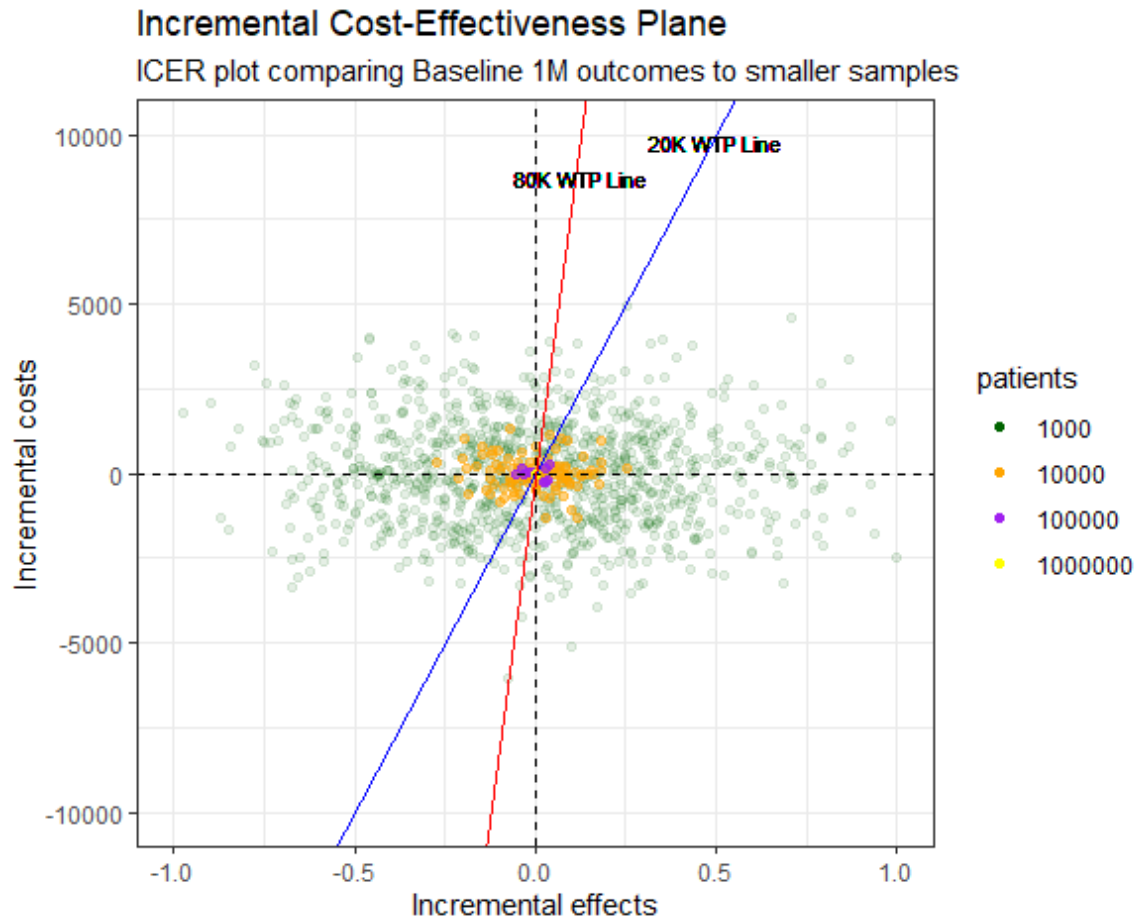


Figure 19: ICER plane showing reproducibility of baseline KPIs in smaller simulations

The more patients are simulated in a run, the more reproducible the result is. The experiment with 100,000 patients resembles the results of the large simulation with 1M patients well, whereas the results with 1000 patients are very diverse and do not provide steady results. Multiple batches of 10,000 patients show ICERs that are likely to cover an area that includes the actual mean. In reality, around 850,000 patients are screened annually in the Netherlands. However, as this number is infeasible due to computing power constraints, a lower number of patients per simulation is chosen. Experiments should be run with total patient populations of around 100,000 patients and in batches with the same starting seeds over the different experiments to net reproducible and accurate results.

A graph was created to get more insights into how large samples should be before they are stable, showing the running average of all patient's utilities. The average total utility experienced (QALY) of the base case with a sample size of 1M is now known to be 72.3. This graph in Figure 20 shows how the average Quality Adjusted Life Years experienced by the patients in a simulation converges towards the mean when the sample size is increased. The confidence intervals around the mean are also given and visibly converge towards the mean as the sample size increases.

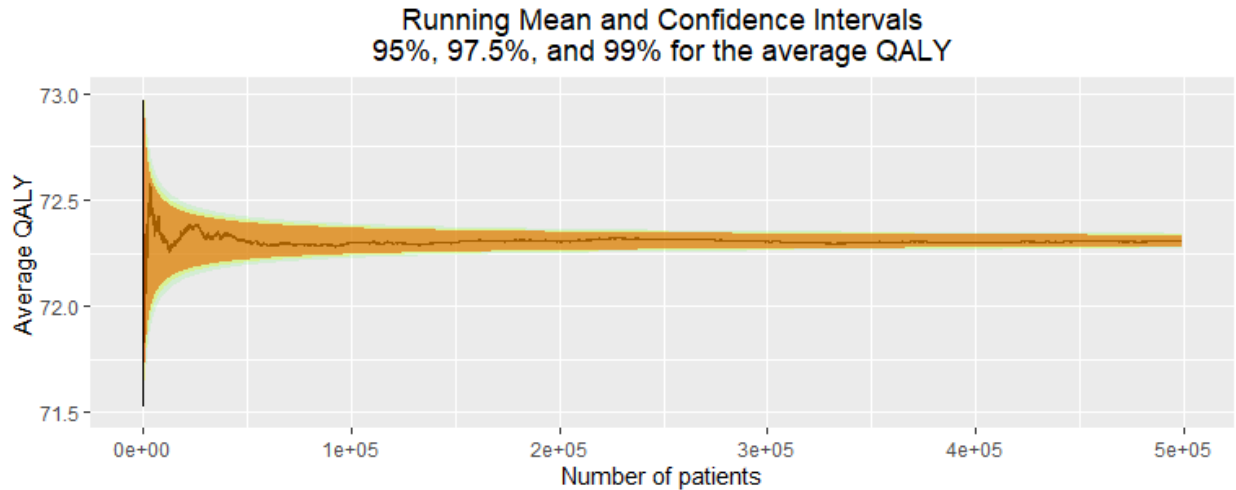


Figure 20: Running mean of simulated populations final total QALY

Table 22: Confidence intervals of mean total QALY at various simulation sizes

N patients	lower95	upper95	lower975	upper975	lower99	upper99	Observed mean
100	70.34197	74.26949	70.05319	74.55827	69.70640	74.90506	71.97282
1000	71.65660	72.95486	71.56317	73.04830	71.45203	73.15943	71.93737
10000	72.11101	72.50045	72.08305	72.52842	72.04981	72.56165	72.32241
100000	72.24393	72.36754	72.23505	72.37641	72.22450	72.38696	72.29801
500000	72.27776	72.33371	72.27374	72.33772	72.26897	72.34250	72.30626

This graph shows the confidence intervals converging as more patients are simulated. The confidence intervals are the colored bands in red, yellow, and green around the main mean. Table 22 shows the 95%, 97.5%, and 99% confidence intervals at various numbers of patients generated. These numbers, taken from the simulation, are given with five decimals to show how the outcomes differ. The 99% confidence interval remains the widest. After simulating 500,000 patients, we can say with 99% confidence that the actual mean lies between 72.27 and 72.34. At 100,000 patients simulated, the gaps of the confidence intervals are larger, but we can still say with 95% confidence that the actual mean lies between 72.24 and 72.37. The convergence of the confidence intervals happens exponentially more slowly, and exponentially more computing power is needed to get small decreases in the confidence intervals. To limit resource use and increase testing speed, it is wise to limit the number of patients generated per experiment to 100,000.

5.2. PSA Results

The probabilistic sensitivity analysis will be based on the experiments presented in Table 20 and Appendix X – 2K Factorial Experiment Design. These experiments are run to determine the effects of changing input variables on the study's outcome. In other words, the PSA is performed to check the stability of the results given changed input variables. It is done to see what would happen if real-life values differed significantly from the model's inputs.

Table 23: PSA Experiments and associated outcomes

Variable Changed	Value Used	ICER compared to Baseline	Mean Costs	Mean QALY
<i>Baseline</i>	-	-	20102 (95% CI: 19805 – 20399)	72.281 (95% CI: 72.218– 72.343)
<i>Screening Disutility</i>	0	-	2010 (95% CI: 19805 – 20399)	72.282 (95% CI: 72.220 – 72.344)
<i>Screening Disutility</i>	0.5	-	20102 (95% CI: 19805 – 20399)	72.280 (95% CI: 72.217 – 72.342)
<i>Diagnosis Disutility</i>	0	-	20102 (95% CI: 19805 – 20399)	72.283 (95% CI: 72.221 – 72.345)
<i>Diagnosis Disutility</i>	0.5	-	20102 (95% CI: 19805 – 20399)	72.279 (95% CI: 72.217 – 72.341)
<i>Screening Uptake</i>	0.5	-	20002 (95% CI: 19704-20298)	72.281 (95% CI: 72.218 – 72.343)
<i>Screening Uptake</i>	1	-	20548 (95% CI: 20249 – 20847)	72.287 (95% CI: 72.225 – 72.349)
<i>False Positives</i>	0%	-	17263 (95% CI: 16979 – 17547)	72.291 (95% CI: 72.229 – 72.353)
<i>False Positives</i>	5%	-	29995 (95% CI: 29661 – 30330)	72.279 (95% CI: 72.217 – 72.342)
<i>Treatment Costs</i>	0.25x	-	15906 (95% CI: 15659 – 16153)	72.281 (95% CI: 72.219 – 72.343)
<i>Treatment Costs</i>	2x	-	25697 (95% CI: 25320 – 26075)	72.281 (95% CI: 72.219 – 72.343)
<i>Societal Costs</i>	0x	-	19472 (95% CI: 19185 – 19760)	72.281 (95% CI: 72.219 – 72.343)
<i>Societal Costs</i>	2x	-	20731 (95% CI: 20425 – 21039)	72.281 (95% CI: 72.219 – 72.343)
<i>PAID Costs</i>	0x	-	9972.1 (95% CI: 9834.8 – 10110)	72.281 (95% CI: 72.219 – 72.343)
<i>PAID Costs</i>	2x	-	30233 (95% CI: 29736 – 30730)	72.281 (95% CI: 72.219 – 72.343)

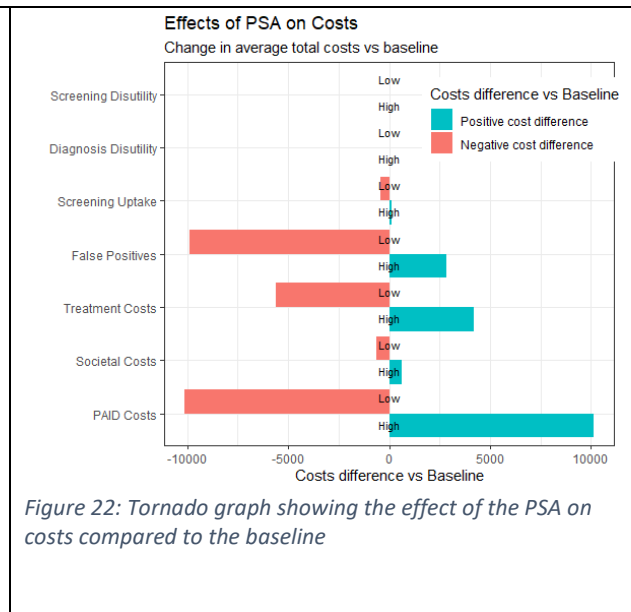
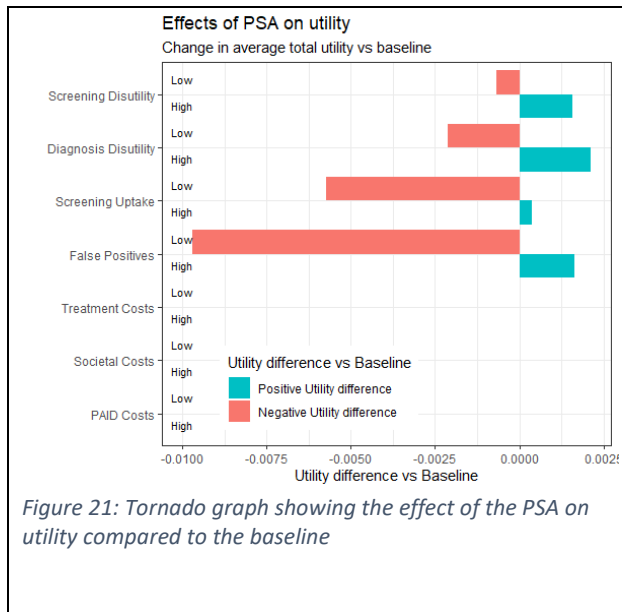
The results of the PSA experiments in Table 23 are based on runs for 100,000 patients, all using the same starting seed used for the different experiments. The experiments were run in 110 batches per experiment, using the same starting seeds for all experiments. The results of the experiments are visible in Table 23 and in the tornado graphs in Figure 21 and Figure 22. In Table 23, the figures are given with five significant figures to show the differences between the simulations.

The effects of changing a variable to a more extreme value can be seen in Table 23. The interventions that only affect the cost do not change the average experienced QALY at all. Interventions expected to impact the QALY, show no significant deviations from the QALY calculated on the baseline. Whilst costs do deviate significantly over experiments, the lack of significant variations in total QALY means no ICER can be calculated.

None of the QALY values are outside the confidence interval. Therefore, the disutility experienced during screening and diagnosis does not significantly affect the overall quality of life in the simulated population. Increasing or decreasing the screening uptake, the fraction of women adhering to their screening invitation, shows small but insignificant effects on both costs and utility.

Some cost differences are significant. Increasing or decreasing the treatment costs used significantly changes the average total accrued costs per patient. Including the effects of PAID and Societal costs in a simulation are significant. Together, they contribute to more than half of the costs associated with breast cancer from a societal perspective. The guideline’s advice on including these types of costs in health economic analyses significantly affects the outcomes of these studies (Zorginstituut Nederland, 2016).

Increasing and decreasing false positive rates does not lead to significant changes in QALY. This is likely because women who get a false positive diagnosis only experience minor disutility for a short amount of time, which is not enough to offset the average QALY of the population. Changing this variable does significantly affect the total costs parameter. False positives cause a significant part of the costs associated with the breast cancer screening program, due to the extra unnecessary diagnoses and treatment, and costs associated with that.



The tornado graphs above show the effect of changing input parameters to extremely high or low values on the costs and utilities compared to the baseline. Changing disutility associated with screening and diagnosis has minimal effects on the total QALYs of a patient compared to the baseline. Changing screening uptake has a more considerable effect. The ratio of false positives impacts costs and utility differences the most. Costs associated with other diseases after breast cancer treatment, the PAID costs, significantly affect the total accrued costs in the simulation. Changing the treatment costs leads to similar results. Increasing or decreasing societal costs and screening uptake slightly affects the total costs parameter.

5.3. Experiment Results

The costs and effects of the baseline experiment will be used in the ICER formula to compare the experiment outcomes to generate the incremental cost-effectiveness ratio of other experiments. The Incremental Cost Effectiveness Ratio is calculated by dividing the difference in costs by the difference in effectiveness. Equation 1 is repeated for clarity.

$$ICER = \frac{C_{Baseline} - C_{Experiment}}{E_{Baseline} - E_{Experiment}}$$

These ICERs are calculated through the simulation for the experiments shown in Table 24. The ICER is calculated by comparing the costs and effects of the experiment to the cost and effects of the baseline, as seen in the table above. All experiments were run for 100,000 patients divided over ten batches. The baseline experiment was performed again, as all experiments here use the same starting seed per batch and the same number of simulated patients for comparability.

Table 24: Results of different screening strategies

Experiment	B0	B1	B2	B3	B4	B5
Screening ages	50-75	None	50-75	50-75	40-80	55-70
Strategy	Bi-annual	-	Annual	Tri-annual	Bi-annual	Bi-annual
Mean Total Costs	20056 (95% CI: 19760-20351)	15401 (95% CI: 15128-15675)	23978 (95% CI: 23668-24287)	18740 (95% CI: 18451-19029)	23107 (95% CI: 22798-23416)	18271 (95% CI: 17986-18557)
Mean Medical Costs	9262.2 (95% CI: 9136.5-9387.9)	5070.5 (95% CI: 4972.1-5168.9)	12611 (95% CI: 12470-12753)	8181.4 (95% CI: 8061.5-8301.3)	12056 (95% CI: 11908-12205)	7782.4 (95% CI: 7666.1-7898.7)
Mean Societal Costs	632.86 (95% CI: 614.44-651.28)	477.78 (95% CI: 464.36-491.17)	880.46 (95% CI: 855.94-904.97)	567.41 (95% CI: 550.87-583.95)	824.61 (95% CI: 798.94-850.27)	567.06 (95% CI: 550.89-583.23)
Mean PAID Costs	10160 (95% CI: 9950.2-10371)	9853.1 (95% CI: 9647.0-10059)	10486 (95% CI: 10269-10701)	9991.0 (95% CI: 9783.8-10198)	10226 (95% CI: 10017-10436)	9921.9 (95% CI: 9715.2-10128)
Mean Total QALY	72.268 (95% CI: 72.206-72.329)	72.328 (95% CI: 72.267-72.389)	72.301 (95% CI: 72.240-72.363)	72.256 (95% CI: 72.194-72.318)	72.274 (95% CI: 72.212-72.335)	72.259 (95% CI: 72.200-72.321)
Mean Total LY	80.103 (95% CI: 80.023-80.183)	80.133 (95% CI: 80.053 – 80.213)	80.159 (95% CI: 80.079-80.239)	80.082 (95% CI: 80.002-80.162)	80.125 (95% CI: 80.046-80.205)	80.077 (95% CI: 79.997-80.157)
ICER	-	-77025	117170	113710	517367	214856
Cancer Incidence %	14.2	12.9	14.3	14.1	14.4	13.9
DCIS %	14.5	9.24	17.5	13.4	16.7	12.7
Stage I %	42.9	34.1	45.6	41.2	45.4	40.5
Stage II %	33.2	37.2	29.4	34.4	30.9	34.2
Stage III %	7.4	14.3	6.1	8.6	5.9	9.6
Stage IV %	2.0	5.1	1.5	2.4	1.4	2.9

The five strategies presented in Table 19 are tested in this experiment and compared to the baseline. Experiment B1 tests no screening at all, experiment B2 tests annual screening, whereas experiment B3 tests screening only once every three years. Experiments B4 and B5 test adjusted age ranges where patients might be eligible for screening. The results of the experiments are shown in Table 24. The ICER is calculated by comparing the costs and effects of the experiment to the cost and effects of the baseline. All experiments were run for 100,000 patients divided over ten batches. The baseline experiment was performed again, as all experiments here use the same starting seed per batch and the same number of simulated patients for comparability.

Experiment B1, with no screening implemented, shows a decrease in average total costs and a non-significant increase in average QALY compared to the baseline. This leads to a negative ICER. Significantly fewer cancer diagnoses are made, but the fraction of later-stage diagnoses is considerably higher. The disutility experienced in total by screening the entire population and treating tumors that would otherwise have had no need for treatment is worse than having some patients end up in a worse disease state.

If Annual screening is implemented, the costs and cancer incidence increase, and the QALY does not significantly increase. An increase in societal costs can be seen as people receive screening more often, bringing with it the extra costs of screening and possible complications. Medical costs also go up, due to more cancer treatments. The incidence of cancer also goes up, but the incidence of late-stage tumors goes down. This leads to a positive ICER of €117,170, meaning that would be the costs associated with gaining on average 1 additional QALY.

Tri-annual screening in experiment B3 shows the opposite effect of experiment B2. Costs go down, cancer incidence goes slightly down, and average quality of life decreases non-significantly. Late-stage tumors become more apparent as cancer incidence, in general, is somewhat lower. This also leads to a positive ICER, but the costs savings are insufficient to compensate for the loss in quality-adjusted life.

Increasing age ranges for bi-annual screening to include everyone between age 40 and 80 has a similar effect as implementing annual screening. The average total costs go up as more people receive screening, cancer incidence goes up as more cancers are found, and late-stage cancers show up only minimally. However, the increase in QALY is not significant, with this small difference in mean leading to an ICER of €517,367.

Decreasing the eligibility for screening, by setting the age range to between 55 and 70 years, shows the opposite effect. The mean total costs go down, and the average QALY goes down, but not significantly. There is less cancer incidence and late-stage tumors are present more often. The costs savings are too little to offset the very small loss in Quality of Life, leading to a high positive ICER.

In Figure 23, all different strategies are plotted on the cost-effectiveness plane. The difference in cost and the difference in effects of all batches compared to the baseline is calculated and plotted on the graph. The average difference in costs and effects is also given using a larger circle with a cross. The Willingness-to-pay lines of 80,000 euros per QALY and 20,000 euros per QALY are also given.

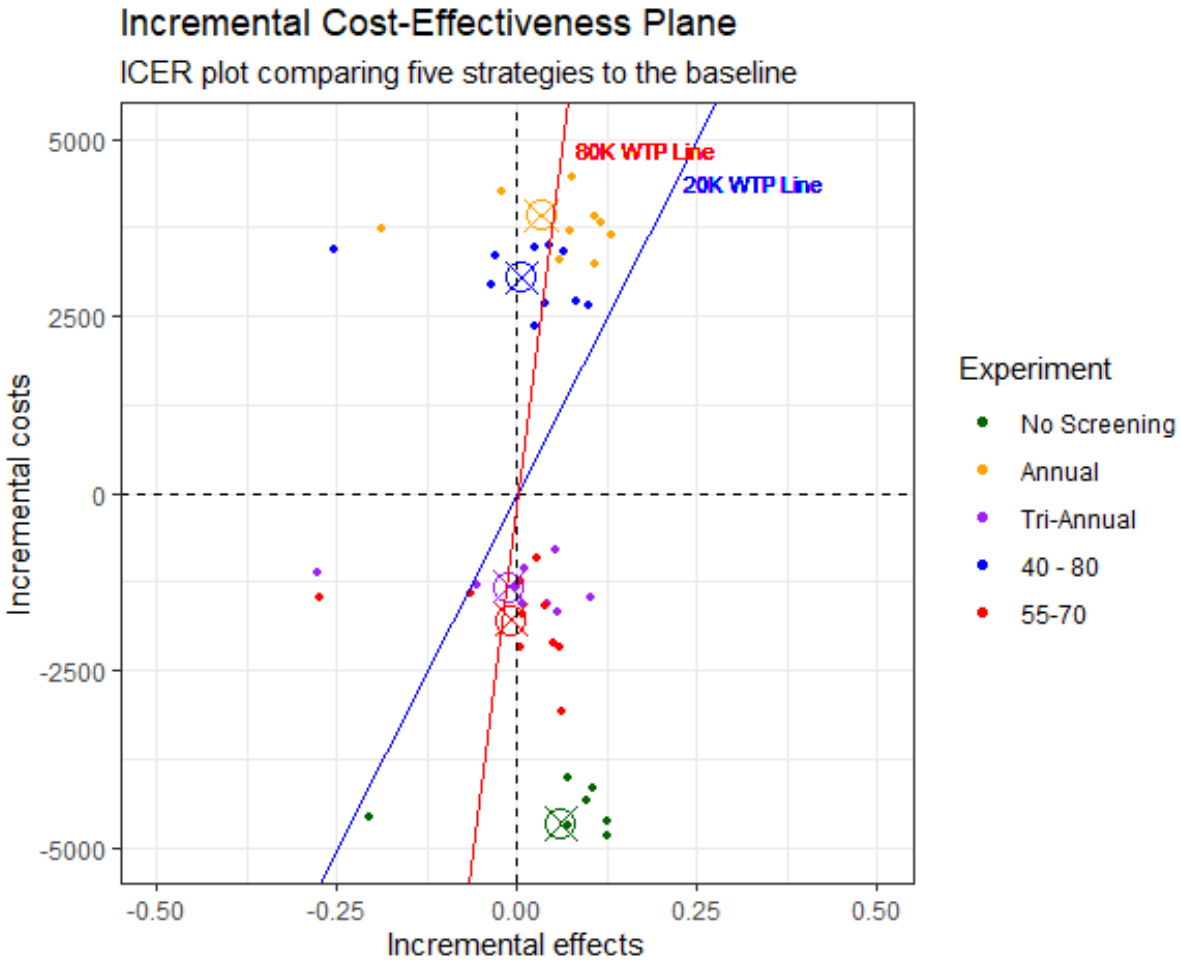


Figure 23: ICER plane comparing cost and utility outcomes of the strategies to the baseline

Comparing this CE-plane to the CE-Plane in Figure 5 in chapter 4.2, shows that no screening compared to the baseline results in values that are in an acceptable area of the ICER plane. This strategy leads to a higher total quality of life, at a savings in cost.

Decreasing screening eligibility to include only women between 55 and 70 and decreasing screening intervals from two to three years show a decrease in costs and a slight decrease in total experienced Quality of Life.

Annual screening and extending the age ranges for screening to between 40 and 80 increases costs associated with screening and treatment significantly, but it also increases the average quality of life. These strategies still end up above the €80K WTP line and are therefore not viable alternatives at these willingness to pay levels.

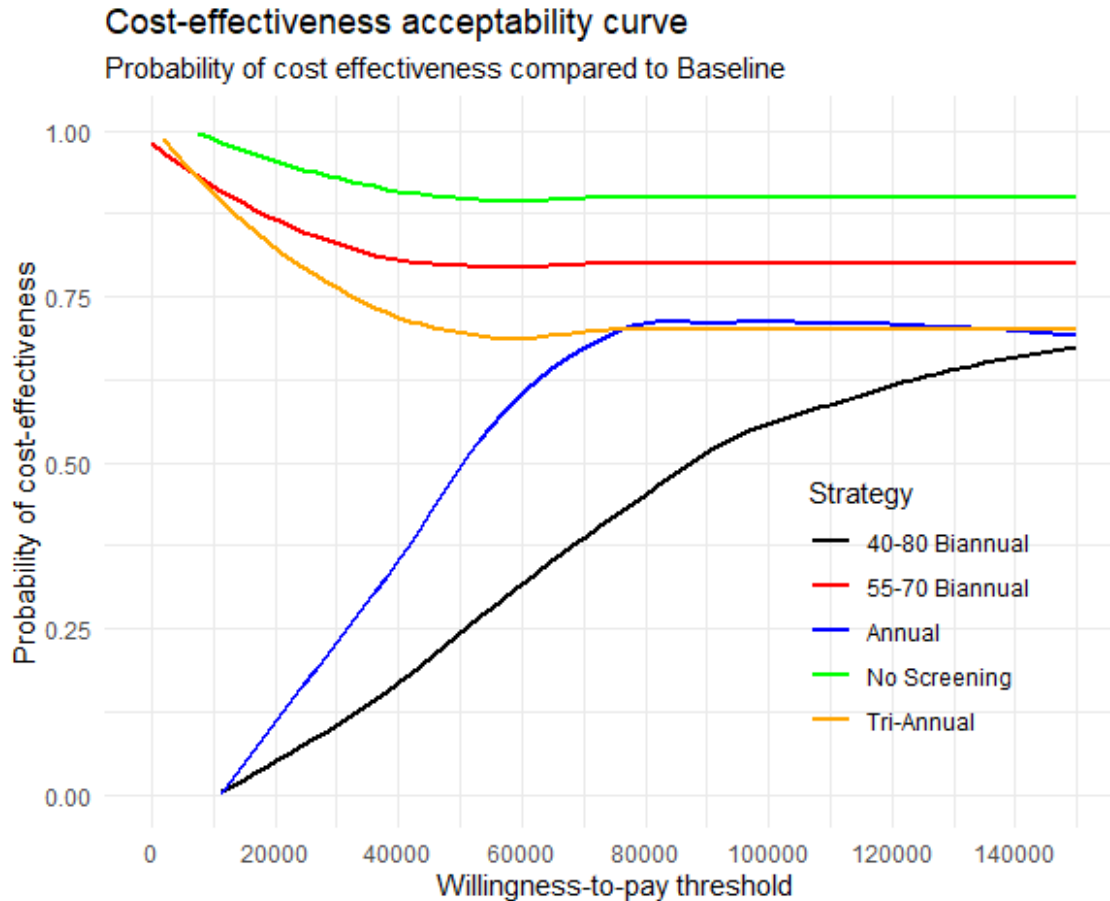


Figure 24: Cost-effectiveness acceptability curve of the screening strategies compared to the baseline

Figure 24 shows the different strategies plotted on the cost-effectiveness acceptability curve. At a willingness to pay threshold of 0, the strategies for no screening, tri-annual screening, and screening between ages 55 and 70 would be viable as they currently lead to a cost saving. As the willingness to pay increases, the probability of cost-effectiveness of the other strategies also increases. At a willingness to pay of 80,000, both annual and tri-annual screening are around 70% likely to be cost-effective. No screening is still the most cost effective, with up to 90% cost-effectiveness at a WTP level of 150,000 per QALY.

Besides the health-economic point of view in the cost-effectiveness plane and the cost effectiveness acceptability curve, an important point to consider is the stage distribution of the tumors found under various screening strategies. In Figure 25, the absolute distribution is shown, and incidence percentages are given per column. The five new strategies and the baseline are shown in this graph. If screening was to be abolished and no screening would be implemented, there would be a significant decrease in found tumors and a sharp incidence shift towards later-stage tumors. The relative and absolute incidence of stage II, III and IV tumors increase compared to all other strategies. In general, abolishing the breast cancer screening program would not lead to a significant change in the average quality of life for the population. However, there will be a portion of the population who has a significantly worse experience than if any other screening strategy would be implemented.

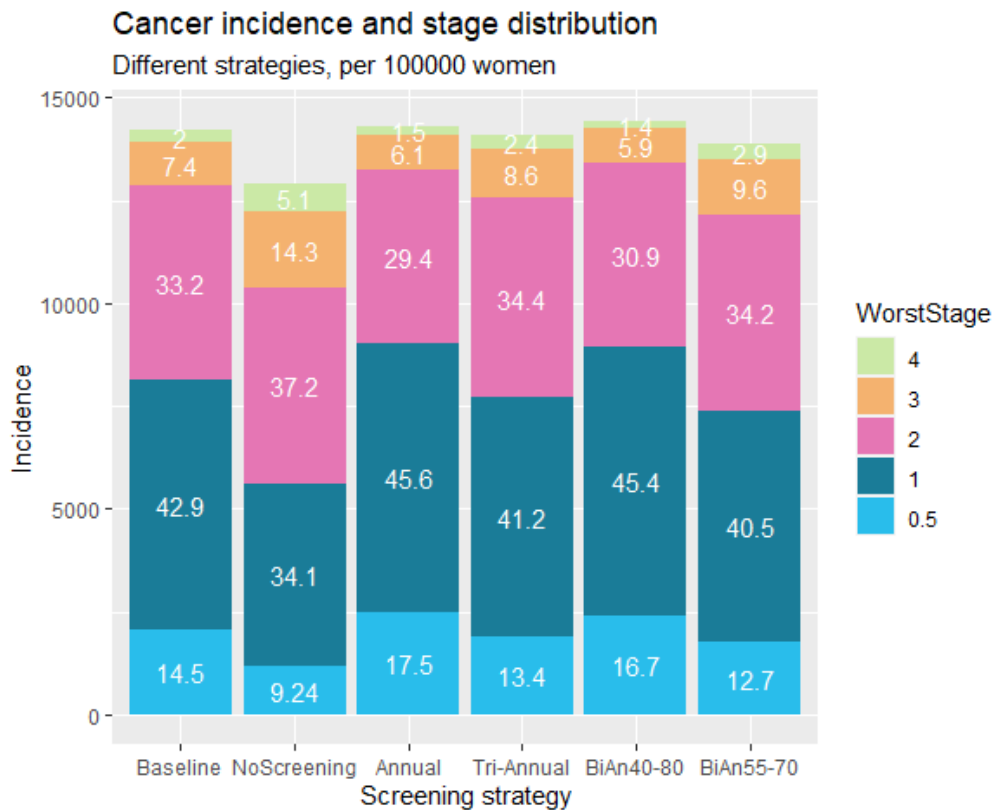


Figure 25: Cancer stage distribution for the strategies. Incidence on y-axis and percentages in bars.

5.4. Strategy Test Results

In 4.5.4., a 2k factorial design was presented to test 256 unique screening strategies, to see which strategy would perform best in the simulation. All strategies were tested for a total of 10,000 patients. The difference in costs and in utility compared to the baseline was calculated. As a baseline, strategy 125 was taken. This strategy tests biannual screening between 50 and 75, and this particular test simulated the same number of patients in the same 16 batches and using the same random seeds as the other 255 tests.

The different strategies were all compared to this baseline, and their difference in costs and effects was plotted on the CE-plane in Figure 26. The color of the points in the plane shows the maximum number of screenings performed. The current strategy, bi-annual screening between 50 and 75, would lead to a maximum of 13 screenings. Screening between 40 and 80 would lead to maximally 20 screenings, and no screening would lead to a maximum of zero.

This shows that most strategies are less cost-effective than the current strategy. A handful of strategies would show better effects at a cost reduction. These are mainly strategies including very little screening in total. These results are in line with the results of the strategy experiments, where no screening led to the best ICER. Some strategies lead to better effects at higher costs, which are akin to the 40-80 strategy

presented in the previous experiments. All other strategies lead to worse effects at various cost differences.

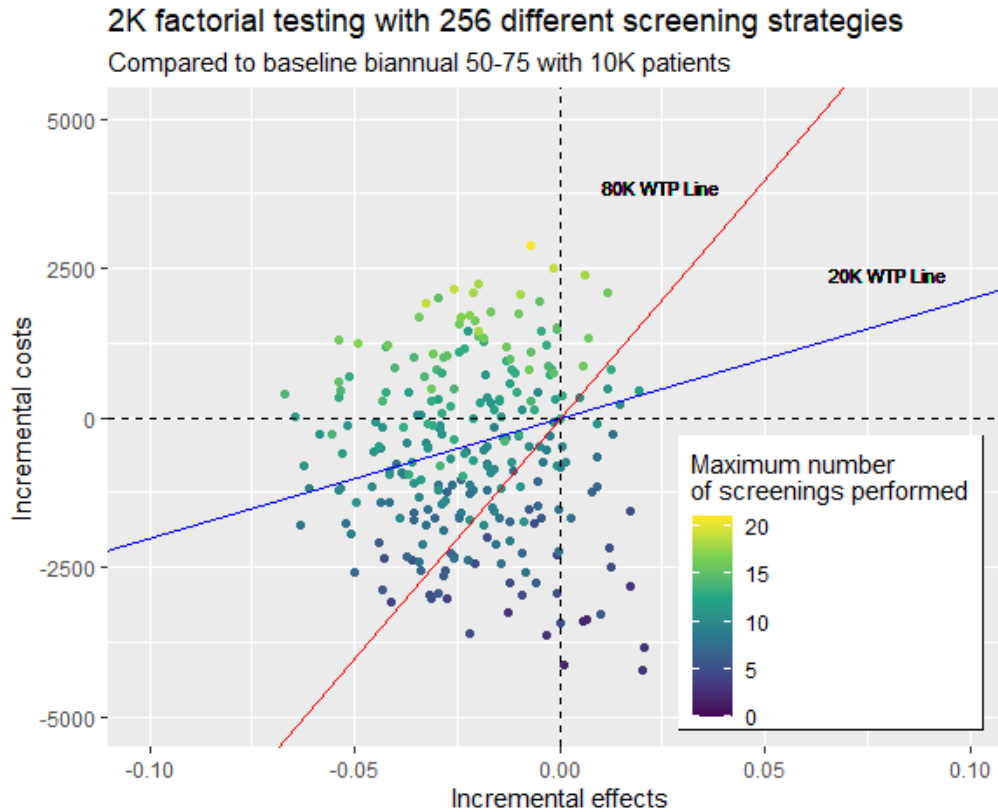


Figure 26: Cost-effectiveness plane showing 255 screening strategies compared to the baseline

5.5. Budget Impact Analysis

Besides difference in average total costs and effects per patient, as well as the likelihood of acceptance of the different strategies, it is also important to analyze the budget impact that implementing various strategies would have on the Dutch healthcare system.

The strategies tested in 5.3 show the average costs and effects per person when simulating a cohort of 100,000 women for their entire lifetimes. The annual cost of the screening program in the Netherlands is around 80 million euros, as presented in the annual report of Bevolkingsonderzoek Nederland (PriceWaterhouseCoopers Accountants N.V., 2022). However, these costs are annual and extend invitations to more than 1.2M women annually. The difference in medical costs per patient per screening round should be multiplied with the number of people eligible for screening to translate the costs from the simulation to the annual budget impact.

Right now, every patient goes through a maximum of 13 screening rounds. Because patients are screened, it costs the Netherlands on average 4192 euros per person more in medical care over their lifetimes. This comes down to around 322 euros extra per screening round. In the Netherlands, on average 1.2M women are invited for screening each year. This leads to a budget impact of 387M per year in extra costs when comparing the current screening strategy to no screening. Table 25 shows the budget impact of medical costs of the tested strategies.

Table 25: Budget impact of various screening strategies compared to the baseline

Experiment	Screening ages	Strategy	Budget impact of medical costs
B0	50-75	Bi-annual	
B1	None	-	€387,000,000
B2	50-75	Annual	- €309,000,000
B3	50-75	Tri-annual	€100,000,000
B4	40-80	Bi-annual	- €258,000,000
B5	55-70	Bi-annual	€137,000,000

Abolishing screening would result in costs savings of 387M annually. Implementing annual screening or extending the age ranges of eligibility would go paired with increased medical spending. Only screening once every three years or only between the ages of 55 and 70 would free up over 100M in budget annually.

With these results, research question 3 can be answered. The health and economic status of the screening program is now known. The health economic outcomes of the simulations include patient utility and costs, and the outcomes vary between the strategies. The societal costs for the screening program have been calculated per person and strategy, and the total costs are shown in the budget impact. Answering question 3.3 proves difficult and opens room for an ethical discussion. The current health economic evaluation metrics suggest that the current costs for the screening program are high. Still, these current standards have yet to be revamped after the suggestion of the inclusion of more costs in these calculations. The health gains are insignificant on a per-person basis. However, the reduction of late-stage tumors due to the screening program is significant. The acceptability of the total costs associated with that are for now outside of often acceptable health economic criteria, but this study opens room for debate on these benefits.

6. Conclusion

This research has determined the health economic performance of the breast cancer screening program in the Netherlands. To answer the first research question, the status of the screening program according to the literature was assessed. The literature does not provide consensus on the current status of the screening program. A systematic literature review was performed, which showed that the health- and economic impact varied greatly (between €5,000 to €54,000 per QALY) over different studies. The studies analyzed all used different methods for determining the current costs and the health harms and benefits, with none of the studies using methods that were up to par with current health economic analysis guidelines. From these studies it became clear that new research should include more costs and quality of life parameters in order to arrive at a meaningful conclusion

The population health of populations with different screening strategies was compared to determine the health effects of the screening program. For individual women in the Netherlands, the results show clear benefits of participating in the screening program for breast cancer. The probability of being diagnosed with a late-stage tumor is significantly less for women who are screened than for those who are not screened. The disutility experienced during screening is negligible over a lifetime with the current mammography screening strategy. Even the effects of false positives are, on average, not significant for the population. The average individual will have better post-diagnosis prognosis when the tumor is detected earlier.

A tool has been created that allows comparison of screening strategies for breast cancer. The model has been built in compliance with the latest guidelines for health economic analyses. Compared to older models, this model considers more factors for utility and disutility, including short time periods for screening and diagnosis. The model also accounts for different types of costs. Previous studies were limited to medical costs for breast cancer treatment. This study also calculates the societal costs related to time costs from missed work and travel costs, as well as costs due to unrelated diseases in gained life years. This shines a broader perspective on the costs involved with the screening program. The model is validated in accordance with the AdViSHE tool recommended by health economic evaluation guidelines. Cost, utility, treatment, survival data and various other input parameters for the model were based on the literature. This model can now serve as a base for further health economic evaluations on screening programs and could, with small modifications, be used for different diseases and proposed screening programs like bowel cancer, ovarian cancer, prostate cancer, and lung cancer.

This research has shined a light on new aspects when evaluating the screening program on a societal level, and these points should be taken as the start of an ethical discussion on the future of the breast screening program. This research shows no significant difference in average quality adjusted life years experienced when comparing the current screening strategy to a strategy without screening (72.27 vs. 72.33), while all costs related to breast cancer per individual are higher (€20,056 vs. €15,401). This means that the screening program currently does not add quality of life but does increase healthcare costs. It is important to keep in mind the aims of screening: reducing disease-related mortality, incidence, and severity. The average severity is reduced, with a 57% and 43% reduction in stage IV and stage III diagnoses, respectively. This in turn leads to lower disease-related mortality, specifically for these groups.

Six different screening strategies have been compared in an extensive analysis. The current strategy of bi-annual screening between ages 50 and 75 was compared to no screening, annual and tri-annual screening

between 50 and 75, bi-annual screening between 40 and 80 and bi-annual screening between 55 and 70. This comparison shows that screening leads to an increase in diagnoses, but a decrease in late-stage diagnoses. This is most apparent between the baseline test and the tests for no screening, but other comparisons show the same result. The 2K factorial analysis shows that the current screening strategy tries to find the best of both worlds: compared to other screening strategies, the current program has about average costs but above average quality-adjusted life years compared to 255 other strategies. Besides this feat, late-stage diagnoses are limited, and increasing screening does not reduce this incidence as significantly as limiting screening would increase the incidence of late-stage tumors. The current screening strategy works rather well compared to other strategies when considering the impact of disease.

An essential part of this research is the Budget Impact Analysis. The budget impact analysis answered the third research question, by putting the health and cost outcomes in a societal perspective. Current staff shortages and increasing healthcare costs are a topic of concern in the Netherlands. The BIA aims to provide insight into the total costs for the healthcare system associated with the screening program. It does not guide where to save costs, but only provides information on what changing the screening strategy could mean for the Dutch healthcare system. A new ethical discussion is recommended to decide what these costs and changes to the program would implicate. The budget impact coming forth from this analysis is significant. The direct costs associated with the screening program are €80M, but comparing the total cost differences between the strategies sees a decrease of €387M, €100M, and €137M when moving to no screening, tri-annual screening or reduced screening between 55 and 70. Increasing screening to annual screening or extending the age ranges to include all women between 40 and 80 would increase costs with €309M and €258M, respectively.

This research provides a new, validated model for health economic analyses for cancer screening programs, with this version focused on breast cancer. The study's results on different breast cancer screening programs aim to provide a basis for an ethical discussion on how benefits and drawbacks should be valued. The current screening program in the Netherlands is one of the best options. Severe cancer incidence is significantly reduced due to the screening program, but that does come with an increase in overall diagnoses and high costs. Besides starting an ethical discussion, there is an explicit recommendation for further research with this model, where cancer institutes and governing bodies should populate the model with their own governed data which was unavailable in this research. This research led to many new insights in the breast cancer screening program, which is valuable for society and science alike.

7. Discussion

This study comes with some limitations. The most important limitations are shown in this discussion section, and advice for further research is offered.

The costs included in this study are more extensive than the costs used in previous studies. This raises a few points for discussion. Previous studies were limited to only include medical costs. The results showed that medical costs comprise only about 50% of the total costs, ranging from 53% in the annual screening scenario to 33% in the no-screening scenario. The other types of costs, the societal and costs for unrelated diseases in gained life years, account for the other half. This leads to higher total costs associated with the screening programs. The willingness to pay values of €20,000 and €80,000 referenced in this study are only used as reference values, but they were previously used in research that only included medical costs and might therefore not translate to an accurate reflection of the willingness to pay value in Dutch health economic analyses in the future, and, given the more comprehensive approach to costs inclusion, a revision of these willingness to pay values should be considered for future research.

The costs included in this study are based on cost figures from the literature. Medical costs are based on previous research on breast cancer, societal costs are based on the guidelines for health economic analyses, and the costs for unrelated diseases in gained life years are based on the PAID 3.0 tool. Future research performed by cancer institutes or governmental bodies could have access to better sources for cost figures, with the possibility to draw more data from the same source. This could significantly impact the final costs, as shown in the PSA experiments.

New screening modalities could impact the results of this research. This study is based on the currently used mammography screening. There are calls to implement MRI screening for specific risk groups, and there is ongoing research into mammogram alternatives that lead to less discomfort for screened women. New screening modalities could lower discomfort but can also change the number of false positives. Different screening methods could decrease the number of false positives due to better systems, but they could also have the opposite effect. If the new diagnostic method is more sensitive to smaller abnormalities, false positives could increase, and overtreatment becomes a larger concern. When new screening modalities are considered, this model could check their effectiveness.

The experiments used a variety of sample sizes. The first baseline experiment used for validation simulated 1 million patients. This analysis showed that experiments of 100,000 patients would lead to stable results. These experiments with 100,000 patients were used in the PSA and strategy experiments. However, these experiments cost a lot of computational power, with experiments of 100,000 patients taking around 40 minutes to finish. Discrete event simulation is a computationally expensive means of analysis in general, and patients in this model experience on average 10,000 attribute updates per simulation, with each update requiring multiple computational operations. The original idea for this research was to only compare two different strategies. By including the PSA and 2k factorial, the number of experiments required increased 100-fold. To limit time spent on running experiments, the choice was made to limit the number of patients in these additional simulations.

Therefore, the 2k factorial experiment used only 10,000 patients per experiment, with all experiments using the exact same seed. The drawback of using such a small sample size is that the results can be skewed. In experiments with 10,000 patients, a 14% cancer incidence and a 5% stage IV incidence means that only around 70 patients are diagnosed with stage IV cancer. In these sample sizes, the seeds used

could significantly influence the result. When possible, these experiments should be replicated with a larger sample size.

The results of the PSA showed that the disutility experienced during screening and diagnosis is minimal. Overtreatment and overdiagnosis do have a significant impact on population health. The most important parameter from the PSA was the cost values used. If treatment costs differ from the values used in the research, the ICER rates will also vastly differ. The probability of these effects is kept to a minimum by only using data from reputable sources. To minimize the potential for a change in results, the difference in treatment costs for the various cancer stages was checked, and the percentual changes were stable over various studies. Still, the costs could differ from reality. Therefore, there is still advice for cancer institutes to repeat all experiments with their input data.

During this research, many ideas for additional research arose. Currently, screening buses tour the country to systematically invite women per area. This means that all eligible women of a certain city or village are invited for screening at the same time. The results from these screenings are also available at similar times after initial screening. Local hospitals then need to accommodate women that require additional diagnostics and possible treatment in waves of increased intensity every two years. It would be interesting to analyze the effect these waves have on the capacity of hospitals, and if these periods of increased capacity requirements can be handled more efficiently.

The model can also easily be expanded to track other interesting parameters. For example, the model could be expanded to measure the environmental impact of the screening program by including additional parameters that track the carbon emissions related to breast cancer screening and travel per patient in the model.

This study provides a starting point for further research. The model provides a means to investigate the cost-effectiveness and budget impact for screening programs and the result provides a basis on the current status of the Dutch breast cancer screening program. Furthermore, the results are meant to start an ethical discussion on how we value disease and treatment, and what levels of discomfort we are willing to experience to reduce the possibility of severe illness. To further validate the results of this study and, more importantly, assess the exact health-economic status of the Dutch breast cancer screening program, there is an explicit request for governmental bodies concerned with public health and cancer institutes to replicate the experiments from this study based on their input data.

Bibliography

- Accenture. (2022a). *Healthcare Technology Trends 2022 | Tech Vision | Accenture*.
<https://www.accenture.com/us-en/insights/health/digital-health-technology-vision>
- Accenture. (2022b). *Quantum Enhanced Supply Chain Optimization*.
- American Cancer Society. (2020). *ACS Breast Cancer Screening Guidelines*.
<https://www.cancer.org/cancer/breast-cancer/screening-tests-and-early-detection/american-cancer-society-recommendations-for-the-early-detection-of-breast-cancer.html>
- Amin, M. B., Edge, S. B., Greene, F. L., Schilsky, R. L., Brookland, R. K., Washington, M. K., Gershenwald, J. E., Compton, C. C., Hess, K. R., Sullivan, D. C., Jessup, M. J., Brierley, J. D., Gaspar, L. E., Balch, C. M., Winchester, D. P., Asare, E. A., Madera, M., Gress, D. M., & Meyer, L. R. (2017). American Joint Committee on Cancer (AJCC). *AJCC Cancer Staging Manual*. In *AJCC Cancer Staging Manual*.
<https://link.springer.com/book/9783319406176>
- Arrospide, A., Rue, M., van Ravesteyn, N. T., Comas, M., Soto-Gordoa, M., Sarriugarte, G., & Mar, J. (2016). Economic evaluation of the breast cancer screening programme in the Basque Country: retrospective cost-effectiveness and budget impact analysis. *BMC Cancer*, *16*(1).
<https://doi.org/10.1186/S12885-016-2386-Y>
- Avanade, Accenture, & Microsoft. (2021). *Healthcare on Azure*.
- Berry, D. A., Inoue, L., Shen, Y., Venier, J., Cohen, D., Bondy, M., Theriault, R., & Munsell, M. F. (2006). Modeling the impact of treatment and screening on U.S. breast cancer mortality: a Bayesian approach. *Journal of the National Cancer Institute. Monographs*, *36*, 30–36.
<https://doi.org/10.1093/JNCIMONOGRAPHS/LGJ006>
- Bevolkingsonderzoek Nederland. (2022). *Bevolkingsonderzoek borstkanker*.
<https://www.bevolkingsonderzoeknederland.nl/en/borstkanker/>
- Blank, C., Park, D. K., & Petruccione, F. (2021). Quantum-enhanced analysis of discrete stochastic processes. *Npj Quantum Information* *2021 7:1*, *7*(1), 1–9. <https://doi.org/10.1038/s41534-021-00459-2>
- Bromley, H. L., Petrie, D., Mann, G. B., Nickson, C., Rea, D., & Roberts, T. E. (2019). Valuing the health states associated with breast cancer screening programmes: A systematic review of economic measures. *Social Science & Medicine*, *228*, 142–154.
<https://doi.org/10.1016/J.SOCSCIMED.2019.03.028>
- Caro, J. J., Briggs, A. H., Siebert, U., & Kuntz, K. M. (2012). Modeling good research practices--overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--1. *Value in Health : The Journal of the International Society for Pharmacoeconomics and Outcomes Research*, *15*(6), 796–803. <https://doi.org/10.1016/J.JVAL.2012.06.012>

- Carter, K. J., Castro, F., & Morcos, R. N. (2018). Insights Into Breast Cancer Screening: A Computer Simulation of Two Contemporary Screening Strategies. *AJR. American Journal of Roentgenology*, 210(3), 564–571. <https://doi.org/10.2214/AJR.17.18484>
- Carter, S. M., & Barratt, A. (2017). What is overdiagnosis and why should we take it seriously in cancer screening? *Undefined*, 27(3). <https://doi.org/10.17061/PHRP2731722>
- Centraal Bureau voor de Statistiek. (2022). *Prognose periode-levensverwachting; geslacht en leeftijd, 2022-2070*. <https://www.cbs.nl/nl-nl/cijfers/detail/85325NED>
- Chia, Y. L., Salzman, P., Plevritis, S. K., & Glynn, P. W. (2004). Simulation-based parameter estimation for complex models: a breast cancer natural history modelling illustration. *Statistical Methods in Medical Research*, 13(6), 507–524. <https://doi.org/10.1191/0962280204SM380RA>
- Clark, S., & Reeves, P. J. (2015). Women’s experiences of mammography: A thematic evaluation of the literature. *Radiography*, 21(1), 84–88. <https://doi.org/10.1016/J.RADI.2014.06.010>
- Clift, A. K., Dodwell, D., Lord, S., Petrou, S., Brady, S. M., Collins, G. S., & Hippisley-Cox, J. (2022). The current status of risk-stratified breast screening. *British Journal of Cancer*, 126(4), 533–550. <https://doi.org/10.1038/s41416-021-01550-3>
- De Glas, N. A., De Craen, A. J. M., Bastiaannet, E., Land, E. G. O. t., Kiderlen, M., Van De Water, W., Siesling, S., Portielje, J. E. A., Schuttevaer, H. M., De Bock, G. T. H., Van De Velde, C. J. H., & Liefers, G. J. (2014). Effect of implementation of the mass breast cancer screening programme in older women in the Netherlands: population based study. *BMJ (Clinical Research Ed.)*, 349. <https://doi.org/10.1136/BMJ.G5410>
- Department Of Pathology. (2023). *Types of Breast Cancer - Breast Pathology*. Johns Hopkins Pathology. <https://pathology.jhu.edu/breast/types-of-breast-cancer>
- D.J. de Leede, & R. van Velzen. (2022). *Beleidskader Bevolkingsonderzoeken naar Kanker*. <https://www.rivm.nl/bibliotheek/rapporten/2022-0137.pdf>
- Dowie, J. (2004). Why cost-effectiveness should trump (clinical) effectiveness: The ethical economics of the South West quadrant. *Health Economics*, 13(5), 453–459. <https://doi.org/10.1002/HEC.861>
- Drossaert, C. H. C., Boer, H., & Seydel, E. R. (2002). Monitoring women’s experiences during three rounds of breast cancer screening: results from a longitudinal study. *Journal of Medical Screening*, 9(4), 168–175. <https://doi.org/10.1136/JMS.9.4.168>
- Economic Evaluation Overview | POLARIS | Policy and Strategy | CDC*. (n.d.). Retrieved December 5, 2022, from <https://www.cdc.gov/policy/polaris/economics/index.html>
- Engelman, K. K., Cizik, A. M., & Ellerbeck, E. F. (2008). Women’s Satisfaction with Their Mammography Experience: Results of a Qualitative Study. http://Dx.Doi.Org/10.1300/J013v42n04_02, 42(4), 17–35. https://doi.org/10.1300/J013V42N04_02
- Federatie Medisch Specialisten. (2022). *Prinsjesdag 2022: reactie Federatie | Federatie Medisch Specialisten*. <https://demedischspecialist.nl/nieuwsoverzicht/nieuws/prinsjesdag-2022-reactie-federatie>

- Fryback, D. G., Stout, N. K., Rosenberg, M. A., Trentham-Dietz, A., Kuruchittham, V., & Remington, P. L. (2006). The Wisconsin Breast Cancer Epidemiology Simulation Model. *Journal of the National Cancer Institute. Monographs*, 36, 37–47. <https://doi.org/10.1093/JNCIMONOGRAPHS/LGJ007>
- Gezondheidsraad. (2008). *Screening: tussen hoop en hype*.
- Ghaemian, N., Tehrani, N. H. G., & Nabahati, M. (2021). Accuracy of mammography and ultrasonography and their BI-RADS in detection of breast malignancy. *Caspian Journal of Internal Medicine*, 12(4), 573. <https://doi.org/10.22088/CJIM.12.4.573>
- Gøtzsche, P. C., & Jørgensen, K. J. (2013). Screening for breast cancer with mammography. *The Cochrane Database of Systematic Reviews*, 2013(6). <https://doi.org/10.1002/14651858.CD001877.PUB5>
- Health Council of the Netherlands. (2014a). *Advies bevokingsonderzoek naar borstkanker: verwachtingen en ontwikkelingen*.
- Health Council of the Netherlands. (2014b). *Population screening for breast cancer: expectations and developments*.
- Heinrich, S. (2003). From Monte Carlo to quantum computation. *Mathematics and Computers in Simulation*, 62(3–6), 219–230. [https://doi.org/10.1016/S0378-4754\(02\)00239-2](https://doi.org/10.1016/S0378-4754(02)00239-2)
- Integraal Kankercentrum Nederland. (2020). *Oncoguide*. <https://oncoguide.nl/#!/projects/7/guideline>
- Integraal Kankercentrum Nederland. (2022). *Borstkankercijfers*. <https://iknl.nl/borstkankercijfers>
- Integraal Kankercentrum Nederland (IKNL). (2022). *Synthetische dataset Nederlandse Kankerregistratie (NKR)*. <https://iknl.nl/NKR/Cijfers-Op-Maat/Synthetische-Data>.
- Khan, S. A., Hernandez-Villafuerte, K. V., Muchadeyi, M. T., & Schlander, M. (2021). Cost-effectiveness of risk-based breast cancer screening: A systematic review. *International Journal of Cancer*, 149(4), 790–810. <https://doi.org/10.1002/IJC.33593>
- Koffijberg, E. (2021). *Advanced Simulation for Health Economic Analysis*.
- Law, A. M., Law, & M., A. (2015). *Simulation Modeling & Analysis*. McGraw-Hill.
- Lead time bias - Catalog of Bias*. (n.d.). Retrieved December 5, 2022, from <https://catalogofbias.org/biases/lead-time-bias/>
- Lee, S. J., & Zelen, M. (2003). Modelling the early detection of breast cancer. *Annals of Oncology : Official Journal of the European Society for Medical Oncology*, 14(8), 1199–1202. <https://doi.org/10.1093/ANNONC/MDG323>
- Lerner, B. H. (2016). “To See Today with the Eyes of Tomorrow”: A History of Screening Mammography. <https://doi.org/10.3138/Cbmh.20.2.299>, 299–321. <https://doi.org/10.3138/CBMH.20.2.299>
- Mandelblatt, J., Schechter, C. B., Lawrence, W., Yi, B., & Cullen, J. (2006). The SPECTRUM population model of the impact of screening and treatment on U.S. breast cancer trends from 1975 to 2000: principles and practice of the model methods. *Journal of the National Cancer Institute. Monographs*, 36, 47–55. <https://doi.org/10.1093/JNCIMONOGRAPHS/LGJ008>

- McNamee, P., Murray, E., Kelly, M. P., Bojke, L., Chilcott, J., Fischer, A., West, R., & Yardley, L. (2016). Designing and Undertaking a Health Economics Study of Digital Health Interventions. *American Journal of Preventive Medicine*, 51(5), 852–860. <https://doi.org/10.1016/j.amepre.2016.05.007>
- Medisch Contact. (2022). 'Zorgkosten lopen in 2060 op tot 18 procent van het bbp' | medischcontact. <https://www.medischcontact.nl/nieuws/laatste-nieuws/nieuwsartikel/zorgkosten-lopen-in-2060-op-tot-18-procent-van-het-bbp-.htm>
- Wet op het bevolkingsonderzoek, (2021) (testimony of Minister van Welzijn Volksgezondheid en Cultuur). <https://wetten.overheid.nl/BWBR0005699/2021-07-01>
- Beschikking van de Staatssecretaris van Volksgezondheid, Welzijn en Sport, houdende verlenging van de vergunning op grond van de Wet op het bevolkingsonderzoek voor het bevolkingsonderzoek naar borstkanker, (2020) (testimony of Welzijn en Sport Ministerie van Volksgezondheid).
- Mittmann, N., Stout, N. K., Tosteson, A. N. A., Trentham-Dietz, A., Alagoz, O., & Yaffe, M. J. (2018). Cost-effectiveness of mammography from a publicly funded health care system perspective. *CMAJ Open*, 6(1), E77–E86. <https://doi.org/10.9778/CMAJO.20170106>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Altman, D., Antes, G., Atkins, D., Barbour, V., Barrowman, N., Berlin, J. A., Clark, J., Clarke, M., Cook, D., D'Amico, R., Deeks, J. J., Devereaux, P. J., Dickersin, K., Egger, M., Ernst, E., ... Tugwell, P. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLOS Medicine*, 6(7), e1000097. <https://doi.org/10.1371/JOURNAL.PMED.1000097>
- Moore et al. (2018). Moore Clinically Oriented Anatomy 8th Edition. *Wolters Kluwer*, 282(15), 1485–1485.
- Mühlberger, N., Sroczyński, G., Gogollari, A., Jahn, B., Pashayan, N., Steyerberg, E., Widschwendter, M., & Siebert, U. (2021). Cost effectiveness of breast cancer screening and prevention: a systematic review with a focus on risk-adapted strategies. *European Journal of Health Economics*, 22(8), 1311–1344. <https://doi.org/10.1007/S10198-021-01338-5/FIGURES/2>
- Nationaal Borstkanker Overleg Nederland, & Nederlandse Internisten Vereniging. (2020, February 7). *Borstkanker - Algemeen - Richtlijn - Richtlijndatabase*. <https://richtlijndatabase.nl/richtlijn/borstkanker/algemeen.html>
- National Cancer Institute. (2022). *CISNET: Modeling Approach*. <https://cisnet.cancer.gov/modeling/index.html>
- Nederland, Z. (2018). Ziektelast in de praktijk. *De Theorie En Praktijk van Het Berekenen van Ziektelast Bij Pakketbeoordelingen*.
- Nederlandse Rijksoverheid. (2021). *Zorguitgaven internationaal vergelijkbaar; functies en financiering | Data overheid*. <https://data.overheid.nl/dataset/641-zorguitgaven-internationaal-vergelijkbaar--functies-en-financiering>
- Nelson, H. D., Fu, R., Cantor, A., Pappas, M., Daeges, M., & Humphrey, L. (2016). Effectiveness of breast cancer screening: Systematic review and meta-analysis to update the 2009 U.S. Preventive services

- task force recommendation. *Annals of Internal Medicine*, 164(4), 244–255.
<https://doi.org/10.7326/M15-0969>
- Nelson, H. D., Pappas, M., Cantor, A., Griffin, J., Daeges, M., & Humphrey, L. (2016). Harms of Breast Cancer Screening: Systematic Review to Update the 2009 U.S. Preventive Services Task Force Recommendation. *Annals of Internal Medicine*, 164(4), 256–267. <https://doi.org/10.7326/M15-0970>
- The iMTA PAID tool version 3.0.* (n.d.). Retrieved March 19, 2023, from <https://imta.shinyapps.io/PAID3code/>
- Parkinson, B., & De Abreu Lourenco, R. (2015). Discounting in Economic Evaluations in Health Care: A Brief Review. *Cancer Research Economics Support Team*. www.crest.uts.edu.au
- Poulos, A., & Llewellyn, G. (2005). Mammography discomfort: a holistic perspective derived from women's experiences. *Radiography*, 11(1), 17–25. <https://doi.org/10.1016/J.RADI.2004.07.002>
- PriceWaterhouseCoopers Accountants N.V. (2022). *Jaarverslag Stichting Bevolkingsonderzoek Nederland 2021*.
- PYPL Popularity of Programming Language index.* (2022). <https://pypl.github.io/PYPL.html>
- R Core Team. (2017). *R: A Language and Environment for Statistical Computing*. <https://www.R-project.org/>
- Rafia, R., Brennan, A., Madan, J., Collins, K., Reed, M. W. R., Lawrence, G., Robinson, T., Greenberg, D., & Wyld, L. (2016). Modeling the Cost-Effectiveness of Alternative Upper Age Limits for Breast Cancer Screening in England and Wales. *Value in Health : The Journal of the International Society for Pharmacoeconomics and Outcomes Research*, 19(4), 404–412.
<https://doi.org/10.1016/J.JVAL.2015.06.006>
- Rainey, L., van der Waal, D., Jervaeus, A., Wengström, Y., Evans, D. G., Donnelly, L. S., & Broeders, M. J. M. (2018). Are we ready for the challenge of implementing risk-based breast cancer screening and primary prevention? *Breast*, 39, 24–32. <https://doi.org/10.1016/J.BREAST.2018.02.029>
- Rim, S. H., Allaire, B. T., Ekwueme, D. U., Miller, J. W., Subramanian, S., Hall, I. J., & Hoerger, T. J. (2019). Cost-effectiveness of breast cancer screening in the National Breast and Cervical Cancer Early Detection Program. *Cancer Causes & Control : CCC*, 30(8), 819–826.
<https://doi.org/10.1007/S10552-019-01178-Y>
- RIVM. (2022, August 22). *Population screening programmes*. RIVM. <https://www.rivm.nl/en/population-screening-programmes>
- RStudio Team. (2020). *RStudio: Integrated Development Environment for R*. <http://www.rstudio.com/>
- Schiller-Fruehwirth, I., Jahn, B., Einzinger, P., Zauner, G., Urach, C., & Siebert, U. (2017). The Long-Term Effectiveness and Cost Effectiveness of Organized versus Opportunistic Screening for Breast Cancer in Austria. *Value in Health : The Journal of the International Society for Pharmacoeconomics and Outcomes Research*, 20(8), 1048–1057. <https://doi.org/10.1016/J.JVAL.2017.04.009>

- Shiell, A., Donaldson, C., Mitton, C., & Currie, G. (2002). Health economic evaluation. *Journal of Epidemiology & Community Health*, 56(2), 85–88. <https://doi.org/10.1136/JECH.56.2.85>
- Sociaal-economische Raad. (2020). *VERKENNING 20/02 | Juni 2020 Zorg voor de toekomst - Over de toekomstbestendigheid van de zorg*. www.ser.nl
- Solbjør, M., Forsmo, S., Skolbekken, J. A., & Saetnan, A. R. (2011). Experiences of Recall After Mammography Screening—A Qualitative Study. [Http://Dx.Doi.Org/10.1080/07399332.2011.565530](http://Dx.Doi.Org/10.1080/07399332.2011.565530), 32(11), 1009–1027. <https://doi.org/10.1080/07399332.2011.565530>
- Strayer, D. S. (David S., Rubin, E., Saffitz, J. E., & Schiller, A. L. (2014). *Rubin's pathology : clinicopathologic foundations of medicine* (7th ed.). https://books.google.com/books/about/Rubin_s_Pathology.html?id=0yj_jgEACAAJ
- Swedish Council on Health Technology Assessment, French National Authority for Health, The Institute for Economic Research, & Institute for Quality and Efficiency in Health Care. (2015). Methods for health economic evaluations-A guideline based on current practices in Europe. *EUnetHTA*. https://www.eunetha.eu/wp-content/uploads/2018/03/Methods_for_health_economic_evaluations.pdf
- Tan, S. Y. G. L., van Oortmarsen, G. J., de Koning, H. J., Boer, R., & Habbema, J. D. F. (2006). The MISCAN-Fadia continuous tumor growth model for breast cancer. *Journal of the National Cancer Institute. Monographs*, 36, 56–65. <https://doi.org/10.1093/JNCIMONOGRAPHS/LGJ009>
- Tina Shih, Y. C., Dong, W., Xu, Y., & Shen, Y. (2019). Assessing the Cost-Effectiveness of Updated Breast Cancer Screening Guidelines for Average-Risk Women. *Value in Health : The Journal of the International Society for Pharmacoeconomics and Outcomes Research*, 22(2), 185–193. <https://doi.org/10.1016/J.JVAL.2018.07.880>
- Tower, H., Ruppert, M., & Britt, K. (2019). The Immune Microenvironment of Breast Cancer Progression. *Cancers 2019, Vol. 11, Page 1375, 11(9)*, 1375. <https://doi.org/10.3390/CANCERS11091375>
- Ucar, I., Hernández, J. A., Serrano, P., & Azcorra, A. (2018). Design and Analysis of 5G Scenarios with simmer: An R Package for Fast DES Prototyping. *IEEE Communications Magazine*, 56(11), 145–151. <https://doi.org/10.1109/MCOM.2018.1700960>
- UT-JupyterLab wiki [Jupyter wiki]*. (n.d.). Retrieved March 19, 2023, from <https://jupyter.wiki.utwente.nl/>
- Van Baal, P. H. M., Wong, A., Slobbe, L. C. J., Polder, J. J., Brouwer, W. B. F., & De Wit, G. A. (2011). Standardizing the inclusion of indirect medical costs in economic evaluations. *Pharmacoeconomics*, 29(3), 175–187. <https://doi.org/10.2165/11586130-000000000-00000/FIGURES/TAB3>
- van den Broek, J. J., van Ravesteyn, N. T., Heijnsdijk, E. A., & de Koning, H. J. (2018). Simulating the Impact of Risk-Based Screening and Treatment on Breast Cancer Outcomes with MISCAN-Fadia. *Medical Decision Making : An International Journal of the Society for Medical Decision Making*, 38(1_suppl), 54S-65S. <https://doi.org/10.1177/0272989X17711928>

- Van Goethem, M., Mortelmans, D., Bruyninckx, E., Verslegers, I., Biltjes, I., Van Hove, E., & De Schepper, A. (2003). Influence of the radiographer on the pain felt during mammography. *European Radiology*, 13(10), 2384–2389. <https://doi.org/10.1007/S00330-002-1686-6>
- van Haperen, V. R. (2018). *Factsheet Bevolkingsonderzoek Borstkanker*. <https://www.rivm.nl/sites/default/files/2019-02/Factsheet%20borstkanker%202018%20DEF.pdf>
- van Luijt, P. A., Heijnsdijk, E. A. M., & de Koning, H. J. (2017). Cost-effectiveness of the Norwegian breast cancer screening program. *International Journal of Cancer*, 140(4), 833–840. <https://doi.org/10.1002/IJC.30513>
- Vemer, P., Corro Ramos, I., van Voorn, G. A. K., Al, M. J., & Feenstra, T. L. (2016). AdViSHE: A Validation-Assessment Tool of Health-Economic Models for Decision Makers and Model Users. *Pharmacoeconomics*, 34(4), 349–361. <https://doi.org/10.1007/S40273-015-0327-2>
- Versteegh, M., Knies, S., & Brouwer, W. (2016). From Good to Better: New Dutch Guidelines for Economic Evaluations in Healthcare. *Pharmacoeconomics*, 34(11), 1071–1074. <https://doi.org/10.1007/S40273-016-0431-Y/METRICS>
- Wilson, J. M. G., & Jungner, G. (1968). PRINCIPLES AND PRACTICE OF SCREENING FOR DISEASE. *The Journal of the Royal College of General Practitioners*.
- Winter, L., & Velden, J. van der. (2022). *U Heeft Nog 25 Jaar*. S2 Uitgevers.
- World Health Organization. (2020). Screening programmes: a short guide. *WHO Regional Office for Europe*.
- Wright, S., Gray, E., Rogers, G., Donten, A., Hainsworth, R., & Payne, K. (2022). *MANC-RISK-SCREEN Version 1.0* (1.0). University of Manchester. <https://research.manchester.ac.uk/en/datasets/manc-risk-screen-version-10>
- Zielonke, N., Geuzinge, A., Heijnsdijk, E. A. M., Heinävaara, S., Senore, C., Jarm, K., de Koning, H. J., & van Ravesteyn, N. T. (2021). Extending Age Ranges in Breast Cancer Screening in Four European Countries: Model Estimations of Harm-to-Benefit Ratios. *Cancers*, 13(13). <https://doi.org/10.3390/CANCERS13133360>
- Zielonke, N., Gini, A., Jansen, E. E. L., Anttila, A., Segnan, N., Ponti, A., Veerus, P., de Koning, H. J., van Ravesteyn, N. T., Heijnsdijk, E. A. M., Heinävaara, S., Sarkeala, T., Cañada, M., Pitter, J., Széles, G., Voko, Z., Minozzi, S., Senore, C., van Ballegooijen, M., ... Priaulx, J. (2020). Evidence for reducing cancer-specific mortality due to screening for breast cancer in Europe: A systematic review. *European Journal of Cancer*, 127, 191–206. <https://doi.org/10.1016/J.EJCA.2019.12.010>
- Zorginstituut Nederland. (2015). *Rapport kosteneffectiviteit in de praktijk*.
- Zorginstituut Nederland. (2016). *Guideline for economic evaluations in healthcare*.

Appendix

Appendix I – Guidelines and Treatment

The flowchart describes the care path, including all possible treatment options for breast cancer in the Netherlands. People that have gotten a screening result of BI-RADS 0,3,4, or 5, will be referred to the hospital through their General Practitioner (GP). After positive diagnosis, a broad spectrum of treatment options is available. There are various surgical procedures that can be carried out, as well as chemo-, radio-, hormone-, and (neo)adjuvant systemic therapies. The proposed type(s) of therapy are based on the tumor stage, hormone response, age, and clinical fitness of the patient (Integraal Kankercentrum Nederland, 2020).

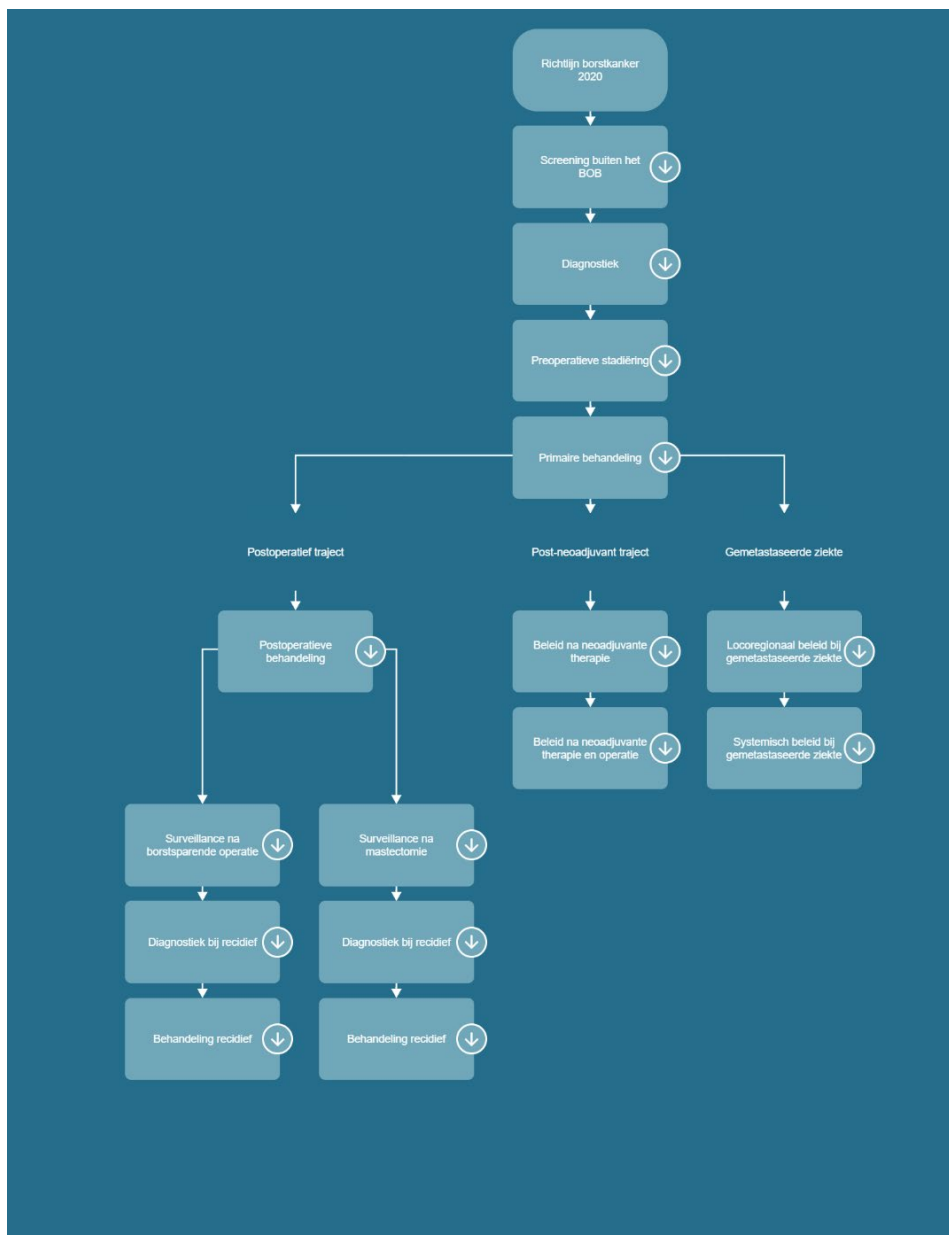


Figure 27: OncoGuide's Breast Cancer care pathway Flowchart

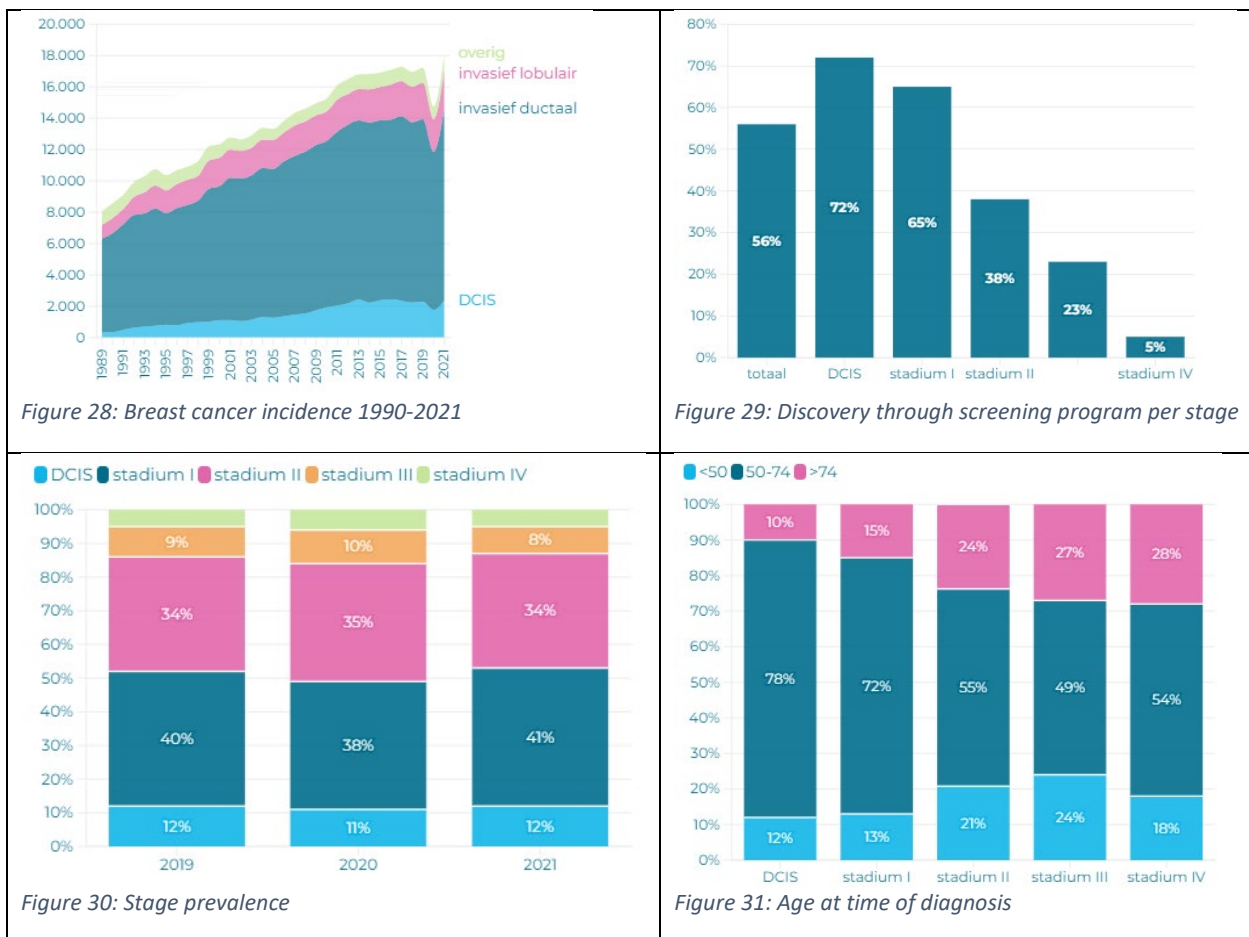
Appendix II – Breast Cancer

This chapter will provide more information on the characteristics and treatment of breast cancer. Basic knowledge, current statistics, and literary conclusions on screening programs are also covered. Furthermore, principles and methods for health economic evaluations are explored through the literature.

Breast Cancer Statistics

IKNL keeps the registry of all known Dutch breast cancer figures. They offer a page containing key statistics on their website. These figures will provide a baseline for the simulation in this research to work towards. If it is possible to replicate those figures, the simulation is working sufficiently for a base case analysis.

In 2021, 18.162 people in the Netherlands were diagnosed with breast cancer. 1 in 7 women will get breast cancer in their lifetime, and it is the most prevalent cancer in women (31% of all cancer). 1 in 27 women will die of breast cancer. 56% of all breast cancer diagnosis was through the screening program. Four graphics containing important statistics can be seen in Figure 28, 8, 9, & 10 (Integraal Kankercentrum Nederland, 2022)



The AJCC’s staging manual also contained tables showing average 5-year DSS per stage and risk profile. This can be seen below.(Amin et al., 2017)

Table 26: Overall survival and disease-specific survival from risk profile and stage

Stage (7th Edition)	Risk Profile	N	5-yr. DSS	95% CI	5-yr. OS	95% CI
I (IA and IB)	0	36	100%		97%	80.4%–99.6%
	1	1173	99.4%	98.7%–99.7%	96.7%	95.4%–97.0%
	2	274	98.8%	96.4%–99.6%	94.6%	91.0%–96.8%
	3	119	96.6%	91.1%–98.7%	93.8%	87.5%–97.0%
IIA	0	31	100%		96.8%	79.2%–99.5%
	1	634	99.4%	97.5%–99.8%	97.1%	94.7%–98.4%
	2	236	97.5%	93.2%–99.1%	94.1%	88.7%–97.0%
	3	98	91.0%	81.8%–95.7%	88.2%	78.5%–93.8%
IIB	0	11	100%		100%	
	1	309	96.9%	92.6%–98.8%	94.6%	89.6%–97.2%
	2	107	92.9%	83.6%–97.1%	89.3%	80.1%–94.4%
	3	40	91.5%	75.6%–97.2%	91.5%	75.6%–97.2%
IIIA	0	3	100%		100%	
	1	134	98.3%	88.2%–99.8%	91.5%	82.6–96.0%
	2	50	92.2%	77.2%–97.5%	90.3%	75.7%–96.3%
	3	7	68.6%	21.3%–91.2%	68.6%	21.3%–91.2%
IIIC	0	0				
	1	39	92.2%	72.1%–98.0%	84.4%	63.7%–93.9%
	2	16	80.8%	51.4%–93.4%	80.8%	51.4%–93.4%
	3	10	33.3%	6.3%–64.6%	33.3%	6.3%–64.6%

Note: There were insufficient numbers of cases with Stage IIIB cancer for analysis

The above information is essential for understanding breast cancer and the risk it poses on society. Proper treatment and diagnosis is crucial. It is therefore necessary to assess whether current methodologies are sufficiently efficient for treating this disease.

Risk Factors

Rubin’s Pathology, one of the most influential books that describes the clinicopathologic foundations of medicine, lists risk factors for breast cancer. They make a distinction between Modifiable and Non-Modifiable data. Some of the risks are quantified: deferring childbearing before age 35 could increase risk up to 2-3 times compared to an average. Smoking and alcohol consumption both increase relative risk with around 20% as well. Rubin’s presents the following table of risk factors: (Strayer et al., 2014)

Table 27: Risk factors for breast cancer development

Not Modifiable	Modifiable
Age	Body mass index
BRCA germline mutations	Diet
Family history	Alcohol
Chest radiation	Exogenous estrogen
Race/ethnicity	Exercise
Height	Smoking
Age at menarche	Reproductive history
Age at menopause	Age at first full-term delivery
Breast density	Lactation
Atypia on prior breast biopsy	

Individuals simulated in the model proposed for this research can have the risk factors as attributes. The model's screening scenarios could then be optimized through risk stratification based on these factors. The difference between medical and non-medical risk factors could be made, to allow for optimization based on publicly available data. Non-medical personal data is data that is available not only through a person's medical file but can be drawn from societal databases. Factors like Age, Race, BMI, Diet, Alcohol, Exercise, Smoking, and children from the table presented in Rubin's could be used. More factors could be identified through literature (Strayer et al., 2014).

Appendix III - Screening Principles & legal grounds and Evaluations

Screening Principles and Evaluations

In 1968, Wilson and Jungner published an article on screening for the World Health Organization. This marked the start of modern screening. The publication stated the initial definition and goals of screening and laid out 10 principles to assess if screening is the right way to improve public health. These ten principles are important, and screening measures should still match all these principles. Some of these principles are more difficult to assess than others, and it is unclear if all screening programs still match with these principles (Wilson & Jungner, 1968; World Health Organization, 2020).

1. The condition should be an important health problem.
2. There should be an accepted treatment for patients with recognized disease.
3. Facilities for diagnosis and treatment should be available.
4. There should be a recognizable latent or early symptomatic phase.
5. There should be a suitable test or examination.
6. The test should be acceptable to the population.
7. The natural history of the condition, including development from latent to declared disease, should be adequately understood.
8. There should be an agreed policy on whom to treat as patients.
9. The cost of case-finding (including a diagnosis and treatment of patients diagnosed) should be economically balanced in relation to possible expenditure on medical care as a whole.
10. Case-finding should be a continuous process and not a “once and for all” project.

The National Health Council (‘Gezondheidsraad’) has published summarizing principles of the existing framework (Gezondheidsraad., 2008):

- Screening should be focused on an important health problem
- Screening should be proven effective with benefits outweighing harm
- Screening should be reliable and valid
- Respect for autonomy should be central
- Resources should be used efficiently and explicit accountability in terms of cost-effectiveness and equity is required.

The Dutch house of representatives (‘tweede kamer’) requested research towards the effectiveness of screening in the Netherlands was published in 2014. It was stated that screening was effective, and that around 17 percent, or 31 prevented deaths per 100,000 women was attributable to population screening for breast cancer. It is stated that the effectiveness met the initial expectations, but the question whether the complete benefits outweigh the harms of screening remains unanswered (Health Council of the Netherlands, 2014b).

Currently, the PRISMA study (Personalised RISK-based MAMmascreening), is researching whether it is possible, economically or socially viable to adapt risk stratification based on breast tissue density, hormone, protein, and DNA tests. The PRISMA study focusses on optimization through medical data. It uses the MISCAN model for evaluating screening programmes (Rainey et al., 2018; Tan et al., 2006).

The most commonly used model for evaluating breast cancer screening is the MISCAN model, developed by the Erasmus MC in the Netherlands. Globally, there are more tools available. Commonly used models

are the Dana-Farber Model, the Erasmus MC MISCAN Model, the Georgetown-Einstein Model, the MD Anderson Model, the Stanford Model, and the Wisconsin-Harvard Model. All these models have been designed in the late 1990s – early 2000s and were initially based on data as old as 1890 up to 1990. Furthermore, there are several limitations in these models that could influence the realistic representation of the outcomes. Limitations could include continuous tumor growth rate and post-treatment utilities that are no longer up to date (Berry et al., 2006; Chia et al., 2004; Fryback et al., 2006; Lee & Zelen, 2003; Mandelblatt et al., 2006; National Cancer Institute, 2022; van den Broek et al., 2018). More information on these and other models is presented in Chapter 2.3.3.

Appendix IV - Screening Legal Grounds and Governmental Policies

Legal Grounds and Governmental Policies

The Netherlands has codified regulations on population health screening into law, into the so called ‘wet op het bevolkingsonderzoek,’ or law on population health screening (WBO). This law describes rules that screening programs should follow, as well as how and when they should be evaluated, and how an acceptable proposition for a screening program can receive a permit for implementation.

A request for a permit will be denied if the expected benefits of the screening program do not outweigh the risks for the researched population if it is scientifically unfounded or if it does not add to public health in general. A permit can be withdrawn if new scientific evidence shows that the added benefits of the screening program are less than initially expected and falls below the required threshold (*Wet Op Het Bevolkingsonderzoek*, 2021).

In 2020, the latest permit for the screening program for breast cancer was granted. This permit stated that the current screening program met all demands from the WBO and that the organizations performing the screening should act in compliance with the policy framework on screening programs for cancer, published by the RIVM. This framework from the RIVM shows more policies that the screening programs should follow. The policies range from public duties as defined in the constitution to quality monitoring, and from optimisation policies to public values (D.J. de Leede & R. van Velzen, 2022; *Beschikking van de Staatssecretaris van Volksgezondheid, Welzijn En Sport, Houdende Verlenging van de Vergunning Op Grond van de Wet Op Het Bevolkingsonderzoek Voor Het Bevolkingsonderzoek Naar Borstkanker*, 2020).

Relevant aspects from the policy are the principles of screening programs. There are three main principles: the Wilson & Jungner criteria, public values, and cooperation in the health care chain. The principles for screening as defined by Wilson and Jungner were discussed in the previous paragraph. The public values that are defined are quality, accessibility, and affordability. Cooperation in the healthcare chain defines the transition between the screening program and traditional healthcare (D.J. de Leede & R. van Velzen, 2022).

The public values, quality, and affordability, are of interest for this research. Literature shows that there is unclarity on the effectiveness, and thus quality and affordability of the screening program. These unclaritys themselves will be discussed in **chapter 1.3**. Quality means that programs are effective: the test characteristics itself, the participation of the target population and the benefit to public health should be sufficient. There are no definitions for what exactly qualifies as sufficient. The value affordability states that the total costs of the program should be insightful, so the government can balance the use of public resources deployed against their use for other government tasks. The programs should also be offered at the lowest possible cost to realize the expected quality, and programs should be cost-effective. No explicit definition of cost-effectiveness is given (D.J. de Leede & R. van Velzen, 2022).

Further research learns that standards on cost-effectiveness have been advised in 2015. The Dutch healthcare institute has, together with the ministry of public health, wellbeing, and sports, defined three criteria of acceptable costs. These can be seen in the table below

Burden of disease

Maximum additional cost (€) per QALY

0.1 – 0.4	€ 20,000
0.41 – 0.7	€ 40,000
0.71 – 1.0	€ 80,000

The acceptability of costs for care in the Netherlands largely depend on Quality-adjusted Life Years, which will be further discussed in 1.3. The idea is that new interventions are only accepted if they deliver a certain quantifiable healthcare gain at a maximum cost threshold, where interventions for serious disease are allowed higher costs, as the societal impact of helping these patients is perceived to be higher (Zorginstituut Nederland, 2015).

From these documents, it appears that cost-effectiveness for healthcare in the Netherlands is an important metric in order to gain approval for a new intervention. This combined with the WHO's advice on reevaluating current screening methods, means that the breast cancer screening program in the Netherlands should be health-economically reevaluated. The best way for evaluating this would be through a cost-effectiveness analysis, with emphasis on how these interventions impact current healthcare budgets.

Appendix V - Drawbacks and Bias

Drawbacks and Bias

There are some biases and drawbacks that make it difficult to assess the effectiveness of screening programs. As stated before, the principles of Wilson and Jungner, as well as the Health Council state that costs should be balanced in comparison to other healthcare spending per health gain. However, the Health Council also stated that this is still an aspect that has not been adequately assessed for breast cancer screening (Health Council of the Netherlands, 2014b). The following drawbacks and biases are therefore important to recognize in order to create a model that does allow for accurate assessment.

- Data quality
- Data size
- Large incidence at start of screening
- Mortality comparison over time
- Lead time bias
- Length time bias

An explanation of these problems, an idea of how to deal with this, and how is actually dealt with this in this research are given in the following subsections.

Data quality

To determine if a screening should be implemented or whether it is effective, data regarding incidence and mortality is needed. This data relies on the reporting of disease and conditions and registration of causes of death. Furthermore, there often needs to be a range of historical data for accurate assessment. Medical data is notoriously hard to collect and store, and data from a long time ago might not be accurate enough to validate current models on (World Health Organization, 2020).

By simulating patient characteristics, disease incidence and mortality, and other factors, based on just recent data, it can be assumed that that generated data is an accurate reflection of the population at this time. There is no need of using historical data from decades ago, as using this data could lead to inaccurate reflection of reality. Nowadays, there are better treatment options and more accurate diagnostics techniques that improve overall disease survival and patient utility. This model used the most recent available data to replicate reality as accurate as possible. The model validation section shows that results from the model create an accurate reflection of reality.

Data size

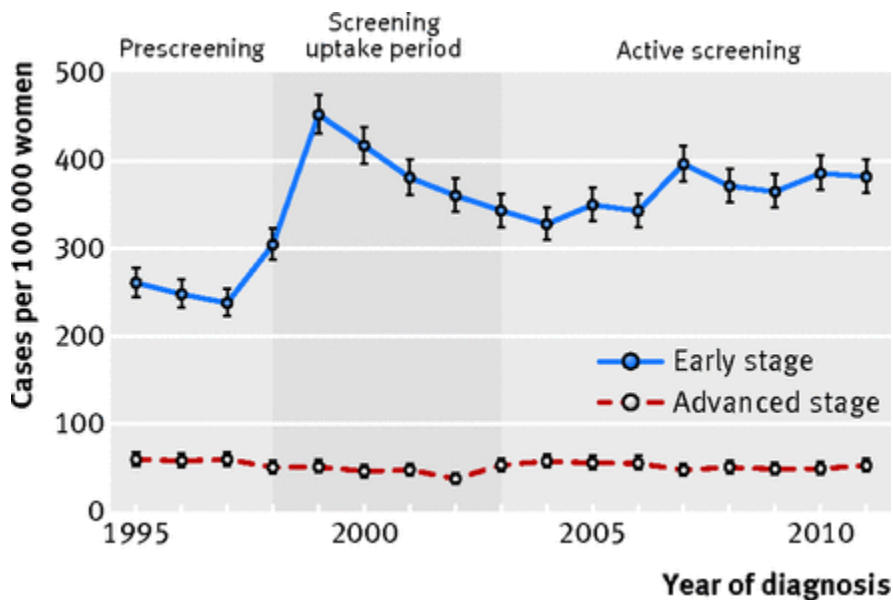
Often, data to be used in studies is limited. This is due to a limited number of study participants, not enough data availability or just that the newly implemented screening method has not been available long enough to gather enough data. Changes between different years could then push the confidence intervals too wide to draw an accurate conclusion (World Health Organization, 2020).

By generating accurate, representative patient data from scratch, practical data size is only limited by computing power. If the generated data is sufficiently accurate, it could reduce variability by increasing sample size. The model for this research is able to simulate populations of any given size, with population sizes of 100,000 giving an accurate representation of the population at acceptable computing times. By using similar random seeds whilst testing different strategies, the initial conditions

of all simulations are the same and patient experiences will differ only based on the type of strategy tested.

Large incidence at start of screening

If a new screening program is introduced, the incidence in the first year is often significantly higher than in the next years. This is called the prevalent or first-round effect. An example of this can be seen in the graph below. The graph represents the breast cancer incidence for women between 70 and 75 in the Netherlands. In 1998, the screening program was extended to include women up to 75 years old. This initially caused a significant spike in incidence, before settling to a stable situation just years later (World Health Organization, 2020).



(De Glas et al., 2014)

The practical use of simulations shines in this aspect as well. By being able to implement incidence rates, an analysis through simulation can give a fair overview of the current situation under current screening measures, as well as a view back in the past if screening methods were implemented earlier or for different criteria. In this simulation, patients are followed from birth to death, with screening starting at a similar age for all patients in the simulation. This starting age differs over the experiments, but there will be no large incidence spike for the entire population, as screening strategies are always tested over an entire lifespan.

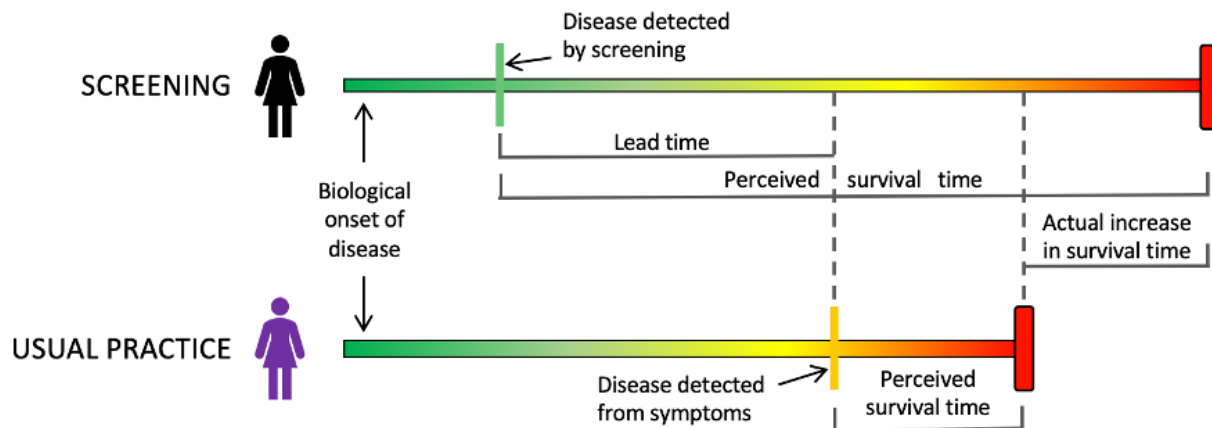
Mortality comparison over time

An improvement in mortality rates over time could be attributable to more than the implementation of a screening program. The improvement in treatment and other diagnostic techniques are also paramount factors in the decrease in disease-related deaths. It is often impossible to orchestrate accurate randomized trials to evaluate screening methods, as the implementation of this is seen as not ethical and it could take decades before enough meaningful data is gathered (World Health Organization, 2020).

When using a microsimulation, the simulation can be repeated with patients with the exact same characteristics. One simulation would involve a screening, and another would not. The results are then exactly comparable. Mortality rates would not differ due to better treatment, as the mortality, utility and expected life years used would be the same for both patients and only based on the stage of the disease at time of diagnosis.

Lead time bias

If extended survival for patients after screening is not corrected for lead time, the longer perceived survival is not necessarily proof of a benefit of early detection. This is called lead time bias. An example of this can be seen in **image**.

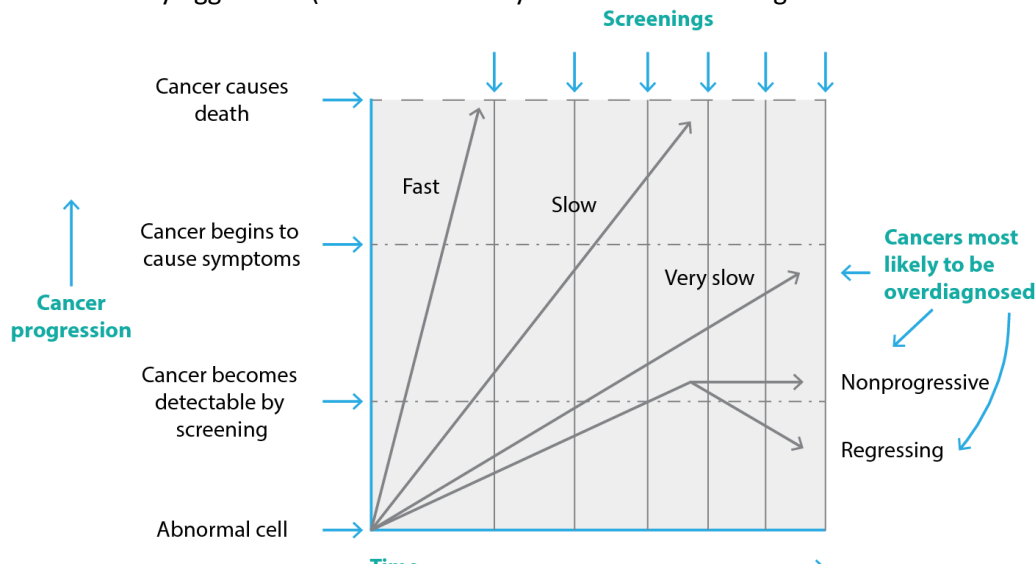


Clinical diagnosis of a disease or condition occurs when a patient exhibits a set of symptoms. People with diseases that are discovered through population screening get a diagnosis sooner, often before symptoms start to show. As a result, estimates of the differences in survival time between those with disease diagnosed through screening and those whose disease is discovered after symptoms appear can be skewed. This is because those with disease detected through screening will appear to have a longer survival time, even if early detection has no impact on the course of the disease and also if survival time is prolonged (*Lead Time Bias - Catalog of Bias, n.d.*).

Length time bias

Screening programs have a higher likelihood of identifying slow-growing, less aggressive cancers than they do faster-growing, more aggressive cancers. A slow-growing cancer may remain dormant in the body for a long period of time. It is therefore more likely to be present during screening. In contrast, a rapidly spreading, more aggressive cancer will likely cause symptoms, which will cause the patient to seek medical attention and receive a clinical diagnosis between regular scans. These "interval cancers" don't indicate that the screening program was ineffective; rather, they show that a small percentage of

cancer is very aggressive (more so than any reasonable screening schedule could catch).



(S. M. Carter & Barratt, 2017)

Screening evaluation studies should take this length time bias into account when researching the effectiveness of the program. Previous evaluation models, like the MISCAN model, uses continuous tumor growth rates. Invasive breast tumors are initiated and are assumed to have a constant growth rate, which differs between tumors. This variation is necessary to account for the length time bias as described above (van den Broek et al., 2018).

This research uses various tumor growth rates, with the possibility for spontaneous regression and growth stagnation. Tumor growth rates are compared to those of other CISNET models and finetuned accordingly. This implementation allows fast- and slow growing tumors to have appearances in a simulation, with various probabilities of detection through screening.

Screening programs have the noble goal to improve population health, and if properly carried out, this can also be the result. For example, breast cancer screening in the Netherlands is assumed to prevent around 850 deaths annually. To create a successful screening program, it is important that Wilson and Jungner's ten principles of screening are adhered to, as has also been recommended by the WHO. Screening programs should be evaluated periodically, for which there are multiple ways. These analyses also need to keep in mind a range of commonly seen biases to prevent overvaluing the screening program.

Health Economic Analyses of Breast Cancer Screening Programs through Simulation: A Systematic Literature Review

G.J.W. van Weezel - February 9, 2023

University of Twente, Industrial Engineering and Management, Supervisor: Dr. Ir. A.A.M. Spil

ABSTRACT: Breast cancer is a serious disease that 1 in 7 women will experience during their lifetime. To reduce disease mortality and diagnose the disease earlier so that treatment may be administered sooner, screening programs are implemented. These claims are however disputed by literature. This systematic review will analyze the state-of-the-art health economic analyses that used simulation models to determine the cost-effectiveness of breast cancer screening programs. For this systematic review, articles from 2016 onwards will be used from databases PubMed and Scopus. The article is based on the PRISMA guidelines. The first query resulted in 74 papers, which were scanned on title and abstract. Scanning the articles resulted in a total of 10 articles that were included in this systematic review. The type of simulation most often used is discrete event. Various screening strategies were tested, with bi-annual screening between 50 and 75 as the most common. The evaluated studies performed health economic analyses through cost-effectiveness, cost-utility, budget impact, harm-benefit and comparative effectiveness analyses. Outcomes in cost-effectiveness and cost-utility varied between €5000 and €53000 per QALY in the analyzed studies.

Key words: Breast Cancer, Screening, Mammography, Health Economic Analysis, Simulation, DES

1 INTRODUCTION

Breast cancer affects around one in every seven women over their lifetime, and one in every 27 women dies from the disease [1]. Breast cancer screening programs (BCSP) are implemented in many healthcare systems to speed up the diagnosis of smaller, asymptotic breast cancers, reduce cancer mortality, and increase overall utility of the population. Mammography is used for most women starting at middle age, while screening recommendations and practices vary by area [2]. The WHO European Region appears to be moving more toward increased health monitoring and screening for noncommunicable illnesses. In some cases, there isn't enough proof for effectiveness. The potential harm of screening, its cost and burden on the health system, and the necessity of rigorous quality assurance are issues that policymakers, healthcare professionals, and the general public frequently appear to be oblivious of [3]. Health Economics provides tools for decision making in healthcare to promote fair resource use. Health economic analyses (HEA) evaluate costs, harms, and benefits of interventions, with the goal to provide the maximum healthcare benefits with constraints on resources [4]. By means of a systematic review, the aim of this study is to assess the processes and outcomes of previous health economic evaluations on breast cancer screening programs. For this systematic review, the following research question was formulated:

- What is the cost-effectiveness of the breast cancer screening program?

Three sub-questions were formulated to answer the main question and to provide context on the methods for performing a health economic analysis on the breast cancer screening program.

- What are health economic analysis methods used to evaluate the cost-effectiveness of breast cancer screening programs?
- What are parameters used in health economic analysis on the cost-effectiveness of breast cancer screening programs?
- What are the reported outcomes of health economic evaluations for breast cancer screening programs?

The systematic review aims to answer these questions to provide an overview of and comparison between HEA methods, parameters and outcomes. These outcomes can then be used for a further, independent HEA on the BCSP.

2 METHODS

This section presents the method of data processing and selection of the found material used for this systematic literature review. This systematic review was performed according to preferred reporting items for systematic reviews and meta-analyses guidelines (PRISMA) [5]

2.1 Data Sources and Search method

For this review, a systematic search is done in the databases of Scopus and PubMed. The search aims to find health economic analyses on breast cancer screening programs using simulations or models. The articles covering this subject are selected between the years 2016 and 2022. The search was based upon the following topics:

- Health economic analysis and/or cost-effectiveness
- Breast cancer screening and/or mammography screening
- Simulation and/or model

The combination of these topics resulted in the following search query shown in Appendix A. The query is used for the systematic literature review and given as input for the selected databases. In these databases, the search is applied for the title, keywords, and abstract. The resulting list of articles is sorted by relevance. The choice was made to include 'Simulation and/or model' in the search query, as leaving this out would result in a significant multitude of results, most of which were focused on health economic outcomes of novel breast cancer screening methodologies, whereas this study focusses on the HEA of the current methods.

2.2 Study selection

One author (van Weezel) independently reviewed the article titles and screened for duplicates. After the duplicate extraction, the same authors reviewed the remaining articles by the abstract and keywords. During the abstract screening, any inconsistencies from previous steps were solved with a full-text screening. Further discrepancies were solved by full-text analysis.

2.3 Inclusion and exclusion criteria

In this systematic review paper, only papers including the keywords Cost-effectiveness, breast cancer screening and simulation are included. However, some synonyms of these keywords were also added to prevent missing out on relevant literature. These can be found in 1. Articles before 2016 were excluded from this search because of the time relevance. Papers not describing specific methods and outcome measures relevant for this research were excluded. Lastly, case studies and systemic review papers were also excluded due to their irrelevance for writing systematic review paper.

2.4 Data extraction and Synthesis

The results gained from the databases, based on the in- and exclusion criteria, were checked by the author by reading the abstract of each article. Then the author decided which articles were relevant. This resulted in ten articles that provided relevant information for this systematic review. These articles were read, and data were extracted according to a form. The publications were reviewed by one person. The following data were extracted from all publications:

- Basic information
 - Title
 - Authors
 - Year of publication
 - Country of research
 - Country of data
- Research question/objective
- Main screening strategy
- Target population
- Type of HEA
- Type of simulation
- Outcome measures
- Cost outcome metric
- Costs used
- Utilities used
- Tumor growth model

Extracting this data allows separation of the findings into different categories, e.g., cost estimations from papers that use Quality Adjusted Life Years (QALYs) can be compared to those that use Life Years Gained (LYG).

2.5 Risk of Bias assessment

The databases used for this systematic literature review are PubMed and Scopus. However, many more databases could have been included in the research. Leaving these databases out makes the risk of leaving out relevant sources more significant. Next, there is the possibility of publication bias where studies showing adverse outcomes are not published. Another bias is the personal bias of the researcher, which is a relevant bias as the research was performed alone. The use of PRISMA guidelines for performing this systematic review is a personal choice as the author worked with this method before. Using this method might introduce a bias, which could result in different a paper selection and results than when another systematic review methodology was followed.

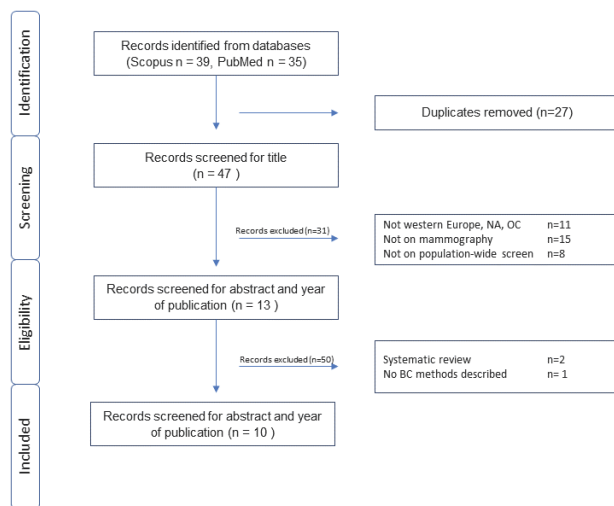


Figure 1: PRISMA flow diagram showing the study selection for this systematic review

Table 1: Synonyms of the used keywords

Cost-effectiveness	Breast cancer screening	Discrete event simulation
Cost benefit	Mammography screening	Simulation
Cost utility	Breast screening	Markov model
Cost of illness		Population model
Budget impact		Cohort model
Health economic		
Value for money		

3 RESULTS

3.1 General characteristics

In table 2 the characteristics of the ten selected articles are displayed. The articles are all published in 2016 or later. The locations were split in European Union and North America (the paper from the UK was from 2016, and is therefore classified as EU). The research methods and outcomes of these papers differs in type of HEA, the main reported outcome, types of costs, utilities, and tumor growth model used. Most articles are researching the costs per non-quality-adjusted or quality-adjusted life year, whereas one article only compares the harms and benefits, and not the costs. The table displays the exact number of publications per category. The most used HEA is a cost-

effectiveness analysis resulting in costs per QALY.

3.2 Study results

The results of the studies are analyzed and summarized in four tables. The first table, 3, discusses the type of health-economic analysis used in the research. The second table, 4, discusses the costs used and the third table, 5, discusses the utilities used in the different papers. The final table, 6, summarizes the results of the papers in the cost per outcome, as well as the type of simulation and the tumor growth model used. Between the articles, there are many similarities. Nine out of ten articles choose a cost-effectiveness analysis as health economic analysis. Nine out of ten articles choose a discrete-event microsimulation as simulation. Even though the start of many research is

Table 2: General characteristics of the reviewed articles

Characteristics	Value	N
Year of publication	2016	2
	2017	2
	2018	3
	2019	2
	2021	1
Location of research	European Union	6
	North America	4
Location of Data	European Union	5
	North America	4
	Other (Norway)	1
Type of HEA	Cost-Effectiveness	9
	Budget Impact	2
	Harm-benefit	1
Main reported outcome	cost/LYG	4
	cost/QALY	6
	other	1
Costs considered	Screening	9
	Diagnosis	9
	Treatment	9
	Societal	1
Utilities considered	Per disease stage	6
	Per event	2
	per adverse effect	1
Tumor Growth Model	Not defined	2
	Markov Model	1
	Growth equation	7

similar, the outcomes vary a lot. The results can be seen in 2. All values were corrected for inflation and converted to 2023 Euros.

4 ANALYSIS OF THE RESULTS

4.1 Cost vs benefit

There is no clear consensus visible on the costs per QALY or per LYG. There is also no statistical significance in whether the calculations for cost per life year gained are higher than the costs per quality adjusted life year gained. The results vary too much to state that screening programs fall above the willingness-to-pay threshold. The Willingness to pay threshold in the Netherlands varies between 20,000 and 80,000 euro per QALY, depending on the severity of the disease [16]. None of the studies included all cost criteria from the Dutch guidelines for health economic evaluations [17].

4.2 Cost vs. region

It appears that screening programs in North America cost more per gained (quality adjusted) life year than in the European Union. This could be due to different health care expenditure in general, qualitatively differing screening techniques, or evaluations that took more costs into account.

4.3 Use of cost measures

Most papers only use the costs of screening, diagnosis, and treatment as inputs to calculate the total cost. Only the paper from Shih [8] additionally used the productivity loss due to early death in their cost calculation. None of the papers provide societal cost estimates from productivity loss for visits to screening facilities. The paper from Schiller [14] notes that the cost estimate is from a societal perspective, and therefore only uses health care related costs that are paid by society. The paper from Arrospe et al. makes an interesting note on the costs of implementation. It states that the costs per gained life year are a difficult indicator, as you cannot implement this intervention for just one patient: in order to create an effective screening program, it needs to be implemented for the entire population. This will result in a significant annual budget impact for the health care system. Other interventions for disease often only incur costs for people that are ill, resulting in lower total costs, even at higher costs/QALY rates [12].

4.4 Use of utility measures

The six papers calculating the added QALYs did this differently. In two papers, minor reductions to the current quality of life (QoL) were given during screening and diagnosis

Table 3: Type of Health Economic Analyses performed in each paper

Main outcome measure	Paper	Type of Health Economic Analysis				
		Harm-benefit	Cost-Effectiveness	Cost-Utility	Budget Impact	Effectiveness
Harm/benefit	Zielonke 2021 [6]	X				
€/QALY	Rim 2019 [7]		X	X		
	Shih 2019 [8]		X	X		
	Mittmann 2018 [9]		X	X	X	
	Rafia 2016 [10]		X	X		
	v.Luijt 2017 [11]		X	X		
€/LYG	Arrospide 2016 [12]		X	X	X	
	Carter 2018 [13]		X			X
	Schiller 2017 [14]		X			
	Koleva 2018 [15]		X			

Table 4: Costs used in the health economic analyses

Main outcome measure	Paper	Costs Used			
		Screening	Diagnostics	Treatment	Social Burden
Harm/benefit	Zielonke 2021				
€/QALY	Rim 2019	X	X	X	
	Shih 2019	X	X	X	X
	Mittmann 2018	X	X	X	
	Rafia 2016	X	X	X	
	v.Luijt 2017	X	X	X	
€/LYG	Arrospide 2016	X	X	X	
	Carter 2018	X	X	X	
	Schiller 2017	X	X	X	
	Koleva 2018	X	X	X	

Table 5: Utilities used in the health economic analyses

Main outcome measure	Paper	Utilities Used			
		Cancer stage	Screening	Diagnostics	Notes
Harm/benefit	Zielonke 2021				NA
€/QALY	Rim 2019				Decremental after detection
	Shih 2019	X			
	Mittmann 2018	X			
	Rafia 2016	X	X		decrease for stress
	v.Luijt 2017	X	X	X	
€/LYG	Arrospide 2016	X			Different Age ranges
	Carter 2018				NA
	Schiller 2017				NA
	Koleva 2018				NA

Table 6: Cost, simulation, and tumor growth characteristics of the reviewed articles

Main outcome measure	Paper	Euro / outcome			
Harm/benefit €/QALY	Zielonke 2021		Microsim (MISCAN)	Exponential	
	Rim 2019	53918	Microsim (Stanford)	Constant	
	Shih 2019	43087	Microsim (MDACC)	Exponential	
	Mittmann 2018	32134	Microsim (Wisconsin)	Gompertz	
	Rafia 2016	21266	Microsim	Exponential	
	v.Luijt 2017	12445	Microsim (MISCAN)	Exponential	
	Arrospide 2016	5117	Microsim	Progressive stochastic	
	€/LYG	Carter 2018	28415	Monte carlo (Treeage)	Gompertz
		Schiller 2017	23918	Microsim	Markov state transition
		Koleva 2018	6582	Microsim (SiMRiSc)	Exponential
Koleva 2018					

[8, 11]. One of the papers updates the QoL for a fixed duration after a certain treatment is received, before returning to a baseline [8]. One of the papers uses an initial decrease after diagnosis and then a continuous decrease per year for the duration of the disease [7]. The paper from Rafia [10] updates the QoL values of patients at the time that the disease progresses to the next stage, rather than after diagnosis. Furthermore, it adds temporary disutility during the screening progress and during an 'anxiety phase' when a patient is awaiting the results of diagnosis. Four out of six papers updated the QoL of the patients at the time that the disease clinically progressed to the next stage, before the patient is notified or diagnosed [9, 12, 10, 11].

4.5 Tumor growth models

Very little information can be found on the initiation of the tumor growth models. Modelling primary tumor growth is important for determining screening diagnosis parameters. As for the tumor growth models, there seems to be consensus on an initial exponential growth phase, with some papers choosing to flatten the curve at a certain tumor size through a Gompertz equation, representing empirical evidence about slowing growth after having reached a certain size. Few papers take into account the possibility of naturally occurring regression [11, 14]. The other papers opt for a constant growth rate at which a tumor keeps growing, albeit constant or entered in an equation determining the size according to a function.

4.6 HEA Methods

Nine out of ten studies analyzed in this systematic review covered a cost-effectiveness-focused health economic analysis. Only one study performed a harm-benefit analysis [6]. Of these nine studies, all nine performed a Cost-Effectiveness analysis with the outcome focus on the cost per gain in life years or quality adjusted life years. These

CEAs were all performed through discrete event simulation (DES). Of the DES models used, 5 papers used a model listed on the CISNET model registry [18]. One of the models was reused from another paper, one model used a ready-to-use computer program and two models were developed specifically for the study. DES seems to be the preferred modelling method for health economic analyses, as other options for modelling were included in the search query and named in the resulting papers, but were not actually used in any of the research analyzed.

4.7 Harm versus benefit

The 2021 paper from Zielonke et al. [6] is the only included paper that performed a HEA on the harms and benefits of breast cancer and did not include a cost calculation. The paper focused the effects of changing the age ranges of the screening program. The increase in BC deaths averted and life years gained was compared to the increase in over-diagnoses and false positives. The research concludes that the current age ranges of 50-75 are optimal for the Netherlands, but that further research to extend the harm-benefit analysis to a cost-effectiveness analysis is necessary to advice on health policy making. These further studies should include additional screening effects and treatment related effects in the quality of life calculation.

5 DISCUSSION

None of the papers calculated the societal burden associated with people visiting screening facilities, despite the possible substantial impacts on the total costs at this scale. This constitutes a significant gap in our understanding, which requires further consideration in implementation and research efforts. In this regard, the Dutch health economic analysis guidelines emphasize the importance of determining societal costs and short-term burdens in such evalua-

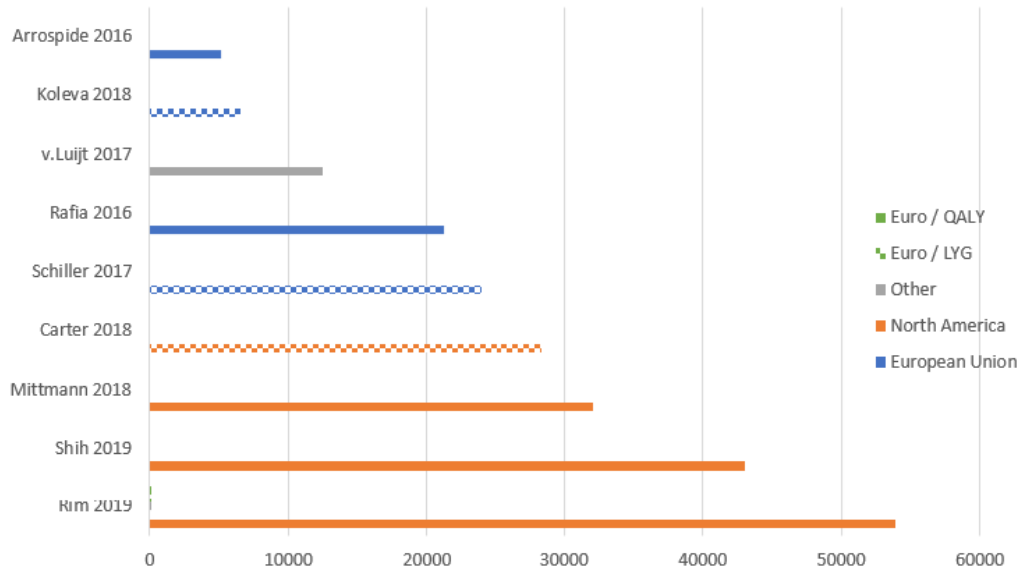


Figure 2: Costs per metric - results of the nine papers listing these outcomes

tions. This view is reinforced by the suggestions made by Carter et al. and Van Luijt et al. [11, 13], who call for more comprehensive analysis of the negative effects of screening and societal burdens in future research.

The way the QALYs are calculated also vary between the papers, and it seems difficult to find consensus on how to calculate this. Papers including more disutility for calculating the QALY, appear to end up on the higher ends of the cost per QALY for that region. Rafia [10] concludes the highest cost/QALY calculation for the EU, and Rim and Shih the highest in NA, both using more QoL adjustments than their counterparts [7, 8]. However, there is too little evidence to conclude causal correlation at this point, due to fractured data from different regions that makes it difficult to draw meaningful comparisons and it remains unclear why there is so much variation in the calculation of the costs per (quality adjusted) life year between the various papers.

Four out of six papers updated the QoL of the patients at the time that the disease clinically progressed to the next stage, before the patient is notified or diagnosed. Dutch guidelines suggest using QoL values of patients at different stages, and these QoL values are only measured after diagnosis. It is unclear if having received the diagnosis lowers a patients perception of his own quality of life. This effect is unaccounted for in research, but presumed to exist in practice.

As for the tumor growth models, the Gompertz functions seem to best fit the natural history model of observed tumor sizes but lacks a regressive component. Further research could include adding the possibility for random regression and growth stagnation at times, which would make for more realistic growth patterns. This is suggested by the paper from Carter et al. in regards to overdiagnosis and false positives. A better understanding of tumor growth and non-progressive tumors would lead to better implementation of overdiagnosis in models. [13]

The studies agree on using discrete event simulation for health economic analysis, and focus on cost-effectiveness in general. The adaptability and expandability of these DES models makes them versatile and often models can be reused and adapted for other research, as shown by studies adapting CISNET models.

No papers from later than 2021 were included, and no papers later than 2019 showed quantitative cost data. Papers in this category could not be found with this search query, and it is therefore assumed that research on the cost-effectiveness of breast cancer screening programs has not taken place or has not been published.

The different papers are based on different data. This makes the results difficult to compare. Costs of the screening program itself, treatment of a disease, normal life expectancy, cancer incidences and screening effectiveness can still differ between these countries. Even though this

review attempted to limit this by only including papers from Western Europe and North America, it cannot be said that the data used for these analyses is interchangeable, and conclusions on one country do not necessarily apply to others.

6 CONCLUSION

Screening has the potential to increase the wellbeing of a population. According to current literature, these health benefits fall within the conventional cost-benefit parameters. However, different research uses a lot of different methods for the approach and calculation of these harms and benefits, and none of the papers manages to catch all factors described in these papers in one research. Furthermore, research from different countries cannot be used to draw conclusions for other countries. Further research is necessary to create a sounding conclusion on the cost-effectiveness of breast cancer screening, factoring in different possibilities of tumor growth, different QoL measures at various points in a patient's life and summing up all costs incurred over a patient's life that are attributable to the screening program. It is a difficult decision to determine what factors should be counted and how they should be counted if they can be counted at all, but progress on determining the cost-effectiveness of screening programs can be made by following (inter)national guidelines more explicitly.

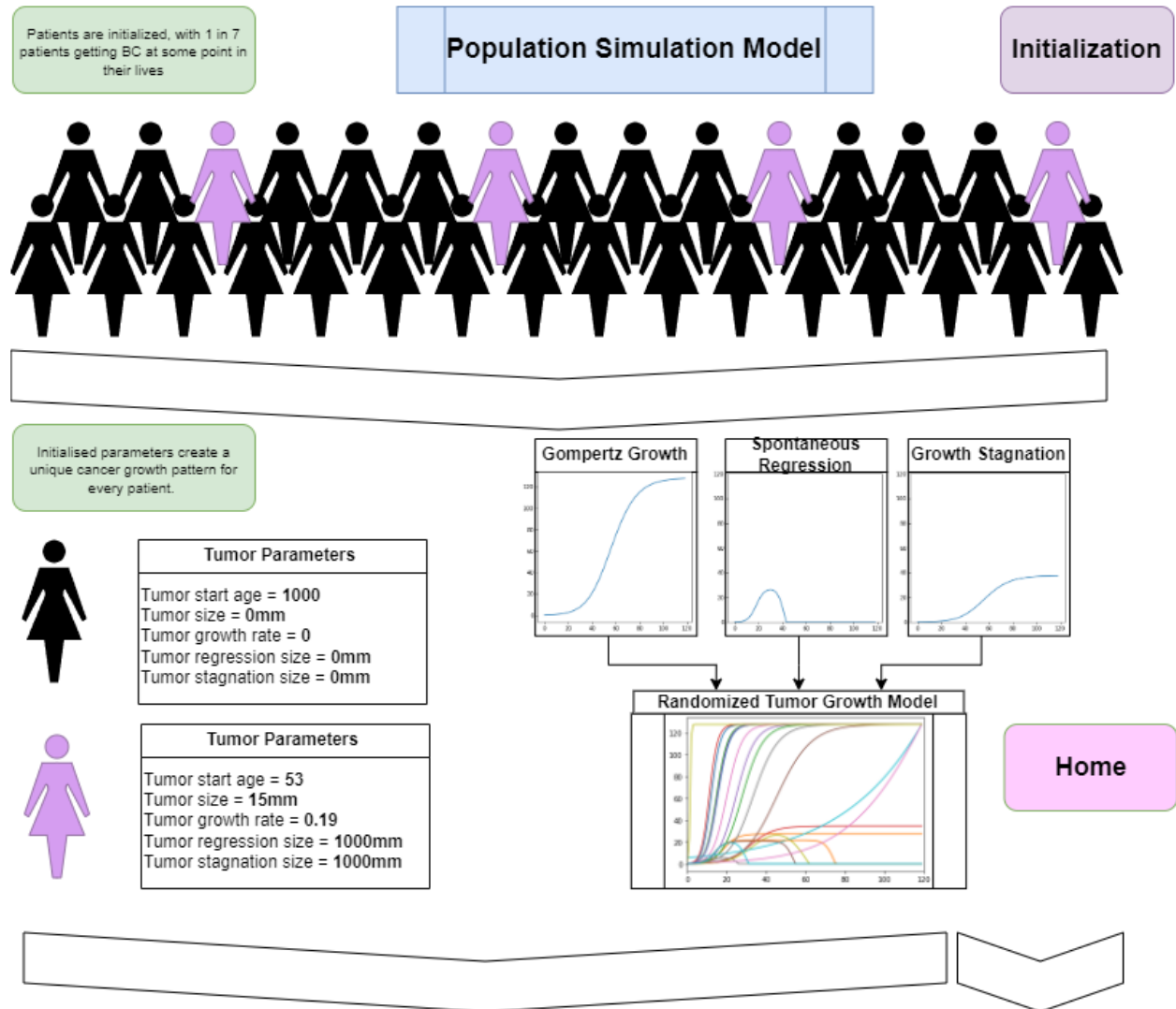
REFERENCES

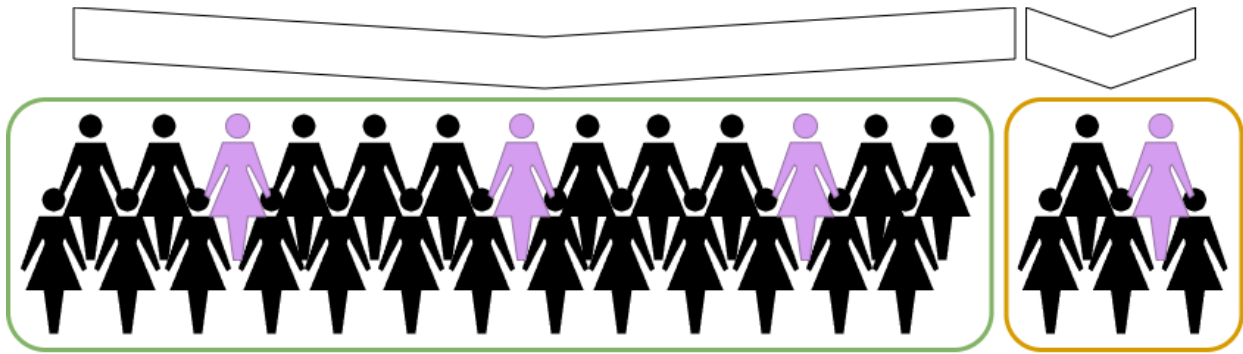
1. Integraal Kankercentrum Nederland, Borstkankercijfers.
2. A.K. Clift, D. Dodwell, S. Lord, S. Petrou, S.M. Brady, G.S. Collins, and J. Hippisley-Cox, The current status of risk-stratified breast screening, *British Journal of Cancer*, (2022), 126(4):533–550.
3. World Health Organization, Screening programmes: a short guide, *WHO Regional Office for Europe*, (2020).
4. Economic Evaluation Overview — POLARIS — Policy and Strategy — CDC.
5. David Moher, Alessandro Liberati, Jennifer Tetzlaff, Douglas G. Altman, Doug Altman, Gerd Antes, David Atkins, Virginia Barbour, Nick Barrowman, Jesse A. Berlin, Jocelyn Clark, Mike Clarke, Deborah Cook, Roberto D'Amico, Jonathan J. Deeks, P. J. Devereaux, Kay Dickersin, Matthias Egger, Edzard Ernst, Peter C. Gøtzsche, Jeremy Grimshaw, Gordon Guyatt, Julian Higgins, John P.A. Ioannidis, Jos Kleijnen, Tom Lang, Nicola Magrini, David McNamee, Lorenzo Moja, Cynthia Mulrow, Maryann Napoli, Andy Oxman, Bá Pham, Drummond Rennie, Margaret Sampson, Kenneth F. Schulz, Paul G. Shekelle, David Tovey, and Peter Tugwell, Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement, *PLOS Medicine*, 7 2009, 6(7):e1000097.
6. Nadine Zielonke, Amarens Geuzinge, Eveline A.M. Heijnsdijk, Sirpa Heinävaara, Carlo Senore, Katja Jarm, Harry J. de Koning, and Nicolien T. van Ravesteyn, Extending Age Ranges in Breast Cancer Screening in Four European Countries: Model Estimations of Harm-to-Benefit Ratios, *Cancers*, 7 2021, 13(3360):3360.
7. Sun Hee Rim, Benjamin T. Allaire, Donatus U. Ekwueme, Jacqueline W. Miller, Sujha Subramanian, Ingrid J. Hall, and Thomas J. Hoerger, Cost-effectiveness of breast cancer screening in the National Breast and Cervical Cancer Early Detection Program, *Cancer causes & control : CCC*, 8 2019, 30(8):819–826.
8. Ya Chen Tina Shih, Wenli Dong, Ying Xu, and Yu Shen, Assessing the Cost-Effectiveness of Updated Breast Cancer Screening Guidelines for Average-Risk Women, *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*, 2 2019, 22(2):185–193.
9. Nicole Mittmann, Natasha K. Stout, Anna N.A. Tosteson, Amy Trentham-Dietz, Oguzhan Alagoz, and Martin J. Yaffe, Cost-effectiveness of mammography from a publicly funded health care system perspective, *CMAJ open*, 2 2018, 6(1):E77–E86.
10. Rachid Rafia, Alan Brennan, Jason Madan, Karen Collins, Malcolm W.R. Reed, Gill Lawrence, Thompson Robinson, David Greenberg, and Lynda Wyld, Modeling the Cost-Effectiveness of Alternative Upper Age Limits for Breast Cancer Screening in England and Wales, *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*, 6 2016, 19(4):404–412.
11. P. A. van Luijt, E. A.M. Heijnsdijk, and H. J. de Koning, Cost-effectiveness of the Norwegian breast cancer screening program, *International Journal of Cancer*, 2 2017, 140(4):833–840.
12. Arantazu Arrospide, Montserrat Rue, Nicolien T. van Ravesteyn, Merce Comas, Myriam Soto-Gordoa, Garbiñe Sarriguarte, and Javier Mar, Economic evaluation of the breast cancer screening programme in the Basque Country: retrospective cost-effectiveness and budget impact analysis, *BMC cancer*, 6 2016, 16(1).
13. Kimbroe J. Carter, Frank Castro, and Roy N. Morcos, Insights Into Breast Cancer Screening: A Computer Simulation of Two Contemporary Screening Strategies, *AJR. American journal of roentgenology*, 3 2018, 210(3):564–571.
14. Irmgard Schiller-Fruehwirth, Beate Jahn, Patrick Einzinger, Günther Zauner, Christoph Urach, and Uwe Siebert, The Long-Term Effectiveness and Cost Effectiveness of Organized versus Opportunistic Screening for Breast Cancer in Austria, *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*, 9 2017, 20(8):1048–1057.
15. Rositsa G. Koleva-Kolarova, Alicja M. Daszczuck, Chris de Jonge, Mohd Kahlil Abu Hantash, Zhuozhao Z. Zhan, Erik Jan Postema, Talitha L. Feenstra, Ruud M. Pijnappel, Marcel J.W. Greuter, and Geertruida H. de Bock, A modelling study to evaluate the costs and effects of lowering the starting age of population breast cancer screening, *Maturitas*, 3 2018, 109:81–88.
16. Zorginstituut Nederland, Ziektelast in de praktijk, *De theorie en praktijk van het berekenen van ziektelast bij pakketbeoordelingen*, (2018).
17. Matthijs Versteegh, Saskia Knies, and Werner Brouwer, From Good to Better: New Dutch Guidelines for Economic Evaluations in Healthcare, *PharmacoEconomics*, 11 2016, 34(11):1071–1074.
18. National Cancer Institute, CISNET: Modeling Approach.

A SEARCH QUERY

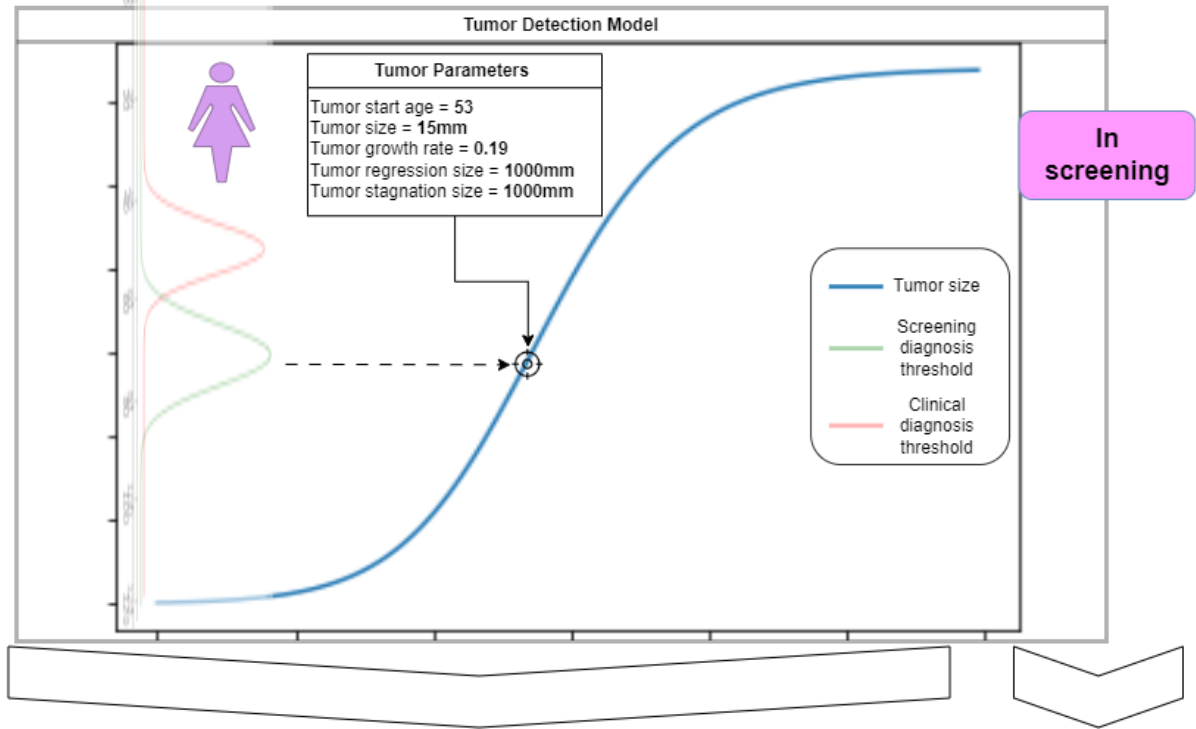
TITLE-ABS-KEY (("cost effectiveness" OR "cost benefit" OR "cost utility" OR "cost of illness" OR "budget impact" OR "Health economic" OR "value for money") AND ("Breast cancer screening" OR "Mammography screening" OR "Breast screening")) AND ("Discrete event simulation" OR "Simulation" OR "Markov model" OR "Population model" OR "Cohort model")) AND PUBYEAR i 2016 AND PUBYEAR j 2023

Appendix VII – Full Model Overview

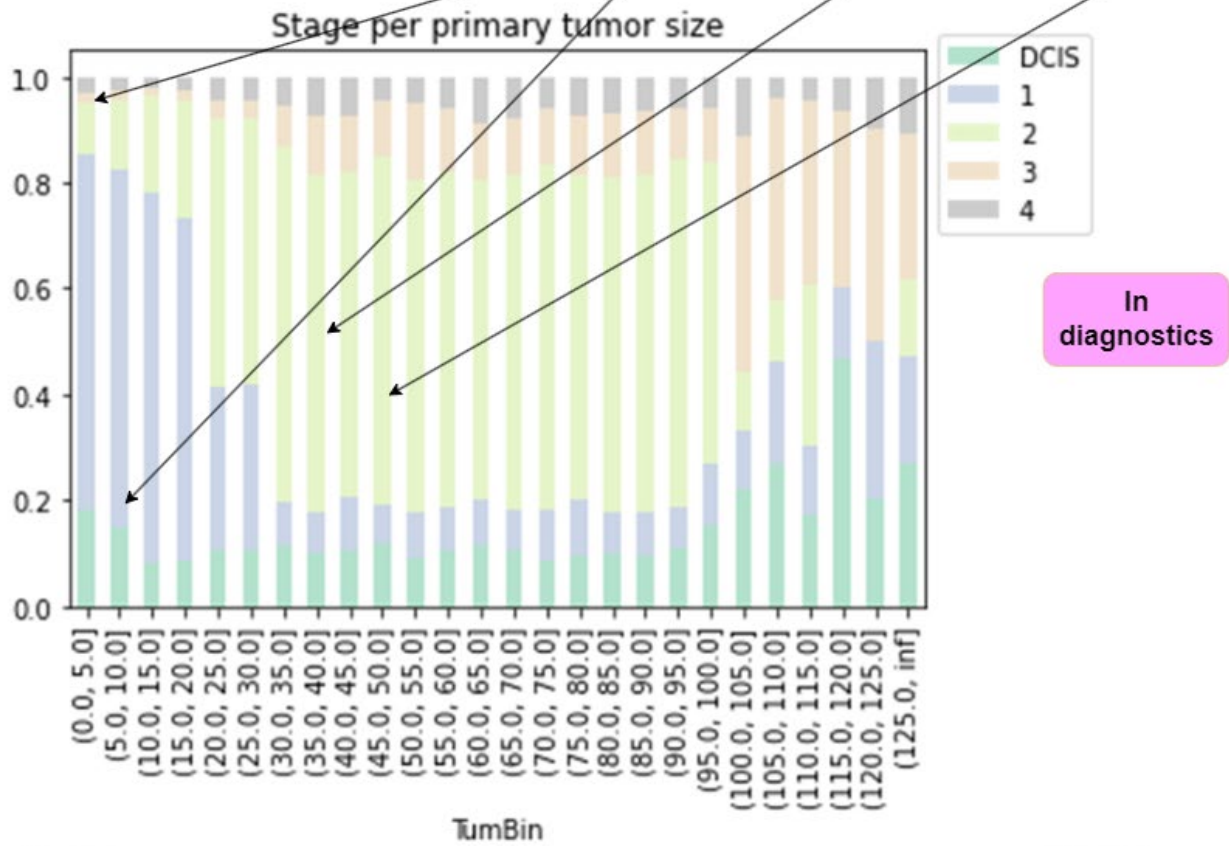
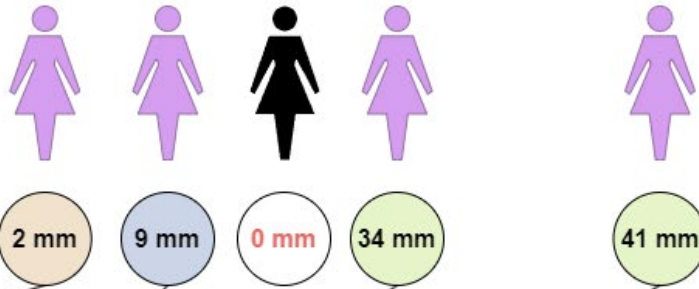




Some women will choose to receive screening, others will not. During screening, the tumor might get detected if the size is larger than the threshold



The size of the primary tumor will determine the stage at diagnosis. There is also a possibility for false positives



After diagnosis, patients will receive treatment based on the stage of their tumor. This treatment is the same, regardless of means of detection

Treatment

Appendix VIII – [Full code](#) explanation

In this appendix, the code will be introduced in commented code snippets with additional explanation. The code will be presented mostly in order as in the actual ran code. This document provides extra explanation on made choices that aren't in the code comments.

The R code is created in one file with different headers visible in the RStudio outline. Various parts of the code were finetuned in smaller code documents, available on [this project's GitHub repository](#).

Contents

Libraries.....	3
Strategy Selector.....	3
Functions.....	4
Cancer	4
Time	9
Next Steps	9
Utilities.....	10
Costs.....	11
Sim.....	14
Parameters.....	14
Patients	14
Data Loading	15
Time	15
Utilities.....	15
Costs.....	16
Tumor.....	16
Detection.....	17
Screen Start Params.....	18
Societal Cost Attributes.....	19
Attributes Recoding	19
Simulation	21
Init	21
Home.....	23
Screening.....	25
Diagnostics	27
Cured.....	29
Residuals	30

Hospital	31
Death.....	33
Base Model	34
Plot	35
Run	36
Reporting	37
Stage distribution Plot 2.....	37
Percentage Plot 1	39
Age of Diagnosis Plot 3.....	39
Survival Curve Plot 4	40
Final Age Distribution.....	42
Histogram Cancer Start age	42
Actual Stage Distribution	43
Tumor Growth Rates Histogram	44
Plot Stage Sizes Histogram.....	45
KPIs.....	46
Grid plots.....	47

Libraries

```
# Intro ---  
## Working Version for testing  
## V1.47  
  
# Libraries ----  
  
#workdir  
setwd('C:/Users/gielv/OneDrive/UT/IEM/001 AFSTUDEREN/RProjects/ThesisModel')  
# ^home  
#setwd('Z:/UT/IEM/001 AFSTUDEREN/RProjects/ThesisModel')  
# ^work  
  
rm(list=ls())  
gc()  
  
library(simmer)  
library(simmer.plot)  
library(simmer.bricks)  
  
library(fitdistrplus)  
library(dplyr)  
library(pracma)  
library(mvtnorm)  
library(gridExtra)  
set.seed(111)
```

In the 'Libraries' section, I keep track of the version, I set the working directory to that of my home pc or my work pc, and clear the console and memory of my pc. Then, the various libraries are loaded.

'Simmer', 'simmer.plot', and 'simmer.bricks' are the main libraries used for the simulation. They contain the building blocks to create trajectories and objects that move through them. These libraries are incredibly helpful for creating a discrete event simulation in R (Ucar et al., 2018).

'Fitdistrplus' is used to fit distributions to datasets. 'dplyr' is a data manipulation toolbox, mainly used for working with dataframes. 'Pracma' is used for more practical numerical math routines and brings more advanced functions for numerical analysis. 'mvtnorm' computes multivariate normal and t probabilities, quantiles, random deviates, and densities for statistical applications. 'ggplot2' is the main library for creating plots, and 'gridExtra' helps with the compact display of these plots.

An attempt was made to keep the amount of libraries to a minimum, to save on application weight and runtime.

'Set.seed' sets the seed of the simulation, so that randomly drawn variables may be reproduced over multiple simulation runs.

Strategy Selector

```
# STRATEGY SELECTOR ----  
  
# here you can select the screening strategy for this simulation run.  
Start_screen_age <- 50  
end_screen_age <- 75  
screen_interval <- 2  
  
manualScreenInput <- F
```

```

manualScreenAges <- c(50,51,52,53,54,55)
if (manualScreenInput==1){
  # enter ages at which patient should be screened here:
  ScreenAges <- manualScreenAges
}

```

The strategy selector is placed at the start of the simulation to create an intuitive overview for the user to determine which screening strategy is to be evaluated. The default is bi-annual screening between ages 50 and 75, which is the case in the Netherlands today. Starting and ending ages as well as the screening interval can be changed, and the simulation will then test a different strategy. If another strategy should be tested, the strategy can be inputted manually by changing the manualScreenAges vector to include the ages at which patients should be screened. The manualScreenInput parameter should then also be changed to True.

Functions

```
# Functions ----
```

The header functions is split up in different categories, highlight functions that are related to those attributes.

Cancer

```
## Cancer ----
```

All functions related to cancer are described under this header.

```

LifeExpAndCancerAge <- function(){
  # Generate one random value from the copula
  random_cancer <- 2
  random_life_exp <- 1
  if (runif(1)<(0.14)){ # Cancer incidence percentage
    while (random_life_exp<random_cancer){
      random_sample <- rmvnorm(1, mean = c(0, 0), sigma = correlation_matrix)

      # Transform copula data to distribution data
      random_life_exp <- qweibull(pnorm(random_sample[1,1]),
                                shape = LifeExpShape,
                                scale = LifeExpScale) *year
      random_cancer <- qnorm(pnorm(random_sample[1,2]),
                            mean = CancerExpMean,
                            sd = CancerExpSD) *year
    }
  } else {
    random_life_exp <- rweibull(1,LifeExpShape,LifeExpScale) *year
    random_cancer <- 1000*year
  }
  set <- c(random_life_exp,random_cancer)
  return(set)
}

```

LifeExpAndCancerAge is a function that determines the patient's healthy life expectancy and age at which she will get cancer. These values are generated through the copula that was derived in chapter 4 of the main thesis. Care is taken that the healthy life expectancy is always higher than the age of cancer onset, if the generated patient is destined to get cancer. Otherwise, the cancer incidence age is set to an unrealistically high number of a 1000 years. The healthy life expectancy is

taken from a Weibull distribution with variable parameters, the cancer incidence is taken from a normal distribution. Variables in these distributions are explained under parameters.

```
GompGrow <- function(age,onsetage,GR,IsCured){
  if(IsCured ==1){return(0)} else{
    t <- (age-onsetage) / month
    if (t<0){
      size<-0
    } else {
      #Determine size of tumour MANC
      Volume <- Vm/(1+((Vm/Vc)^0.25-1)*exp(-0.25*GR*t))^4
      #tumour volume at time t
      size <- 2*(Volume/(4/3*pi))^(1/3)
    }
    return(size)
  }
}
```

GompGrow is a simpler version of function GompGrow2, which contains only the function to determine the size of a tumor at a given time. This function resembles the growth of a non-recessive, non-stagnating tumor as described in chapter 4.

```
FindTimeAtSize <- function(size, onsetage, GR, max_size) {
  # Solve for the time delta using a binary search algorithm
  low <- 0
  high <- 1000*year # set an arbitrarily high upper bound for time delta
  while (high - low > 1e-6) {
    mid <- (low + high) / 2
    #size_mid <- GompGrow(onsetage + mid, onsetage, GR, IsCured=0)
    Volume_mid <- Vm/(1+((Vm/Vc)^0.25-1)*exp(-0.25*GR*mid))^4
    #tumour volume at time t
    size_mid <- 2*(Volume_mid/(4/3*pi))^(1/3)
    if (size_mid < size) {
      low <- mid
    } else {
      high <- mid
    }
  }
  # Return the estimated time when the given size was reached
  return(onsetage + high)
}
```

This function uses the GompGrow function to estimate the tumor size at different time points, and performs a binary search to find the exact time at which the given size was used. The algorithm starts with a lower bound of 0 and an arbitrarily high upper bound. It iteratively halves the search range until the found time converges to the time point at which the desired tumor size is found. This function is used in GompGrow2 below to determine the time at which the tumor started regressing, so that the new size of the tumor at this new time point can be calculated.

```
GompGrow2 <- function(age,onsetage,GR,IsCured,
  max_size,RegressionSize,StagnateSize){
  age = age/month
  onsetage=onsetage/month
  if (IsCured ==1){return(0)}
  Vm = (4/3)*pi*(max_size/2)**3
  size <- 0
```

```

delta = age-onsetage
if (delta<0){
  size <-0
  return (size)
}
#Determine size of tumour
Volume <- Vm/(1+((Vm/Vc)^0.25-1)*exp(-0.25*GR*delta))^4
#tumour volume at time t
size <- 2*(Volume/(4/3*pi))^(1/3)
if (size>StagnateSize){
  size <- StagnateSize
}
if (size>RegressionSize){
  size <- RegressionSize
  time_reached = FindTimeAtSize(RegressionSize,onsetage,GR,max_size)
  delta2 = delta-time_reached
  size2 = size - (size*pnorm(delta2,50,25))
  return(size2)
}
return(size)
}

```

GompGrow2 takes in a patient's age, cancer onset age, tumor growth rate, whether the patient is cured, the maximum size the tumor can reach, the size at which it will regress and the size at which it will stagnate to get an accurate size for the patient's tumor at that age. For this, it first checks to see if the patient is cured. If the patient is cured, the returned size will be zero. Otherwise, the maximum volume of the tumor will be determined and the time delta between onset and current age is calculated. If the delta is negative, the size of the tumor will be set to zero.

The volume of the tumor is determined according to the Gompertz equation explained in chapter four, and the diameter is then calculated. If the size is larger than the patient's predetermined stagnation size, that size will be returned instead. If the size is larger than the patient's regression size, the size will first be reset to this regression size, then the FindTimeAtSize function is used to search for the time at which this regression started, and the delta between regression onset and now is calculated. The actual size now is calculated by multiplying the original size with the probability from a normal distribution at that time and subtracting that from the original size.

This function is now always able to return a tumor size for a patient at a given point in time.

```

RegStagSizes <- function(){
  p <- runif(1)
  RS = 1000*year
  SS=1000*year
  if (p < 0.05){
    RS = rnorm(1,25,3)
  } else if (p<0.10){
    SS = rnorm(1,25,3)
  }
  return(c(RS,SS))
}

```

This function determines the sizes at which tumor growth will stagnate or start spontaneous regression. The probability that one of these sizes is given to a particular patient, is five percent for each. The sizes around which this happens is drawn from a normal distribution with a mean of

25mm and a standard deviation of 3. These values are taken from Cisnet model profiles (National Cancer Institute, 2022; Trentham-Dietz et al., 2021).

```
Staging <- function(size, p) {
  size_bin <- SizeTable[SizeTable$TumBinL <= size & SizeTable$TumBinR > size,]
  if (nrow(size_bin) == 0) {
    return(0)
  }
  cumulative_p <- cumsum(size_bin[,-c(1,2)])
  col_index <- which(cumulative_p >= p)[1]
  if (col_index == 1){
    return(0.5)
  } else if (col_index==2){
    return(1)
  } else if (col_index==3){
    return(2)
  } else if (col_index==4){
    return(3)
  } else if (col_index==5){
    return(4)
  }
  #return(colnames(SizeTable)[col_index + 2])
}
```

This function determines the stage of a tumor given the size and a probability. The probability is taken from a patient attribute, so that multiple diagnoses at the same size from the same patient will deliver the same result. The size of the tumor is compared to the stage-by-size matrix, which is explained in chapter 4. From this matrix, the stage of the tumor can be read and returned so that the patient is diagnosed.

```
CancerSurvival <- function(stage,lifeExp,currAge){
  p = runif(1)
  if (stage == 0.5){
    NLE = lifeExp #NLE= new life exp
  } else if (stage == 4) {
    NLE = tail(which(SurvivalDf$Stage_IV >= p * 100) - 1,1)*year + currAge
  } else if (stage == 3) {
    NLE = tail(which(SurvivalDf$Stage_III >= p * 100) - 1,1)*year+ currAge
  } else if (stage == 2) {
    NLE = tail(which(SurvivalDf$Stage_II >= p * 100) - 1,1)*year+ currAge
  } else if (stage == 1) {
    NLE = tail(which(SurvivalDf$Stage_I >= p * 100) - 1,1)*year+ currAge
  } else{NLE=lifeExp}

  if (NLE > lifeExp){NLE <- lifeExp}

  return(NLE)
}
```

The CancerSurvival function takes in the stage of a patient's tumor, healthy life expectancy and current age. It then returns the New Life Expectancy based on the survival curve for different stages of breast cancer. This approach was taken instead of modelling complete treatments strategies, and nets the same result without taking up too much simulation resources. Once a tumor is diagnosed, the path for treatment and survival is the same, regardless of diagnosis through screening or clinical discovery.

```

ScreenResult <- function(TumorSize){
  min_dia <- rweibull(1,MeanMinDetSize,SDMinDetSize) # from miscan paper
  # define BIRADS based on how much above/below this size
  if (TumorSize == 0){
    birads = ifelse(runif(1)<FP_Perc,3,1) #False Positive
  } else if (TumorSize<0.95*min_dia){
    birads=1
  } else if (TumorSize>0.95*min_dia & TumorSize<1.25*min_dia){
    birads = 3
  } else if (TumorSize>1.25*min_dia & TumorSize<2*min_dia){
    birads = 4
  } else if (TumorSize>2*min_dia){
    birads = 5
  }
  return(birads)
}

```

The function screenresult determines a BIRADS value given the size of the tumor. This is used for determining the result of the screening procedure. There is a probability for false positives built in, which can be changed under the parameters settings. If the tumor is of a given size, then the screening will detect it based on a minimal diameter determined through a Weibull distribution. The BIRADS scores of 3,4 or 5 will send a patient to diagnostics, where the actual stage will be determined.

```

ClinicalCheckAtHome <- function(TumorSize,ClinSizeCheck,PatentAge){
  if (PatentAge %in% ScreenAges){return(0)}
  if ((PatentAge/month) %% 2 == 0){ #round(runif(1,1,2))
    if (TumorSize>ClinSizeCheck){
      return(1)
    } else {return(0)}
  }else {return(0)}
}

```

A clinical check at home will take place once every one or two months. Women are currently advised to check their breasts at least monthly themselves, but are not expected to fully adhere to this. Therefore, a slightly lower occurrence of clinical checking was chosen. If patients are also invited for screening at this time, then no clinical check will take place. To determine if the clinical check nets a positive result, the size of the tumor is compared to the minimal size needed for clinical diagnosis, which is different for each patient. If a tumor has been detected clinically, the patient will be referred to diagnostics.

```

GrowRater <- function(startAge){
  GRmultiplier <- 0
  GR <- rgamma(1,shape = Grow_Gamma_shape, rate = Grow_Gamma_rate)
  # print(GR)
  startAge <- startAge/year
  if (startAge <50){
    GRmultiplier = -(62-startAge)/80
  }
  if (startAge >75){
    GRmultiplier <- -(startAge-62)/80
  }
  GR <- GR + GRmultiplier
  if (GR<0){
    GR <- GR - GRmultiplier
  }
}

```



```

return(GR)
}

```

The GrowRater function will slightly adjust tumor growth rates based on patient's ages. Tumors generally grow slower or faster at different ages, and these values were found through iterative finetuning in order to recreate the graph of tumor stages found at different ages as seen on the IKNL website (Integraal Kankercentrum Nederland, 2022).

Time

```
## Time ----
```

Time functions are used to determine how long a patient will remain in a certain trajectory before going to the next step.

```

WaitAtHome <- function(EntryAge){
  # find next month
  div <- EntryAge %% month
  return(month - div)
}

```

WaitAtHome takes in the patient's age at entry and creates a timeout until the next month when the next steps are determined.

```

ScreenTime <- function(){
  return(max(0.1,rnorm(1,MeanScreenTime,SDScreenTime)))
}
DiagTime <- function(){
  return(max(0.1,rnorm(1,MeanDiagTime,SDDiagTime)))
}

```

ScreenTime and DiagTime return a duration for how long patients should remain in that trajectory. They are drawn from normal distributions with average durations and standard deviations. A minimum is also built in to prevent returning negative durations.

Next Steps

The function AfterHome determines the next step for the patient: what trajectory should they follow after exiting this trajectory.

```

## Next Steps ----
AfterHome <- function(Invite,Screened,Birads,ClinicalCheck,
  Referral,PatientAge,HealthyLifeExpectancy,TumorSize){
  # Next Steps update
  #All five steps possible:
  #a. Go to screening after invite-> 2
  #b. Go to diagnostics due to tumor size or referral after screening ->3
  #c. Go to hospital due to referral -> 4
  #d. Die due to age or illness-> 5
  #e. Continue in Home trajectory, nothing's needed -> 1
  #PatientAge <- PatientAge / year
  s <- 0
  if (Invite>0){
    if (PatientAge %in% ScreenAges){ # check if go for screen
      if (Invite==1){
        upt = uptakefirstscreen
      }else if (Invite > 1){
        upt = uptakeotherscreen
      } else {upt = uptakenoscreen}
      if (upt > runif(1)){
        s <- 2
      }
    }
  }
}

```

```

    }
  }
}

if (ClinicalCheck ==1 | Birads == 3 | Birads == 4 | Birads ==5){
  # check if diagnostics
  s <- 3
}

if (Referral ==1){
  s <-4
}

if (PatientAge > HealthyLifeExpectancy){
  s <- 5 # die
}
if (s==0){s<-1}# stay at home
return(s)
}

```

The function determines if a patient should go to screening, diagnostics, hospital, death or stay at home as the next step. In order to determine this, it takes in the patient parameters Invite, Screened, Birads, ClinicalCheck, Referral, PatientAge, HealthyLifeExpectancy, and TumorSize.

First, the patient is checked to see if she should go to a screening test. If a patient should go to screening, she should have an invite and the patient's age should be in the vector containing the ages at which patients are screened. Then, the probability at which the patient actually adheres to the invitation is determined. Not all patients go for screening, with around 75% going to their first screening, and 90% going to screenings after their first. If no screening has been done but more than one invite has been received, then the probability of going to the screening drops to 25%. These probabilities are checked against a random value from a uniform distribution, and then the patient might get a NextStep value of 2, meaning she will visit the screening trajectory next.

If the patient has found a tumor clinically or through screening, based on the parameters ClinicalCheck and Birads, the patient will go to diagnostics next.

If after diagnostics the patient is referred to the hospital, the patient will go to the hospital as their next step.

If a patient is older than their healthy life expectancy, they should die in the next step.

If none of the above conditions are true, then the patient will remain in the home trajectory for another month.

Utilities

In the Utilities section, two functions for determining a patient's current utility are presented.

```

## Utilities ----
BaseAgeUtil <- function(Age){
  Age <- Age/year
  if (Age < 31){
    return(1)
  }
  Age <- ceiling((Age-30)/5) # steps of 5
  return(utility_ages[Age,2])
}

```

The base utility of a patient is dependent on age. At higher ages, a higher base utility is used. This function looks up a patient's base utility in the utility_ages table and returns this.

```
CancerBasedUtil <- function(currentUtil, stage, age, startAge){
  if (stage==0){
    s=1
  } else if (stage==0.5){
    s=2
  } else if (stage==1){
    s=3
  } else if (stage==2){
    s=4
  } else if (stage ==3){
    s=5
  } else if (stage==4){
    s=6
  }
  delta <- (age-startAge) / year
  utilDec <- utility_decrements[s,2]
  discUtilDec <- utilDec/((1+UtilDiscount)^floor(delta))
  newUtil = currentUtil * (1-discUtilDec)
  return(newUtil)
}
```

If a patient has cancer, the utility or quality of life of that patient is lower. This function takes in the current utility, often the base utility, the stage of the cancer, the patient's age and the age at which the tumor started. First, it determines the index of which utility decrement to look up based on the cancer's stage. Then the time delta between tumor onset and patient age is determined. After that, the new utility decrement is taken from the utility_decrements table.

Based on the time difference between tumor onset and current age of the patient, the actual value of the utility decrement is found through discounting. A 0.10 discount in utility now should only be a decrement of 0.086 in ten years, as utility is discounted at 1.5%. The new utility is found by multiplying the base utility by 1 minus the decrement. This new utility is then returned as patient parameter.

Costs

```
## Costs----
```

Under Costs, various functions for determining costs of pre-medical procedures, societal occurrences, end-of-life costs and cost discounting are given.

```
ScreenDiagInvitesCost <- function(screened,diaged,invite){
  if (is.na(screened)){screened=0}
  if (is.na(diaged)){diaged=0}
  if (is.na(invite)){invite=0}
  cost<- 0
  cost<- cost + screened* AvgScreenCost
  cost<- cost + diaged*AvgDiagCost
  cost<- cost + invite*AvgInvCost
  return(cost)
}
```

This function determines a patient's costs associated with the number of screenings she's received, how often she has been through diagnosis and how often she has received an invite. It is calculated at the end of a patient's life based on the attributes screened, diagnosed, and invites.

```

TreatmentCost <- function(stage,age,lifeExp){
  # cost is determined by:
  # initial cost for first 12 months post treatment
  # ^always counted
  # costs per month after the first 12, up to the last 6 months
  # counted if this interval exists
  # Costs for the last six months of care
  # ^always counted
  delta <- (lifeExp - age) / month
  #delta is how long we have left
  c<-0 # cost variable
  #determine right row from table:
  if (stage==0.5){
    row <- 2
  } else if (stage ==0){
    row <- 1
  } else {
    row<- stage+2
  }
  InitC <- rnorm(1,TreatCostTable$Initial12Cost[row],
    TreatCostTable$InitialSD[row]) /12 #round to monthly
  TerminalC <- rnorm(1,TreatCostTable$Terminal6Costs[row],
    TreatCostTable$TerminalSD[row]) /6 # to monthly
  MidC <- rnorm(1,TreatCostTable$ContinuousMonthlyCost[row],
    TreatCostTable$ContinuousSD[row]) /2

  CInit <- min(delta,12) * InitC #no discount, this is year 0
  CTerminal <- min(delta,6) * TerminalC /
    (1 + CostDiscount)^((round((lifeExp - age)/year)) - 1)
  # discount with 4% per year post diagnosis
  CMid <- annuity_cost_monthly(MidC,delta,CostDiscount)

  c <- CInit + CTerminal + CMid
  return(c)
}

```

Treatment costs is the function that determines the costs of treating a patient's cancer. These costs are based on cancer stage, age and the life expectancy of the patient. The costs are determined in three steps: first the initial costs for the first year of treatment is calculated, then the costs for the final six months of a patient are retrieved and discounted, and then the costs for the months in between initial and final treatment are determined.

The initial costs are taken from the TreatCostTable based on the patient's stage. The costs are then divided by twelve to create monthly costs. Based on the time the patient has left to live, found by creating the delta variable by subtracting the patient's current age from her life expectancy, the costs for this first year are determined. These costs are not discounted, as time of diagnosis is t_0 in the simulation.

The final costs are taken by multiplying 6 or less months by the discounted costs of final treatment. If a patient has less than six months to live at time of diagnosis, not all six months are counted. The costs are discounted at 4% yearly to the time the patient will die in the future. If a patient still has 25 years left to live and the final costs of treatment would be 10000 euros today, the discounted value would be 3750 euros.

For the costs of treatment for all the months between the initial 12 and final 6 months of a diagnosed patients life, the function annuity_cost_monthly was created to sum the discounted

monthly treatment costs. At the end of the function, the three costs variables are summed up and returned to the attribute.

```

VisitationCosts<- function(screened,diaged,hosps){
  costs<- 0
  hosps <- hosps/day #time to days

  #travel costs
  costs <- costs +
    (ifelse(runif(1)<perc_car,1,0) * dist_to_screen * cost_per_km) * screened
  # for screening visits
  costs <- costs +
    (ifelse(runif(1)<perc_car,1,0) * dist_to_hosp * cost_per_km) * diaged
  # for diagnosis in hospital
  costs <- costs +
    (ifelse(runif(1)<perc_car,1,0) * dist_to_hosp * cost_per_km) * hosps

  # productivity loss costs:
  # screening & diag takes average time, hospital takes full working day
  costs <- costs + (screened*MeanScreenTime*prod_cost_hour) +
    (diaged*MeanDiagTime*prod_cost_hour) +
    (hosps*8*hour*prod_cost_hour)

  return(costs)
}

```

The function VisitationCosts determines the societal costs now mandated by the Dutch guidelines on health economic analyses. The costs consists of travel costs to screening, hospital, and diagnosis and are based on standard values from the guidelines (Zorginstituut Nederland, 2016). The costs of productivity loss are also calculated based on the total time a patient has spent in diagnostics, screening and hospital. This time in hours is multiplied with the guideline-given standard value for productivity loss costs per hour. The total societal costs is then returned.

```

PaidCosts <- function(lifeExp,StartAge){
  costs<- 0
  # PAID tool gives insight in expected rest of life costs for if someone
  # lives longer after cancer treatment.
  if (lifeExp>StartAge){
    costlist <- WomenCostList[(round(StartAge)/year):(round(lifeExp)/year)]
    yearlist <- 0:(length(costlist)-1)
    disclist<- costlist / (1+0.04)^yearlist
    costs = sum(disclist)
  }
  return(costs)
}

```

The Guidelines also give instructions to include the costs for future illnesses in the calculation of societal costs. These costs are calculated through PAID, which provides a csv file of costs related to other diseases than breast cancer at various ages through their online tool (PAID 3.0, n.d.). These costs are calculated for the extra years of life a patient has to live after diagnosis and treatment, and are discounted at 4% to their current value.

```

annuity_cost_monthly <- function(payment, months, rate) {
  monthly_rate <- (1 + rate)^(1/12) - 1
  total_cost <- 0
  if (months<13){
    return(0)
  }
}

```

```

months_left = min(0,months-6) #don't count the final 6
for (i in 13:months_left) {
  discounted_payment <- payment / (1 + monthly_rate)^(i - 1)
  total_cost <- total_cost + discounted_payment
}
return(total_cost)
}

```

The `annuity_cost_monthly` function is used for calculating the mid part of the medical costs. It iteratively sums the monthly discounted values of the costs of treatment post first 12 months and pre final six months, and returns the costs for this section.

Sim

```
## Sim ----
```

Under Sim, two functions are placed that are taken from the UT course Advanced Simulation for Health Economic Analysis. These function allow for the tracking of one or multiple attributes during the whole simulation, so that patients might be traced and the simulation can be checked for errors.

```

getSingleAttribute <- function(attribute,
                              output,
                              all=F) sapply(unique(output[, "name"]),
                                             function(entity) {
list(
  if(all) {
    output[output[, "name"]==entity & output[, "key"]==attribute, "value"]
  } else {
    tail(output[output[, "name"]==entity &
              output[, "key"]==attribute, "value"], n=1)
  }
)
})

```

GetSingleAttribute takes one single attribute as an argument and will find all occurrences of it in the final results data frame if the monitoring of the system is set to true. Then, all these occurrences can be checked in a new data frame.

```

getMultipleAttributes <- function(attributes, output)
as.data.frame(sapply(attributes, function(attribute)
as.numeric(getSingleAttribute(attribute, output))))

```

GetMultipleAttributes applies the function `getSingleAttribute` multiple times and returns a dataframe containing more than one attribute over time.

Parameters

```
# Parameters ----
```

In the Parameters section, data is loaded and values are set for various variables. Most data is taken from literature and some is found through finetuning.

Patients

```

## Patients----
n.patients=1000;
mon.patients <- ifelse(n.patients<100,2,0);
LifeExpShape <- 7.937

```

```
LifeExpScale <- 86.788
```

Under Patients, the number of patients to be simulated is set. If the number is low enough, the advanced monitoring will kick in which allows for individual patient tracing through the model using the `getSingle-` and `getMultipleAttributes` functions.

The `LifeExpShape` and `LifeExpScale` parameters have been found through analysis of CBS data on dutch life expectancy, and are used in a Weibull distribution in the function `LifeExpAndCancerAge`.

Data Loading

```
## Data loading ----  
  
# Transform copula data to distribution data  
correlation_matrix <- matrix(c(1, 0.5, 0.5, 1), ncol = 2)  
#for age & cancer incidence
```

```
SizeTable <- read.csv('StagingFromSizes.csv')
```

Under data loading, first the correlation matrix for `LifeExpAndCancerAge` is created. The correlation between the two is set to 50%, which netted the best results.

The `SizeTable` data is loaded, which contains the data used to classify a tumor of a certain size as a certain stage. This data is discussed in chapter 4.

Time

```
## Time-----  
hour = 1;  
minute = hour/60;  
second=minute/60;  
day=hour*24;  
year = day*365;  
month = year/12  
  
MeanScreenTime <- 4*hour  
SDScreenTime <- 30*minute  
MeanDiagTime <- 6*hour  
SDDiagTime <- 30*minute
```

Various timing related variables are initiated under 'time'. First of all, the base variables are set so to make it easier to work with simulation time. A base of 1 hour is chosen, and the variables minute, second, day, year, and month are derived from that. These variables are used throughout the simulation and in various functions to make it easier to work with time.

The variables for the mean and standard deviation for screening and diagnosis are also set, so that every patient has slightly different times in those stations. This is also an example of how the time variables of hour and minute are used in defining other time related variables.

Utilities

```
## Utilities ----  
# these are base utilities per age-range,  
# taken from the MANC model  
utility_ages<-data.frame(c(30,35,40,45,50,  
55,60,65,70,75,  
80,85,90,95,100),  
c(0.9383,0.9145,0.9069,0.8824,0.8639,  
0.8344,0.8222,0.8072,0.8041,0.779,  
0.7533,0.6985,0.6497,0.6497,0.6497)))
```

```
utility_decrements <- data.frame(c('Healthy','DCIS','StageI',
                                'StageII','StageIII','StageIV'),
                                c(0,0.1,0.15,0.20,0.25,0.40))
UtilDiscount <- 0.015
```

In the utilities section, various parameters and data frames related to the quality of life of patients are initiated. First, the data frame `utility_ages` is created. This contains the base utility for patients of a certain age. For example, a 50-year-old woman has a base utility of 0.8639. This data is based on the Manchester model (Wright et al., 2022)

Then, the utility decrements for various stages of cancer are given. Healthy will net a decrement of 0, and stageIV a decrement of 40%. This data is based on Arrosptide et al. (Arrosptide et al., 2016)

Finally, the mandated 1.5% discount rate for utilities is initiated in the variable `UtilDiscount` (Zorginstituut Nederland, 2016).

Costs

```
## Costs ----
TreatCostTable <- read.csv('TreatmentCostsVanLuijt2023Eur.csv')
CostDiscount <- 0.04
AvgScreenCost <- 100
AvgInvCost <- 5
AvgDiagCost <- 212
```

In the costs section, the treatment costs table is loaded. This contains the data for costs of treatment per stage. It is split up in initial 12 months, mid-months, and final 6 months. The explanation of usage of this data is found under the function `TreatmentCosts`. The data can be seen in table x

Stage	Initial12Cost	InitialSD	Continuous MonthlyCost	ContinuousSD	Terminal6Costs	TerminalSD
0	0	0	0	0	0	0
0.5	9583.123	406.9691	140.3392	13.83695	23672.61	3798.378
1	14497.39	312.0097	222.954	11.39514	18138.93	1899.189
2	29104.01	651.1506	423.9329	18.0423	22905.14	1220.907
3	35756.04	2170.502	589.8203	97.80824	18846.1	2848.784
4	33628.54	3527.066	1236.087	157.0901	24758.81	2848.784

The data comes from the study of VanLuijt from 2016, and the amounts are corrected to represent 2023 euros. Reliable recent Dutch data was hard to find. A 2013 paper of de Bock, Siesling et al does contain information on costs for breast cancer treatment based on size, but the data was from 2000 and therefore deemed unrealistic, even with corrections for inflation. To make sure the data used in this paper is reliable, the factors for treatment for different stages were checked. In the 2013 paper, the costs for treatment of advanced cancer were up to five times higher than that of simple cancers. This is comparable to the data used in this paper.

Furthermore, in this section the discount rate for costs is set to 4%, which is used everywhere costs are calculated.

The average costs for screening, invites and diagnosis are also given. The costs for screening, invites and diagnosis are based on the annual report of `BevolkingsonderzoekNederland` and are checked with various studies from the systematic review (Arrosptide et al., 2016; PriceWaterhouseCoopers Accountants N.V., 2022; Rafia et al., 2016; van Luijt et al., 2017).

Tumor


```
## Tumor ----
```

In the tumor section, all variables related to the tumor are set, including age of onset, growth parameters and survival probabilities.

```
#Set tumour growth rate parameters
CancerExpMean = 61.871
CancerExpSD = 14.14

Grow_Gamma_shape <- 1.567742
Grow_Gamma_rate <- 1.933883

max_size <- 128 #mm diameter
start_size <- 0.25 #starting size of tumours, diameter in mm
Vc = (4/3)*pi*(start_size/2)^3 #Volume at start
Vm = (4/3)*pi*(max_size/2)^3 #Max volume

#DF to show survival probability after X years at X stage
SurvivalDf <- data.frame(Years_After_Diagnosis = c(0:10),
  Stage_I = c(100, 100, 100, 99, 99,
    98, 97, 97, 96, 95, 95),
  Stage_II = c(100, 99, 97, 95, 93,
    91, 89, 88, 86, 85, 83),
  Stage_III = c(100, 96, 89, 83, 78,
    73, 69, 65, 62, 60, 58),
  Stage_IV = c(100, 70, 53, 38, 29,
    22, 17, 13, 11, 9, 7))
```

The variables CancerExpMean and CancerExpSD are the parameters used in the normal distribution to determine the age of cancer onset using the copula in the function LifeExpAndCancerAge. These values are found through fitting a distribution on the synthetic data of the IKNL (Integraal Kankercentrum Nederland (IKNL), 2022).

The grow parameters are used in the function GrowRater to determine the growth rate of the tumors. How these growth rates are determined, is explained in chapter 4.

Furthermore, the max_size, start size, start volume and max volume of the tumors are defined. These parameters are based on cisnet model profiles, and tuning of this is explained in chapter 4. Not that to determine Vc and Vm, the formula to create the volume of a sphere from a diameter is used (National Cancer Institute, 2022; Trentham-Dietz et al., 2021).

The SurvivalDf is also initiated. This contains the probabilities of survival per stage, and is used in the function CancerSurvival to determine how long someone has left to live after diagnosis.

Detection

```
## Detection ----
```

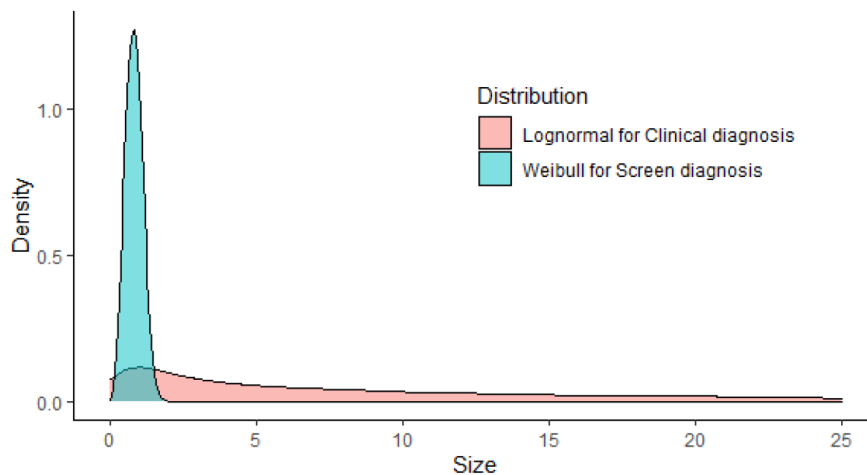
Under detection, the various parameters for detection of tumors are set.

```
MeanMinDetSize <- 0.91 # for screening
SDMinDetSize <- 0.34 # for screening

MeanLogClin <- 3.22
SdLogClin <- 2.25
AlwaysFoundSize <- 100 # buld in a max! so min(100,clindiagsize)
```

```
FP_Perc <- 0.005 #FalsePositive
```

First the parameter for screen detection are set. These parameters are used in the Weibull distribution in the function ScreenResult, so the minimum diameter varies every time someone visits screening. The parameters for clinical diagnosis are higher on average, and drawn from a lognormal distribution instead of Weibull. The minimum size for clinical diagnosis is set once at the beginning of a patients life instead of having a new one drawn every check. There is a maximum size of detection of 100mm built in, as women are expected to always notice a tumor themselves if it has reached that size. The overlay of the two distributions is shown in the figure below:



Here, the false positive percentage is also set. The false positive percentage is set at 0.5%, which is lower than the currently found false positives. One thing to consider is that this percentage is taken every screen, so a lot of patients will get checked against this percentage 10 times, meaning that the cumulative probability of getting a false positive once in your lifetime is a lot higher than 0.5%, at 4.9% over 10 screenings. Furthermore, part of the false positives can also come from tumors that do show up during screening, but have spontaneously regressed by the time of diagnosis.

Screen Start Params

```
## Screen Start Params ----
```

In Screen Start Params, various parameters for the screening strategy are set.

```
Start_screen_age <- Start_screen_age * year
end_screen_age <- end_screen_age * year
screen_interval <- screen_interval * year
ScreenAges = seq(Start_screen_age, end_screen_age, by=screen_interval)
ScreenAges <- append(ScreenAges, 1000 * year)
if (manualScreenInput){
  # enter ages at which patient should be screened here:
  ScreenAges <- manualScreenAges * year
}
# ScreenAges = c(50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74) * year
```

```
#Screening uptake MANC model to Dutch
uptakefirstscreen<- 0.78#0.605
uptakeotherscreen<-0.9#0.852
uptakenoscreen<-0.25#0.191
```

First, the parameters inputted by the user under STRATEGY SELECTOR are transformed to the right time format to be used in the simulation.

Then, the parameters for the uptake of screening are set. These are used in the AfterHome function to determine if the patients actually visit screening after receiving an invite. The idea is based on the Manchester model, and the numbers are taken from the monitor bevolkingsonderzoek 2020 (van Haperen, 2018; Wright et al., 2022).

Societal Cost Attributes

```
## Societal Cost Attributes ----
```

Here, the various standard values for calculating societal costs are set. These values are based on the **guidelines**.

```
dist_to_hosp <- 7
dist_to_screen <- 1.1
avg_park_cost <- 3
perc_car <- 0.8 # most go by car, definitely for hospital visits

cost_per_km <- 0.19 # for Public transport & car

prod_cost_hour <- 31.6 # for women, Dutch guidelines

AdditionalLifeYearCosts <-
read.csv('PAID_Cost_additional_year_post_BC_Costs_Living_Year_Longer_2023-03-
07.csv')
WomenCostList <- AdditionalLifeYearCosts$Unrelated_Women
```

The average distance to hospital is used for diagnosis and treatment, the average distance to screening is used for screening and based on distance to GP. Average parking costs, percentage by car, costs per km and productivity costs per hour are also taken from the guidelines (Zorginstituut Nederland, 2016). The AdditionalLifeYearCosts reads a csv file from the PAID tool discussed earlier. WomenCostList extracts the one useful column from the file. This contains the costs related to other diseases after treatment for breast cancer.

Attributes Recoding

```
# Attributes Recording ----
```

Attributes recording contains one function and one list, which is used to record the final status of all attributes of a patient at the end of her life. The function also returns a 0, to indicate the patient is now dead.

```
allAttributes = c("Index", "Name", "Alive", "PatientAge", "HealthyLifeExpectancy",
  "TumorStartAge", "LastAge", "TotalCosts",
  "MedicalCosts", "SocietalCosts", "PAIDCosts", "TotalUtility",
  "CurrentUtility", "TumorSize", "WorstCancer",
  "WorstStage", "CancerStageT", "IsCured", "StagingProb",
  "TumorGrowthRate", "max_size",
  "RegressionSize", "StagnateSize", "ClinSizeCheck",
  "Invite", "Screened", "BIRADS",
```

```
"Referral", "NextStep", "EntryAge", "TimeAtHome",
"ClinicalCheck", "TimeAtDiagnostics",
"TotalTimeAtDiagnostics", "DiagnosedThrough",
"DiagnosticVisits", "TimeAtHospital", "TotalTimeAtHospital",
"OriginalLifeExp", "HospitalVisits", "TimeAtScreening",
"TotalTimeAtScreening", "FirstStage")
```

AllAttributes is a list containing all the attributes used in the simulation.

```
EndDatDf <- data.frame(matrix(rep('test',length(allAttributes)),
                             ncol=length(allAttributes)))
colnames(EndDatDf) <- allAttributes
```

EndDatDf is a data frame containing a column for each attribute, and by the end of the simulation a row for each patient. The column names are set to the names of the attributes.

```
addAtts <- function(EndDatDf){
  addList <- c(nrow(EndDatDf),
              get_name(basic_sim),
              get_attribute(basic_sim, 'Alive'),
              get_attribute(basic_sim, 'PatientAge'),
              get_attribute(basic_sim, 'HealthyLifeExpectancy'),
              get_attribute(basic_sim, 'TumorStartAge'),
              get_attribute(basic_sim, 'LastAge'),
              get_attribute(basic_sim, 'TotalCosts'),
              get_attribute(basic_sim, 'MedicalCosts'),
              get_attribute(basic_sim, 'SocietalCosts'),
              get_attribute(basic_sim, 'PAIDCosts'),
              get_attribute(basic_sim, 'TotalUtility'),
              get_attribute(basic_sim, 'CurrentUtility'),
              get_attribute(basic_sim, 'TumorSize'),
              get_attribute(basic_sim, 'WorstCancer'),
              get_attribute(basic_sim, 'WorstStage'),
              get_attribute(basic_sim, 'CancerStageT'),
              get_attribute(basic_sim, 'IsCured'),
              get_attribute(basic_sim, 'StagingProb'),
              get_attribute(basic_sim, 'TumorGrowthRate'),
              get_attribute(basic_sim, 'max_size'),
              get_attribute(basic_sim, 'RegressionSize'),
              get_attribute(basic_sim, 'StagnateSize'),
              get_attribute(basic_sim, 'ClinSizeCheck'),
              get_attribute(basic_sim, 'Invite'),
              get_attribute(basic_sim, 'Screened'),
              get_attribute(basic_sim, 'BIRADS'),
              get_attribute(basic_sim, 'Referral'),
              get_attribute(basic_sim, 'NextStep'),
              get_attribute(basic_sim, 'EntryAge'),
              get_attribute(basic_sim, 'TimeAtHome'),
              get_attribute(basic_sim, 'ClinicalCheck'),
              get_attribute(basic_sim, 'TimeAtDiagnostics'),
              get_attribute(basic_sim, 'TotalTimeAtDiagnostics'),
              get_attribute(basic_sim, 'DiagnosedThrough'),
              get_attribute(basic_sim, 'DiagnosticVisits'),
              get_attribute(basic_sim, 'TimeAtHospital'),
              get_attribute(basic_sim, 'TotalTimeAtHospital'),
              get_attribute(basic_sim, 'OriginalLifeExp'),
              get_attribute(basic_sim, 'HospitalVisits'),
              get_attribute(basic_sim, 'TimeAtScreening'),
              get_attribute(basic_sim, 'TotalTimeAtScreening'),
              get_attribute(basic_sim, 'FirstStage'))
```

```

)
EndDatDf<- rbind(EndDatDf,addList)
return(0)
}

```

The function addAtts takes in the EndDatDf, so that it is always working with the most recent version. Then, it collects all the patients attributes and adds this to a new vector. Using rbind and a double <-, the new list is connected to the global attribute EndDatDf. A zero is returned to indicate the patient is now dead.

Simulation

```

# Simulation ----
## Trajectories ----

```

In the Simulation section, the various trajectories are discussed, the trajectories are visualized and the run settings are given.

Init

```

### Init ----
# trajectory for setting all patient-specific starting attributes

```

Init is the first trajectory all patients pass through. They pass through it only once, and in this trajectory some attributes are initialized at 0, other are set to their patient specific values.

```

Initialization <- trajectory()%>%
  set_attribute(key='Alive',value=1) %>%
  set_attribute(key='PatientAge',value=0) %>%
  set_attribute(key=c('HealthyLifeExpectancy','TumorStartAge'),
               value= function() LifeExpAndCancerAge())%>%
  set_attribute(key='LastAge',value=0)%>%
  set_attribute(key='TotalCosts',value=0)%>%
  set_attribute(key='MedicalCosts',value=0) %>%
  set_attribute(key='SocietalCosts',value=0) %>%
  set_attribute(key='TotalUtility',value=0)%>%
  set_attribute(key='CurrentUtility',value=1) %>%
  set_attribute(key='TumorSize',value=0) %>%
  set_attribute(key='WorstCancer',value=0) %>%
  set_attribute(key='WorstStage',value=0) %>%
  set_attribute(key='CancerStageT',value=0) %>%
  set_attribute(key='IsCured',value=0) %>%
  set_attribute(key='StagingProb',value=function() runif(1)) %>%
  set_attribute(key = 'TumorGrowthRate',
               value = function() GrowRate(get_attribute(basic_sim,
                                                         'TumorStartAge')))%>%
  set_attribute(key = 'max_size',
               value=function() rnorm(1,max_size,10)) %>%
  set_attribute(key = c('RegressionSize','StagnateSize'),
               value= function() RegStagSizes())%>%
  set_attribute(key='ClinSizeCheck',
               value=function() min(AlwaysFoundSize,
                                   rlnorm(1,MeanLogClin,SdLogClin))) %>%
  # set size needed for clinical find for this patient
  set_attribute(key = 'Invite',value=0)%>%
  set_attribute(key = 'Screened',value=0)%>%
  set_attribute(key = 'BIRADS',value=10) %>%#Invalid BIRADSvaluebefore checkup
  set_attribute(key = 'Referral',value=0) %>%# No referral

```

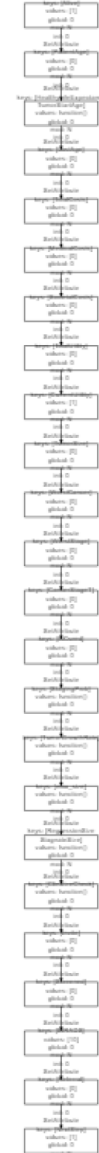
```

set_attribute(key='NextStep',value=1) # go home for first iteration
#timeout(30*year) %>%
#log_(paste0('Going Home'))

```

The following happens in the Init trajectory:

1. Attribute `Alive` is initialized at `1`
2. Attribute `PatientAge` is initialized at `0`
3. Random values are drawn from function `LifeExpAndCancerAge` to set the attributes `HealthyLifeExpectancy` and `TumorStartAge`.
4. Attribute `LastAge` is initialized at `0`
5. Attribute `TotalCosts` is initialized at `0`
6. Attribute `MedicalCosts` is initialized at `0`
7. Attribute `SocietalCosts` is initialized at `0`
8. Attribute `TotalUtility` is initialized at `0`
9. Attribute `CurrentUtility` is initialized at `0`
10. Attribute `TumorSize` is initialized at `0`
11. Attribute `WorstCancer` is initialized at `0`
12. Attribute `WorstStage` is initialized at `0`
13. Attribute `CancerStageT` is initialized at `0`
14. Attribute `IsCured` is initialized at `0`
15. Attribute `StagingProb` is set to a random uniform value between `0` and `1`
16. Attribute `TumorGrowthRate` is drawn from the function `GrowRate`, and is based on previously set attribute `TumorStartAge`.
17. Attribute `max_size` is set, drawn from a normal distribution around previously set parameter `max_size` (`128`) with a standard deviation of `10`.
18. Attributes `RegressionSize` and `StagnateSize` are set using previously discussed function `RegStagSizes`.
19. Attribute `ClinSizeCheck` is set using the parameters `AlwaysFoundSize` and a lognormal distribution with `MeanLogClin` and `SdLogClin`.
20. Attribute `Invite` is initialized at `0`
21. Attribute `Screened` is initialized at `0`
22. Attribute `BIRADS` is initialized at `10`, to make sure it is invalid at the first check.
23. Attribute `Referral` is initialized at `0`
24. Attribute `NextStep` is initialized at `1`, meaning the patient should go to the `Home` trajectory next.
25. An optional `Timeout` of `30` years is built in, to skip the first `30` years of patient's lives.
26. An optional `log_` is built in for bugfixing and active patient tracking.



The visualisation of the `Init` trajectory can be seen in the figure. The resolution of the `simmer.plot` function is quite poor.

Home

```
### Home ----
```

The Home trajectory is where patients spend most of their lives. Every month, patients are checked to see if they should go to another trajectory.

```
VisitHome <- trajectory()%>%
  seize(resource='Home')%>%
  set_attribute(key= 'EntryAge',
               value=function() now(basic_sim)) %>% # record entry age
  set_attribute(key='TimeAtHome',
               value=function() WaitAtHome(
                 get_attribute(basic_sim,'EntryAge')))) %>%
  # determine how long at home (either next step or until next month)
  # time update
  timeout(function() get_attribute(basic_sim,'TimeAtHome')) %>%
  set_attribute(key = 'PatientAge',
               value = function() get_attribute(basic_sim,'TimeAtHome'),
               mod='+',init=0) %>%
  # Next Steps update
  #All five steps possible:
  #a. Go to screening after invite-> 2
  #b. Go to diagnostics due to tumor size or referral after screening ->3
  #c. Go to hospital due to referral -> 4
  #d. Die due to age or illness-> 5
  #e. Continue in Home trajectory, nothing's needed -> 1
  set_attribute(key = 'Invite',
               value= function() ifelse(get_attribute(basic_sim,'PatientAge')
                                       %in% ScreenAges,1,0),
               mod='+')%>% # point a
  set_attribute(key='ClinicalCheck',
               value=function() ClinicalCheckAtHome(
                 get_attribute(basic_sim,'TumorSize'),
                 get_attribute(basic_sim,'ClinSizeCheck'),
                 get_attribute(basic_sim,'PatientAge')))) %>%
  set_attribute('NextStep',
               value=function() AfterHome(
                 get_attribute(basic_sim,'Invite'),
                 get_attribute(basic_sim,'Screened'), # to check point
                 get_attribute(basic_sim,'BIRADS'),
                 get_attribute(basic_sim,'ClinicalCheck'),# to check pointb
                 get_attribute(basic_sim,'Referral'), # point c
                 get_attribute(basic_sim,'PatientAge'),
                 get_attribute(basic_sim,'HealthyLifeExpectancy'),
                 get_attribute(basic_sim,'TumorSize') # point D
                 )) %>%
  # Update tumor & cancer parameters:
  set_attribute('TumorSize',
               value=function() GompGrow2(
                 get_attribute(basic_sim,"PatientAge"),
                 get_attribute(basic_sim,"TumorStartAge"),
                 get_attribute(basic_sim,'TumorGrowthRate'),
                 get_attribute(basic_sim,'IsCured'),
                 get_attribute(basic_sim,'max_size'),
                 get_attribute(basic_sim,'RegressionSize'),
                 get_attribute(basic_sim,'StagnateSize')))) %>%
  set_attribute('WorstCancer',
               value = function() ifelse(
```

```

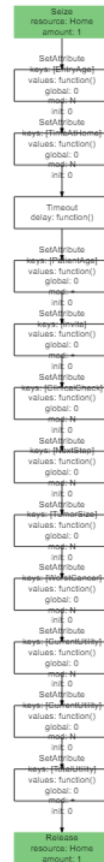
get_attribute(basic_sim,'TumorSize')>
get_attribute(basic_sim,'WorstCancer'),
get_attribute(basic_sim,'TumorSize'),
get_attribute(basic_sim,'WorstCancer')) %>%

# Update utilities
set_attribute(key='CurrentUtility',
             value = function() BaseAgeUtil(
               get_attribute(basic_sim,'PatientAge')) %>%
set_attribute(key='CurrentUtility',
             value = function() CancerBasedUtil(
               get_attribute(basic_sim,'CurrentUtility'),
               get_attribute(basic_sim,'WorstStage'),
               get_attribute(basic_sim,'PatientAge'),
               get_attribute(basic_sim,'TumorStartAge')) %>%
set_attribute(key='TotalUtility',
             value=function() get_attribute(
               basic_sim,'TimeAtHome')*
               get_attribute(basic_sim,'CurrentUtility'),
             mod='+')%>%
release(resource='Home')

```

The following steps take place in the Home Trajectory:

1. The Resource Home is seized, to track time spent at Home.
2. EntryAge of the patient is recorded.
3. TimeAtHome is set, using the function WaitAtHome. This waits until the next month.
4. A timeout takes place, to update simulation time.
5. The attribute PatientAge is updated to reflect this time spent at home.
6. Attribute Invite is set to 1 if the patient is eligible for screening.
7. ClinicalCheck is set based on the function ClinicalCheckAtHome, to see if the patient has a tumour that can be detected by the patient herself.
8. NextStep is determined using the AfterHome function.
9. The patient's TumorSize is updated using the GompGrow2 function
10. The patient's attribute WorstCancer keeps track of the largest cancer size recorded in the patient.
11. The patient's utility is updated. First, the base utility is found based on the patients age
12. Then the current utility is found using the CancerBasedUtil function
13. Then the TotalUtility is recorded by multiplying the patient's current utility with the time spent at home and adding it to the total.
14. Resource Home is released and patients are returned to the main trajectory's branch and rollback.



Screening

```
### Screening ----
```

Screening is the trajectory where patients go if they are invited and willing to go to their screening appointment. They are checked to see if they have a tumour.

```
VisitScreening <- trajectory()%>%
  seize('Screening') %>%
  set_attribute(key= 'EntryAge',value=function() now(basic_sim)) %>%
  # record entry age
  set_attribute(key='TimeAtScreening',value= function() ScreenTime()) %>%
  # determine how long at home (either next step or until next month)
  # time update
  timeout(function() get_attribute(basic_sim,'TimeAtScreening')) %>%
  set_attribute(key = 'PatientAge',
               value = function() get_attribute(basic_sim,'TimeAtScreening'),
               mod='+',
               init=0) %>%
  set_attribute(key= 'TotalTimeAtScreening',
               value = function() get_attribute(basic_sim,'TimeAtScreening')),
```

```

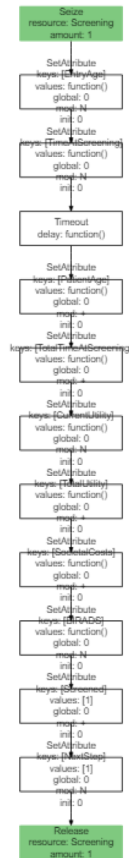
        mod='+',
        init=0) %>%

# add utility
set_attribute(key = 'CurrentUtility',
              value= function() get_attribute(
                basic_sim,'CurrentUtility') - 0.15) %>%
# disutility for during screening
set_attribute(key='TotalUtility',
              value=function() get_attribute(
                basic_sim,'TimeAtScreening')*
                get_attribute(basic_sim,'CurrentUtility'),
              mod='+')%>%
# perform screening
set_attribute(key = 'BIRADS',
              value =function() ScreenResult(
                get_attribute(basic_sim,'TumorSize')) %>%
set_attribute(key='Screened',value=1,mod='+',init=0)%>%
#check next step
set_attribute(key = 'NextStep',value= 1)%>% #go home after screening
release('Screening')

```

The following happens in the Screening Trajectory:

1. Resource Screening is seized to keep track of how often and for how long it is used.
2. EntryAge of the patient is recorded.
3. The time the patient stays at screening is determined with function ScreenTime.
4. The simulation time is updated using a timeout.
5. The PatientAge attribute is updated to reflect this.
6. The attribute TotalTimeAtScreening is updated to keep track of the total time a patient has spent at screening
7. The CurrentUtility of a patient, which was set in the Home trajectory, is updated with a minor disutility for the duration of the screening appointment
8. This utility is multiplied with the time spent at screening and added to the TotalUtility parameter
9. The screening is performed using the function ScreenResult based on the actual size of the tumor.
10. Attribute Screened is updated to reflect the number of times the patient has been screened.
11. NextStep is set to home, so that the patient will always return home after screening.
12. Resource Screening is released.



Diagnostics

```
### Diagnostics ----
```

Diagnostics is the trajectory that patients follow if they have either a BIRADS of equal or more than 3, or if they have clinically found a tumour themselves and want this diagnosed.

```
VisitDiagnostics <- trajectory()%>%
  seize('Diagnostics') %>%
  set_attribute(key= 'EntryAge',value=function() now(basic_sim)) %>%
  # record entry age
  set_attribute(key='TimeAtDiagnostics',value= function() DiagTime()) %>%
  # determine how long at home (either next step or until next month)
  # time update
  timeout(function() get_attribute(basic_sim,'TimeAtDiagnostics')) %>%
  set_attribute(key = 'PatientAge',
    value = function()
get_attribute(basic_sim,'TimeAtDiagnostics'),
  mod='+',
  init=0) %>%
```

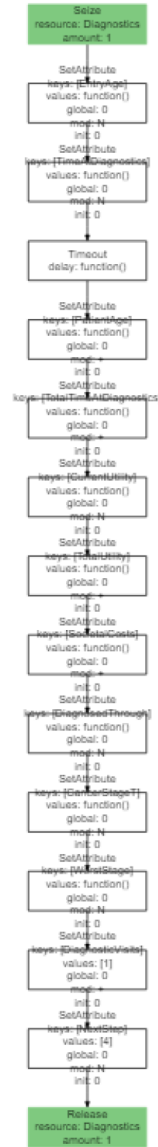
```

set_attribute(key= 'TotalTimeAtDiagnostics',
value = function()
get_attribute(basic_sim,'TimeAtDiagnostics'),
mod='+',init=0)%>%
# add utility
set_attribute(key = 'CurrentUtility',
value= function() get_attribute(
basic_sim,'CurrentUtility') - 0.25) %>% #disutility for diag
set_attribute(key='TotalUtility',
value=function() get_attribute(
basic_sim,'TimeAtDiagnostics')*
get_attribute(basic_sim,'CurrentUtility'),
mod='+')%>%
# add costs
#set_attribute(key='MedicalCosts',value= 600,mod='+',init=0) %>%
set_attribute(key='SocietalCosts',
value=function() get_attribute(
basic_sim,'TimeAtDiagnostics')/hour*35,
mod='+') %>%
set_attribute(key = 'DiagnosedThrough',
value = function() ifelse(
get_attribute(basic_sim,'ClinicalCheck'),1,0)) %>%
#through screening or clinical?
set_attribute(key = "CancerStageT",
value= function() Staging(
get_attribute(basic_sim,'TumorSize'),
get_attribute(basic_sim,'StagingProb')) %>%
set_attribute(key = "WorstStage",
value= function() ifelse(
get_attribute(basic_sim,'CancerStageT')>
get_attribute(basic_sim,'WorstStage'),
get_attribute(basic_sim,'CancerStageT'),
get_attribute(basic_sim,'WorstStage')) %>%
set_attribute('DiagnosticVisits',value = 1, mod= '+',init=0) %>%
set_attribute(key = 'NextStep',
value= function() ifelse(
get_attribute(basic_sim,'CancerStageT')== 0,
1,
4))%>%
release('Diagnostics')

```

The following happens in the VisitDiagnostics Trajectory:

1. Resource `Diagnosics` is seized to keep track of how often and for how long it is used.
2. `EntryAge` of the patient is recorded.
3. The time the patient stays at diagnostics is determined with function `DiagTime`.
4. The simulation time is updated using a timeout.
5. The `PatientAge` attribute is updated to reflect this.
6. The attribute `TotalTimeAtDiagnostics` is updated to keep track of the total time a patient has spent at diagnostics
7. The `CurrentUtility` of a patient, which was set in the `Home` trajectory, is updated with a minor disutility for the duration of the `Diagnosics` appointment
8. This utility is multiplied with the time spent at diagnostics and added to the `TotalUtility` parameter
9. The attribute `DiagnosedThrough` is set to keep track of if a patient was diagnosed through screening or clinically.
10. The stage of the cancer, recorded in attribute `CancerStageT`, is determined using function `Staging` based on the `TumorSize` and the patients initialized probability variable.
11. The attribute `WorstStage` is updated to keep track of the worst type of cancer a patient has experienced.
12. The attribute `DiagnosticVisits` is updated to reflect the number of times a patient visited `Diagnosics`
13. The attribute `NextStep` is set to 4 or 1, to send patients to the hospital if they need treatment and send them home if they have nothing.
14. Resource `Diagnosics` is released.



Cured

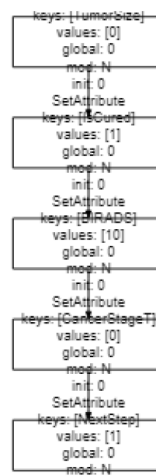
Cured ----

Cured is the trajectory that patients visit if they are cured of their cancer, with no chance of recurrence. They visit this trajectory at the end of the hospital trajectory, but it has to be initiated before the hospital trajectory.

```
Cured <- trajectory() %>%
  #log_('cured') %>%
  set_attribute(key='TumorSize', value = 0) %>%
  set_attribute(key='IsCured',value=1) %>%
  set_attribute(key='BIRADS',value=10) %>%
  set_attribute(key='CancerStageT',value=0) %>%
  set_attribute(key = 'NextStep',value=1)
```

In Cured, the following takes place:

1. The patient's TumorSize is reset to 0
2. The attribute IsCured is updated to 1
3. The attribute BIRADS is set to 10, a number where no steps take place.
4. The attribute CancerStageT is reset to 0
5. The nextStep is set to 1, returning the patient home



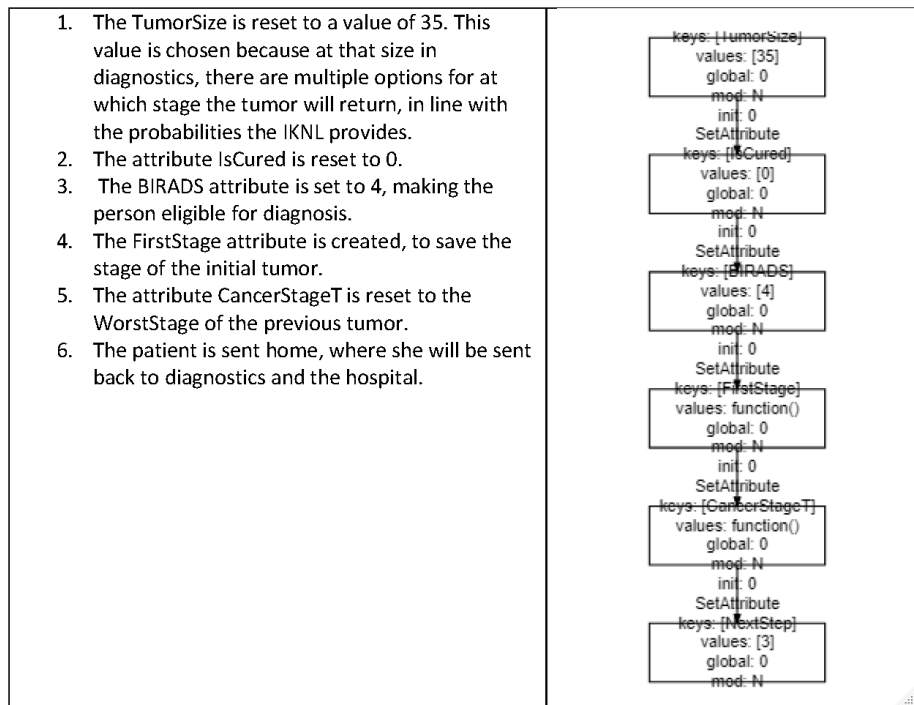
Residuals

```
### Residuals ----
```

Residuals is where patients are sent if they have a tumour that will recur at a later point in time. It is one of the two options at the end of the hospital trajectory.

```
Residual <- trajectory() %>%
  set_attribute('TumorSize', value= 35) %>%
  set_attribute(key='IsCured',value=0) %>%
  set_attribute(key='BIRADS',value=4) %>%
  set_attribute(key='FirstStage',
    value=function() get_attribute(basic_sim,'WorstStage')) %>%
  set_attribute(key='CancerStageT',
    value=function() get_attribute(basic_sim,'WorstStage')) %>%
  # go back home
  set_attribute(key = 'NextStep',value=1)
```

The following happens in this trajectory:



Hospital

```
### Hospital ----
```

In the Hospital trajectory, the treatment of the patients is simulated in a basic way. It uses averages with variation of treatments per stage instead of simulating the entire treatments. Treatment is similar for patients diagnosed clinically or through screening, so there is no difference there. The choice was made to keep this as simple as possible without compromising the results.

```
VisitHospital <- trajectory()%>%
  seize('Hospital') %>%
  set_attribute(key= 'EntryAge',
               value=function() now(basic_sim)) %>% # record entry age
  set_attribute(key='TimeAtHospital',
               value= function()
                 3^get_attribute(basic_sim,'CancerStageT')*day)%>%
  # determine how long at home (either next step or until next month)
  # time update
  timeout(function() get_attribute(basic_sim,'TimeAtHospital')) %>%
  set_attribute(key = 'PatientAge',
               value = function() get_attribute(basic_sim,'TimeAtHospital'),
               mod='+',
               init=0) %>%
  set_attribute(key= 'TotalTimeAtHospital',
               value = function() get_attribute(basic_sim,'TimeAtHospital'),
               mod='+',
               init=0) %>%
```

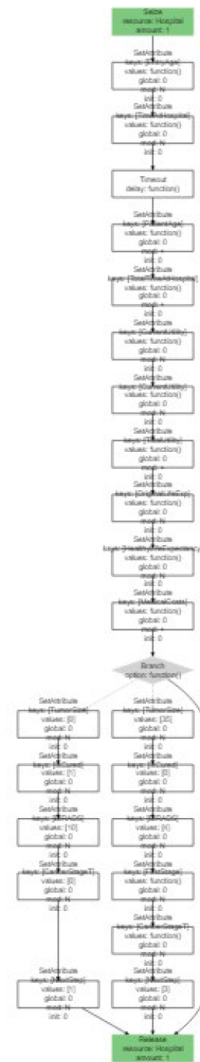
```

# update utilities
set_attribute(key='CurrentUtility',
             value = function() BaseAgeUtil(
                 get_attribute(basic_sim, 'PatientAge'))) %>%
set_attribute(key='CurrentUtility',
             value = function() CancerBasedUtil(
                 get_attribute(basic_sim, 'CurrentUtility'),
                 get_attribute(basic_sim, 'CancerStageT'),
                 get_attribute(basic_sim, 'PatientAge'),
                 get_attribute(basic_sim, 'TumorStartAge'))) %>%
set_attribute(key='TotalUtility',
             value=function() get_attribute(
                 basic_sim, 'TimeAtHospital')*
                 get_attribute(basic_sim, 'CurrentUtility'),
             mod='+')%>%
#instead of cure, go to adjust life exp
set_attribute(key='OriginalLifeExp',
             value=function() get_attribute(
                 basic_sim, 'HealthyLifeExpectancy')) %>%
set_attribute(key = 'HealthyLifeExpectancy',
             value=function() CancerSurvival(
                 get_attribute(basic_sim, 'WorstStage'),
                 get_attribute(basic_sim, 'HealthyLifeExpectancy'),
                 get_attribute(basic_sim, 'PatientAge'))) %>%
set_attribute(key='MedicalCosts',
             value=function() TreatmentCost(
                 get_attribute(basic_sim, 'WorstStage'),
                 get_attribute(basic_sim, 'PatientAge'),
                 get_attribute(basic_sim, 'HealthyLifeExpectancy')),
             mod='+') %>%
branch(option= function() ifelse(runif(1)<0.88,1,2),continue=c(T,T),
       Cured,
       Residual) %>%
release('Hospital')

```

The following steps are taken in the Hospital trajectory. Note that the plot of the trajectory also includes the trajectories for Cured and Residuals.

1. Resource Hospital is seized to keep track of how often and for how long it is used.
2. EntryAge of the patient is recorded.
3. The time the patient stays at Hospital is dependant on the stage of the tumor, with later stage tumor taking exponentially longer to treat than low stage tumors.
4. The simulation time is updated using a timeout.
5. The PatientAge attribute is updated to reflect this.
6. The attribute TotalTimeAtHospital is updated to keep track of the total time a patient has spent in treatment
7. The CurrentUtility of a patient is updated to reflect the patient's base utility
8. The CurrentUtility is updated with the function CancerBasedUtil
9. This utility is multiplied with the time spent in the Hospital trajectory and added to the TotalUtility parameter
10. The attribute HealthyLifeExpectancy is saved under new attribute OriginalLifeExpectancy
11. The HealthyLifeExpectancy is adjusted to reflect the time this patient has left to live based on their stage of cancer, using function CancerSurvival.
12. The medical costs for treatment are calculated using the function TreatmentCost
13. The trajectory now branches with a 12% chance to residual or cured and the patient ends up in one of the trajectories discussed above.
14. After the cured or residual trajectory, they are returned to this trajectory, and the patients are returned home after the Hospital resource is released



Death

```
### Death ----
```

In the Death trajectory, some calculations are made on the patients final attributes, which are also recorded. The patient's trajectories end here and the patients now die.

```
Death <- trajectory()%>%
  #log_('die') %>%
  #compute total costs
  set_attribute('MedicalCosts',
```

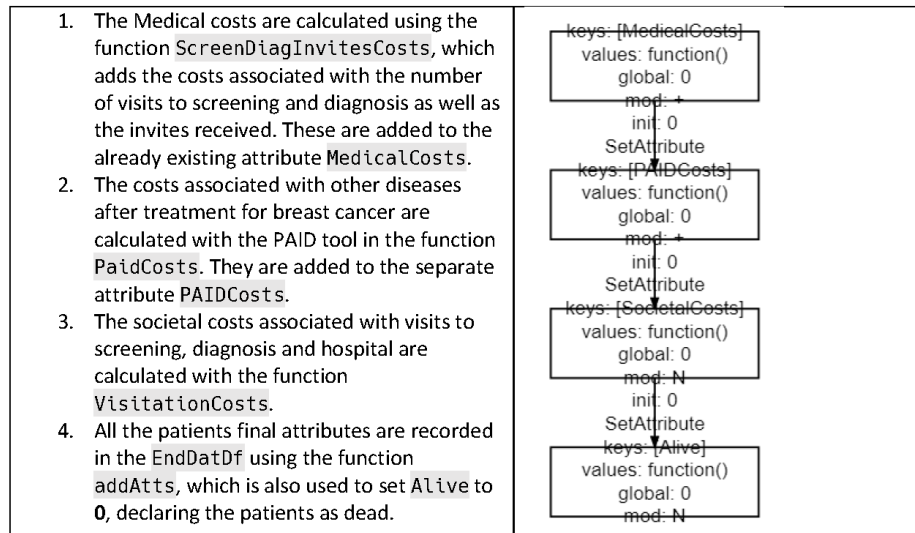
```

value = function() ScreenDiagInvitesCost(
  get_attribute(basic_sim, 'Screened'),
  get_attribute(basic_sim, 'DiagnosticVisits'),
  get_attribute(basic_sim, 'Invite')),
mod = '+' ) %>%
set_attribute(key = 'PAIDCosts',
  value = function() PaidCosts(
    get_attribute(basic_sim, 'HealthyLifeExpectancy'),
    get_attribute(basic_sim, 'TumorStartAge')),
  mod = '+' ) %>%
set_attribute(key='SocietalCosts',
  value=function() VisitationCosts(
    get_attribute(basic_sim, 'Screened'),
    get_attribute(basic_sim, 'DiagnosticVisits'),
    get_attribute(basic_sim, 'TotalTimeAtHospital')))) %>%

# extract all attributes
set_attribute(key='Alive',value = function() addAtts(EndDatDf))

```

The following happens in the trajectory when the patients die.



Base Model

```
### Base Model ----
```

The Base Model stitches all the other trajectories together and uses one branch and rollback to select where patients should be sent next.

```

#stitching the blocks together
basic_model <- trajectory()%>%
  # first initialize
  join(Initialization) %>%

  # 'store' patients at home until it's:

```

```

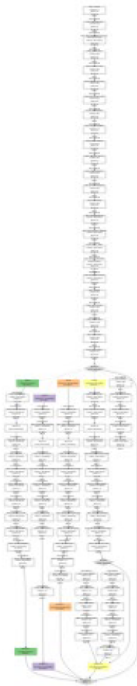
# time to go to screening
# time to go to diagnostics
# time to go to hospital
# time to die
branch(option=function() get_attribute(basic_sim, 'NextStep'),
       continue=c(T,T,T,T,F),
       VisitHome,
       VisitScreening,
       VisitDiagnostics,
       VisitHospital,
       Death) %>%

rollback(amount=1)

```

This trajectory creates a new trajectory object, which is first joined with the Initialization trajectory. After this, a branch is created that retrieves the patient's NextStep attribute, and sends the patient to either Home, Screening, Diagnostics, Hospital or Death. Patients are returned to this trajectory after finishing one of the first four, and are removed from the simulation after finishing the death trajectory. After finishing the branch, the patients are rolled back to enter another branching event.

The full model then looks like this:



Plot

```
## Plot ----
```

In plot, there are various lines of code to plot the trajectories, of which the visualisations have been shown in the separate sections above.

```
plot(basic_model,verbose=TRUE)

# plot(Initialization,verbose=T)
# plot(VisitHome,verbose=T)
# plot(VisitScreening,verbose=T)
# plot(VisitDiagnostics,verbose=T)
# plot(Cured,verbose=T)
# plot(Residual,verbose=T)
# plot(VisitHospital,verbose=T)
```

Run

```
## Run----
```

In Run, the simulation is initiated and ran. Various monitoring objects are also retrieved from the simulation.

```
basic_sim <- simmer()%>%
  add_resource('Home',capacity=Inf) %>%
  add_resource('Screening',capacity=Inf) %>%
  add_resource('Diagnostics',capacity=Inf) %>%
  add_resource('Hospital',capacity=Inf) %>%
  add_generator(name_prefix='patient',
               trajectory=basic_model,
               distribution=at(rep(x=0, times=n.patients)),
               mon=mon.patients)

start_time <- Sys.time()
basic_sim %>%
  reset() %>%
  run(progress=progress::progress_bar$new()$update,until=110*year,steps=1000)
end_time <- Sys.time()
RunTime <- end_time - start_time
print(RunTime)

patient_monitor <-
  get_mon_arrivals(basic_sim) %>%
  transform(wait = end_time - start_time - activity_time) %>%
  transform(totalTime = end_time - start_time)
patient_attributes <- get_mon_attributes(basic_sim)

testdf <- get_mon_attributes(basic_sim)

EndDatDf2 = EndDatDf[-1,]
```

First, the Simmer simulation is initialized. The resources Home, Screening, Diagnostics and Hospital are added, and the generator patient is added as movable object. It is send on the basic_model trajectory, from where it will reach different other trajectories. There is no distribution, instead all patients are spawned at the beginning of the simulation. They are monitored for either 0 or 2, based on the number of patients in the model. A '2' allows for tracking of every attribute change of every patient, a 0 will return no attributes at all.

The start time is recorded and the simulation is reset and started. A progress bar is shown, showing the time from 0 to 110 years, giving plenty of time for all patients to live their lives. After the simulation is finished, the end time is recorded and the run time is shown.

The patient_monitor, patient_attributes and testdf are only used if the monitor is set to '2'. If it's set to 0, no attributes will be recorded. The EndDatDf2 is a copy of EndDatDf created in the Death trajectory, to remove the first row containing only the word 'test'.

Reporting

```
# Reporting ----
```

In the reporting section, various KPIs are collected and plots are created.

```
incpercentages <- c()
PercThroughScreens <- c()

df <- EndDatDf

numsdF = df[-1,]
getmode <- function(v) {
  uniqv <- unique(v)
  uniqv[which.max(tabulate(match(v, uniqv)))]
}
meanAge = round(mean(round(as.numeric(numsdF$PatientAge)/year)),2)
medianAge = median(round(as.numeric(numsdF$PatientAge)/year))
modeAge = getmode(round(as.numeric(numsdF$PatientAge)/year))
patients_total <- nrow(df)
cancers_found <- table(as.numeric(df$WorstStage))
inc_perc <- round((cancers_found[2]+
                  cancers_found[3] +
                  cancers_found[4]+
                  cancers_found[5])/ patients_total * 100,2)
```

First some lists are initialized and the EndDatDf is copied to df. Then the KPIs meanAge, medianAge, modeAge, patients_total, cancers_found and inc_perc are collected to be displayed at the end.

Stage distribution Plot 2

```
## Stadiumverdeling plot 2 ----
subdf <- df %>%
  dplyr::select(WorstStage,DiagnosedThrough) %>%
  na.omit()

subdf2 <- table(subdf)
subdf3 <- as.data.frame(subdf2)
WorstStage = rep(c(0.5,1,2,3,4),3)
ThroughClin = subdf2[,2][2:6]
ThroughScreen = subdf2[,1][2:6]
ThroughTotal = ThroughClin + ThroughScreen
throughlist = append(ThroughTotal,ThroughClin)
throughlist = append(throughlist,ThroughScreen)
percThroughScreen = round(sum(ThroughScreen) /
                          (sum(ThroughScreen)+ sum(ThroughClin)) *100,2)

incpercentages<-append(incpercentages,inc_perc)
PercThroughScreens<-append(PercThroughScreens,percThroughScreen)

subdf <- df %>%
  dplyr::select('WorstStage','DiagnosedThrough','FirstStage') %>%
  mutate(WorstStage = ifelse(is.na(FirstStage),
                             WorstStage,
                             ifelse(WorstStage>FirstStage,
                                     WorstStage,
                                     FirstStage))) %>%
  select('WorstStage','DiagnosedThrough') %>%
  na.omit()
```

```

subdf2 <- table(subdf)
subdf3 <- as.data.frame(subdf2)
WorstStage = rep(c(0.5,1,2,3,4),3)
ThroughClin = subdf2[,2][1:5]
ThroughScreen = subdf2[,1][1:5]
ThroughTotal = ThroughClin + ThroughScreen
throughlist = append(ThroughTotal,ThroughClin)
throughlist = append(throughlist,ThroughScreen)
through = c('Total','Total','Total','Total','Total',
            'Clinical','Clinical','Clinical','Clinical','Clinical',
            'Screening','Screening','Screening','Screening','Screening')

data3 = data.frame(WorstStage,throughlist,through)

data3$WorstStage <- factor(data3$WorstStage,levels = c("4", "3",
                                                    "2", "1", "0.5"))

# Create a vector of colors for each level of WorstStage
colors <- c("#cbe9a6", "#f4b26f", "#e576b4", "#1a7c98", "#29bdeb")

# Calculate the total count for each level of through
total_count <- aggregate(data3$throughlist, by = list(data3$through), sum)

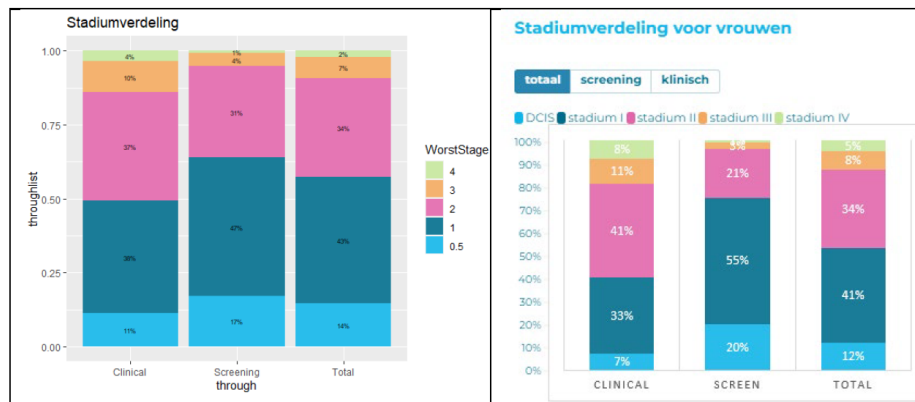
# Calculate the percentage for each combination of WorstStage and through
data3$percent <- data3$throughlist /
  total_count[match(data3$through, total_count$Group.1), "x"] * 100

plot.new()
p2<-ggplot(data3,aes(fill=WorstStage,y=throughlist,x=through))+
  geom_bar(position='fill',stat='identity')+
  labs(title='Stadiumverdeling')+
  scale_fill_manual(values=colors) +
  geom_text(aes(label = paste0(round(percent), "%")),
            position = position_fill(vjust = 0.5),size=2)

plot2 <- recordPlot()
plot.new()

```

With this code, the plot on the left is created. It aims to resemble the plot on the right, which is from IKNL data.



Percentage Plot 1

```
## Percentage plot 1 ----

subdf4 <- subdf3 %>%
  filter(DiagnosedThrough==0)

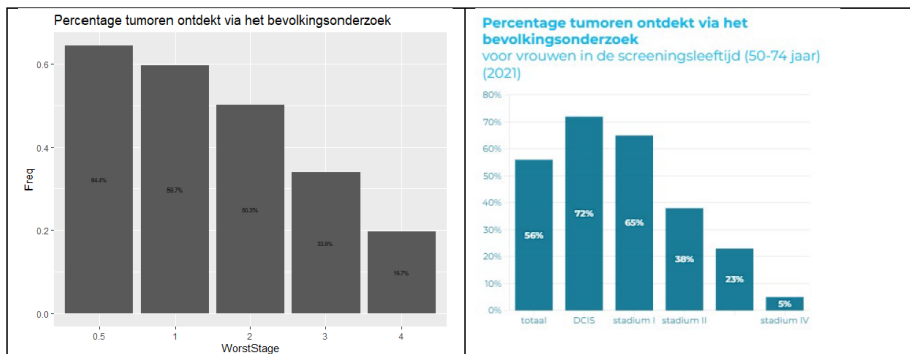
# filter subdf3 by DiagnosedThrough == 0 and drop the DiagnosedThrough column
subdf_filtered <- subset(subdf3, DiagnosedThrough == 0,
  select = -c(DiagnosedThrough))

# calculate the total count of Freq values for each WorstStage
freq_total <- aggregate(Freq ~ WorstStage, subdf3, sum)$Freq

# divide the Freq column by the total count of Freq values for each WorstStage
subdf_filtered$Freq <- subdf_filtered$Freq / freq_total
# subdf_filtered <- subset(subdf_filtered, row_number() != 6)

p1<-ggplot(subdf_filtered,aes(x=WorstStage,y=Freq))+
  geom_bar(stat='identity')+
  labs(title='Percentage tumoren ontdekt via het bevolkingsonderzoek')+
  scale_fill_brewer(palette="Spectral")+
  geom_text(aes(label = paste0(round(Freq*100, 1), "%"),
    position = position_stack(vjust = 0.5),size=2)
plot1 <- recordPlot()
plot.new()
```

With this code, the plot on the left is created. It aims to resemble the plot on the right, which is from IKNL data. Note that this version did not yet include a total tally, the leftmost column of the IKNL graph.



Age of Diagnosis Plot 3

```
## leeftijd diagnose plot 3 ----

subdf5 <- df %>%
  dplyr::select('WorstStage', 'TumorStartAge') %>%
  na.omit() %>%
  filter(TumorStartAge<100*year) %>%
  filter(WorstStage>0) %>%
  mutate(TumorStartAge = as.numeric(TumorStartAge)) %>%
  mutate(ages = cut(TumorStartAge,breaks=c(0,50*year,74*year,Inf))) %>%
  dplyr::select(ages,WorstStage)

subdf6 = rbind(as.data.frame(table(subdf5)))
```

```

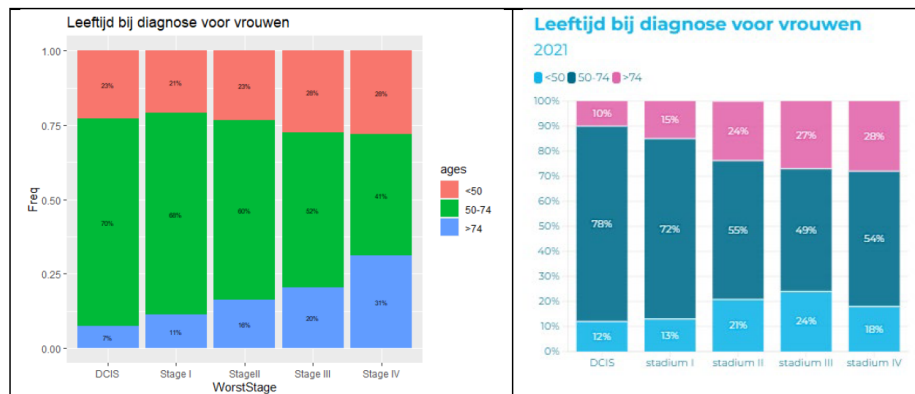
# Calculate the total count for each level of through
total_count2 <- aggregate(subdf6$Freq, by = list(subdf6$WorstStage), sum)

# Calculate the percentage for each combination of WorstStage and through
subdf6$percent <- subdf6$Freq /
  total_count2[match(subdf6$WorstStage, total_count2$Group.1), "x"] * 100

p3<- ggplot(subdf6,aes(fill=ages,y=WorstStage,x=Freq))+
  geom_bar(position='fill',stat='identity')+
  coord_flip()+
  scale_fill_discrete(labels=c('<50', '50-74', '>74'))+
  scale_y_discrete(labels=c("0.5" = "DCIS",
                             "1" = "Stage I",
                             "2" = "StageII",
                             "3" = "Stage III",
                             "4" = "Stage IV"))+
  labs(title='Leeftijd bij diagnose voor vrouwen')+
  geom_text(aes(label = paste0(round(percent), '%'),
                position = position_fill(vjust = 0.5),size=2))
plot3 <- recordPlot()
plot.new()

```

With this code, the plot on the left is created. It aims to resemble the plot on the right, which is from IKNL data.



Survival Curve Plot 4

```
## Survival curve plot 4 ----
```

```

subdf7 <- df%>%
  dplyr::select(WorstStage, PatientAge, TumorStartAge, OriginalLifeExp)%>%
  filter(TumorStartAge<100*year) %>%
  filter(WorstStage>0) %>%
  mutate(WorstStage = as.numeric(WorstStage)) %>%
  mutate(TumorStartAge = as.numeric(TumorStartAge)/year) %>%
  mutate(PatientAge = as.numeric(PatientAge)/year) %>%
  mutate(OriginalLifeExp = as.numeric(OriginalLifeExp)/year) %>%
  mutate(Survival = PatientAge - TumorStartAge) %>%
  mutate(ShouldveSurvived = OriginalLifeExp - TumorStartAge)%>%
  mutate(One = ifelse(round(Survival)>=1,1,
                      ifelse(round(ShouldveSurvived)>=1,0,NA))) %>%
  mutate(Two = ifelse(round(Survival)>=2,1,
                      ifelse(round(ShouldveSurvived)>=2,0,NA))) %>%

```



```

mutate(Three = ifelse(round(Survival)>=3,1,
                      ifelse(round(ShouldveSurvived)>=3,0,NA))) %>%
mutate(Four = ifelse(round(Survival)>=4,1,
                      ifelse(round(ShouldveSurvived)>=4,0,NA))) %>%
mutate(Five = ifelse(round(Survival)>=5,1,
                      ifelse(round(ShouldveSurvived)>=5,0,NA))) %>%
mutate(Six = ifelse(round(Survival)>=6,1,
                     ifelse(round(ShouldveSurvived)>=6,0,NA))) %>%
mutate(Seven = ifelse(round(Survival)>=7,1,
                       ifelse(round(ShouldveSurvived)>=7,0,NA))) %>%
mutate(Eight = ifelse(round(Survival)>=8,1,
                       ifelse(round(ShouldveSurvived)>=8,0,NA))) %>%
mutate(Nine = ifelse(round(Survival)>=9,1,
                      ifelse(round(ShouldveSurvived)>=9,0,NA))) %>%
mutate(Ten = ifelse(round(Survival)>=10,1,
                     ifelse(round(ShouldveSurvived)>=10,0,NA)))

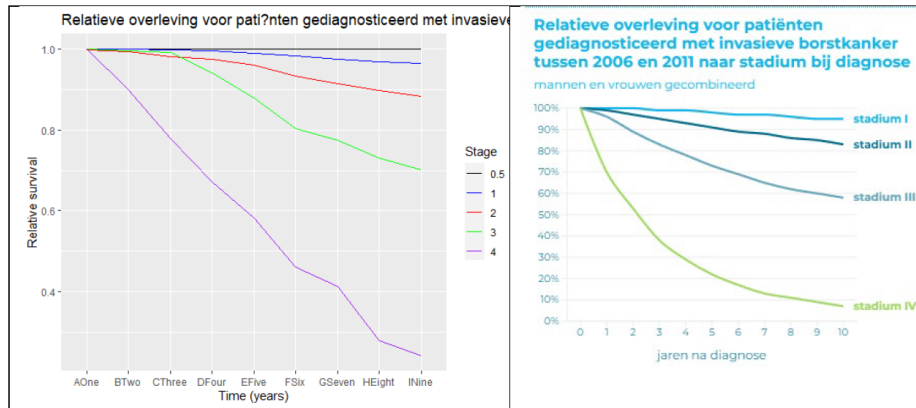
subdf7 <- subdf7 %>%
group_by(WorstStage) %>%
summarise(
  AAZero = c(1,1,1,1,1),
  AOne = mean(One, na.rm = TRUE),
  BTwo = mean(Two, na.rm = TRUE),
  CThree = mean(Three, na.rm = TRUE),
  DFour = mean(Four, na.rm = TRUE),
  EFive = mean(Five, na.rm = TRUE),
  FSix = mean(Six, na.rm = TRUE),
  GSeven = mean(Seven, na.rm = TRUE),
  HEight = mean(Eight, na.rm = TRUE),
  INine = mean(Nine, na.rm = TRUE),
  JTen = mean(Ten, na.rm = TRUE)
)

subdf7
# convert from wide to long format
df_long <- tidyr::pivot_longer(subdf7, cols = c(2:11),
                               names_to = "Column", values_to = "Value")

# create the line plot
p4<-ggplot(data = df_long, aes(x = Column, y = Value,
                              group = WorstStage, color =
factor(WorstStage))) +
  geom_line() +
  scale_color_manual(values = c("black", "blue", "red", "green", "purple")) +
  labs(x = "Time (years)", y = "Relative survival", color = "Stage",
       title='Relatieve overleving voor patiënten gediagnosticeerd met
invasieve borstkanker naar stadium bij diagnose')
plot4 <- recordPlot()
plot.new()

```

With this code, the plot on the left is created. It aims to resemble the plot on the right, which is from IKNL data.

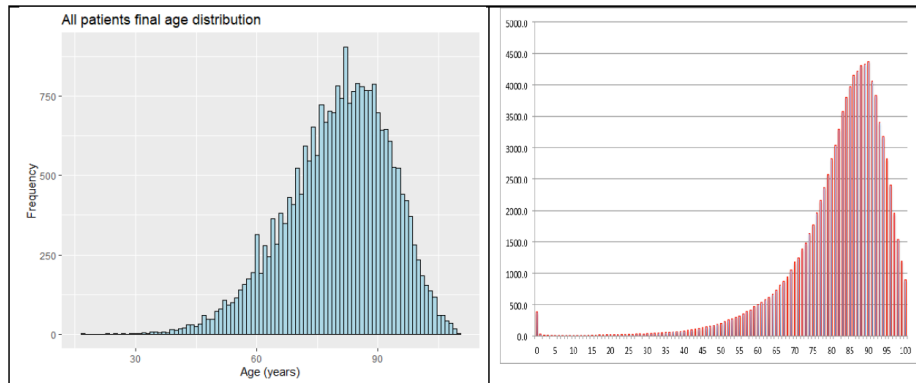


Final Age Distribution

More plots for finetuning

```
## Final age distribution ----
p5 <- ggplot(df, aes(x = as.numeric(PatientAge) / year)) +
  geom_histogram(binwidth = 1, color = "black", fill = "lightblue") +
  labs(title = "All patients final age distribution",
       x = "Age (years)",
       y = "Frequency")
```

With this code, the plot on the left is created showing the distribution of the final ages of all patients. This compares to the actual life expectancy in the Netherlands. The graph on the right is included for demonstrative purposes and based on data from the UK: Numbers of women expected to die at each age, out of 100,000 born, assuming mortality rates stay the same as 2010-2012. The expectation is 83, median 86 (*Why "life Expectancy" Is a Misleading Summary of Survival | Understanding Uncertainty, n.d.*).

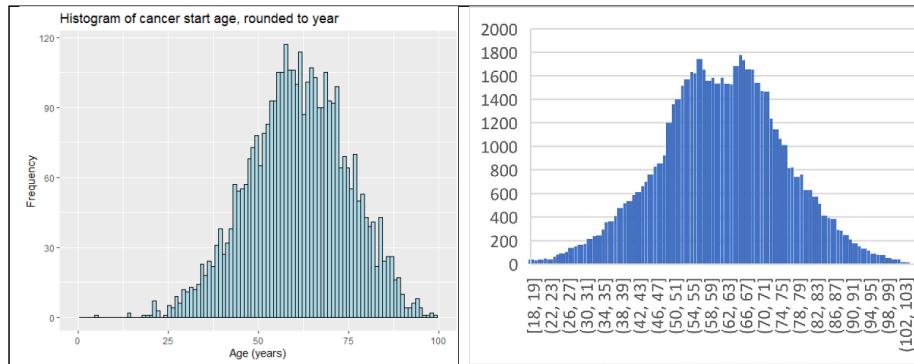


Histogram Cancer Start age

```
## Histogram cancer start age ----
subdf8 <- df %>%
  filter(TumorStartAge < 1000 * year)
```

```
p6 <- ggplot(subdf8, aes(x = as.numeric(TumorStartAge) / year)) +
  geom_histogram(bins = 100, color = "black", fill = "lightblue") +
  xlim(0, 100) +
  labs(title = "Histogram of cancer start age, rounded to year",
       x = "Age (years)",
       y = "Frequency")
```

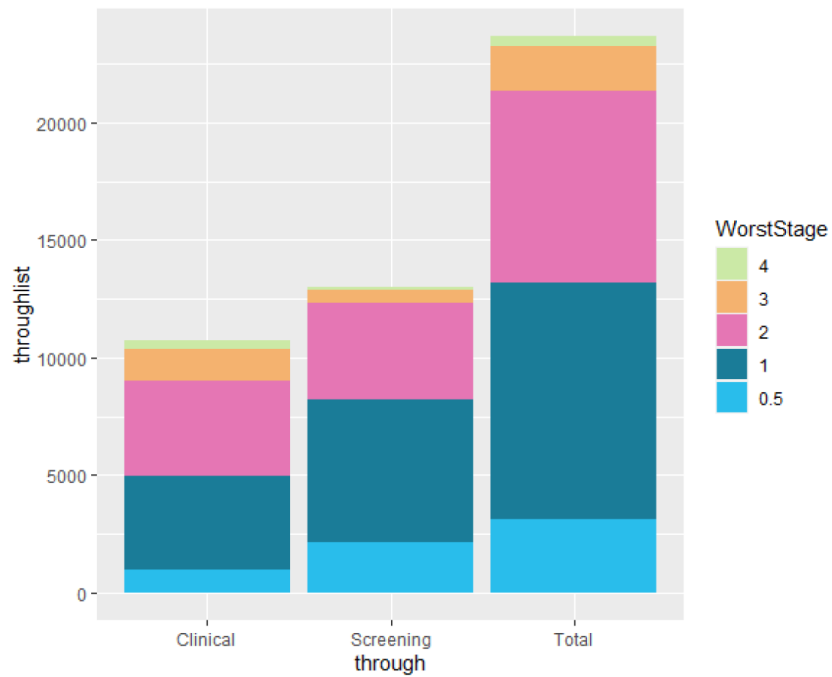
This plot shows the histogram of ages at which patients get tumor. The plot from the simulation is on the left, the histogram from the IKNL data is on the right.



Actual Stage Distribution

```
## actual stage distribtuion ----
p7 <- ggplot(data3, aes(fill=WorstStage, y=throughlist, x=through)) +
  geom_bar(position='stack', stat='identity') +
  scale_fill_manual(values=colors)
```

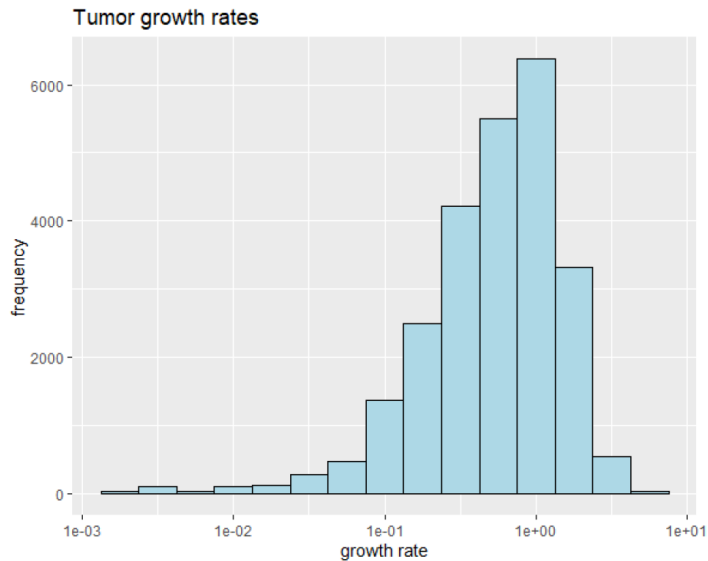
This plot shows the non-normalized stage distributions, to visually inspect which percentage is due to screening and which due to clinical diagnosis.



Tumor Growth Rates Histogram

```
## Tumor growth rates hist ----
p8 <- ggplot(df, aes(x=as.numeric(TumorGrowthRate)))+
  geom_histogram(binwidth = 0.25, color='black', fill='lightblue')+
  labs(title='Tumor growth rates',
        x='growth rate',
        y='frequency')+
  scale_x_log10()
```

This plot shows a histogram of tumor growth rates on a logistic scale.



Plot Stage Sizes Histogram

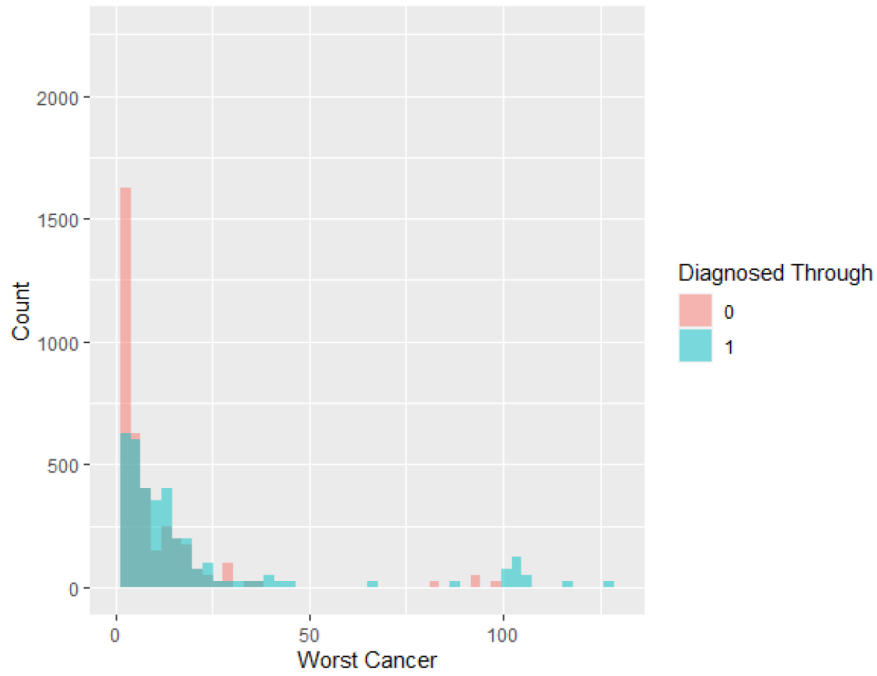
```
#finetune stagefromsize diagram plot
## plot stage size hist ----

SizeTestdf<- df %>%
  dplyr::select(DiagnosedThrough,WorstCancer,WorstStage) %>%
  filter(WorstStage == 1) %>%
  mutate(WorstCancer = as.numeric(WorstCancer))

ggplot(SizeTestdf, aes(x = WorstCancer, fill = factor(DiagnosedThrough))) +
  geom_histogram(alpha = 0.5, position = "identity", bins = 50) +
  scale_fill_discrete(name = "Diagnosed Through") +
  labs(title = "Histogram of Worst Cancer by Diagnosis Method",
       x = "Worst Cancer", y = "Count")+
  xlim(c(0,130))
```

This plot gives an overview of at which sizes certain stage tumors have been diagnosed most often, for finetuning purposes and gaining insights in the workings of the model.

Histogram of Stage I Cancer by Diagnosis Method



KPIs

The following code renders the KPIs and displays them.

```
patients_total <- nrow(df)
cancers_found <- table(as.numeric(df$WorstStage))
inc_perc <- round((cancers_found[2]+
                  cancers_found[3] +
                  cancers_found[4]+
                  cancers_found[5])/ patients_total * 100,2)
percThroughScreen = round(sum(ThroughScreen) /
                          (sum(ThroughScreen)+ sum(ThroughClin)) *100,2)

numsdF = df[-1,]
meanAge = round(mean(round(as.numeric(numsdF$PatientAge)/year)),2)
medianAge = median(round(as.numeric(numsdF$PatientAge)/year))
modeAge = getmode(round(as.numeric(numsdF$PatientAge)/year))

CostUtilDf <- df %>%
  select(TotalUtility,TotalCosts,MedicalCosts,SocietalCosts,PAIDCosts) %>%
  mutate_all(function(x) as.numeric(x)) %>%
  mutate(TotalCosts = SocietalCosts + MedicalCosts+PAIDCosts)

meanQALY <- mean(CostUtilDf$TotalUtility,na.rm = T) /year
meanTotalCosts <- mean(CostUtilDf$TotalCosts,na.rm = T)
meanMedicalCosts <- mean(CostUtilDf$MedicalCosts,na.rm = T)
meanSocietalCosts <- mean(CostUtilDf$SocietalCosts,na.rm = T)
meanPaidCosts <- mean(CostUtilDf$PAIDCosts,na.rm=T)
```

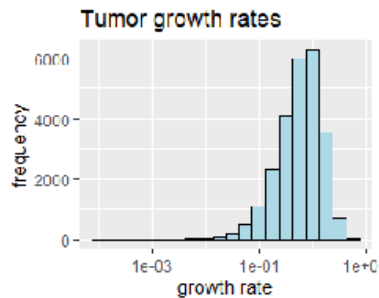
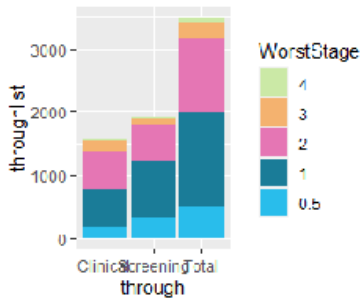
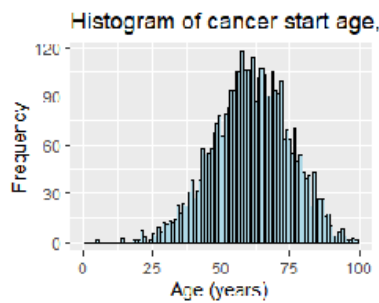
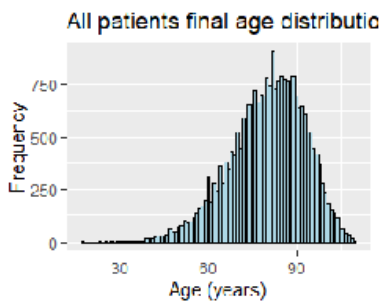
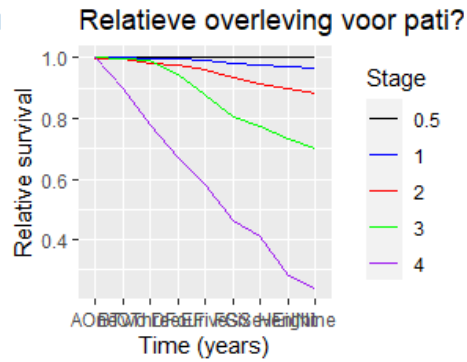
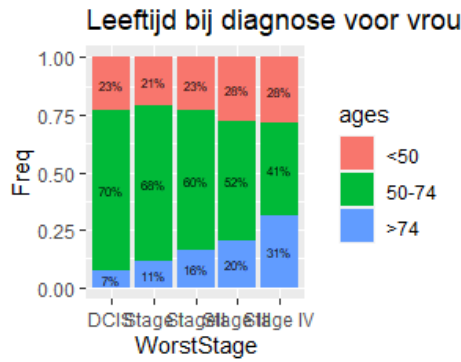
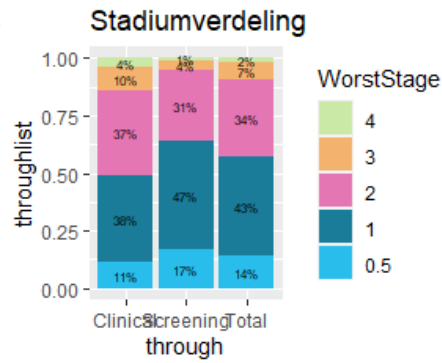
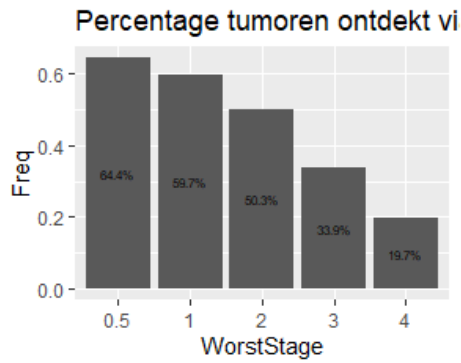
```
## KPIs ----
cat(paste0(strrep("_", 70), '\n',
  'OUTPUT KPIs', '\n',
  strrep("_", 70), '\n',
  'Patients generated: ', patients_total, '\n',
  'incidence percentage: ', inc_perc, '%', '\n',
  'Percentage found through screening: ', percThroughScreen, '%', '\n',
  'Patient Ages: mean: ', meanAge, ' mode: ', modeAge, ' median: ',
  medianAge, '\n',
  'Mean Quality Adjusted Life Expectancy: ', meanQALY, '\n',
  'Mean Total Costs: ', meanTotalCosts, '\n',
  'Mean Medical Costs: ', meanMedicalCosts, '\n',
  'Mean Societal Costs: ', meanSocietalCosts, '\n',
  'Mean PAID Costs: ', meanPaidCosts, '\n'))
```

```
OUTPUT KPIs
Patients generated: 24969
incidence percentage: 13.64%
Percentage found through screening: 54.54%
Patient Ages: mean:80.27 mode: 82 median: 82
Mean Quality Adjusted Life Expectancy: 71.743632132996
Mean Total Costs: 18479.1574273357
Mean Medical Costs: 6873.01761111848
Mean Societal Costs: 499.81419110646
Mean PAID Costs: 11106.3256251108
```

Grid plots

The following code shows two 2x2 grids containing the first eight plots, so that each simulation immediately provides output.

```
grid.arrange(p5,p6,p7,p8,nrow=2)
grid.arrange(p1,p2,p3,p4,nrow=2)
```



Bibliography (for full code explanation appendix)

- Arrospide, A., Rue, M., van Ravesteyn, N. T., Comas, M., Soto-Gordoa, M., Sarriguarte, G., & Mar, J. (2016). Economic evaluation of the breast cancer screening programme in the Basque Country: retrospective cost-effectiveness and budget impact analysis. *BMC Cancer*, *16*(1). <https://doi.org/10.1186/S12885-016-2386-Y>
- Integraal Kankercentrum Nederland. (2022). *Borstkankercijfers*. <https://iknl.nl/borstkankercijfers>
- Integraal Kankercentrum Nederland (IKNL). (2022). *Synthetische dataset Nederlandse Kankerregistratie (NKR)*. <https://iknl.nl/NKR/Cijfers-Op-Maat/Synthetische-Data>.
- National Cancer Institute. (2022). *CISNET: Modeling Approach*. <https://cisnet.cancer.gov/modeling/index.html>
- PAID 3.0. (n.d.). Retrieved March 19, 2023, from <https://imta.shinyapps.io/PAID3code/>
- PriceWaterhouseCoopers Accountants N.V. (2022). *Jaarverslag Stichting Bevolkingsonderzoek Nederland 2021*.
- Rafia, R., Brennan, A., Madan, J., Collins, K., Reed, M. W. R., Lawrence, G., Robinson, T., Greenberg, D., & Wyld, L. (2016). Modeling the Cost-Effectiveness of Alternative Upper Age Limits for Breast Cancer Screening in England and Wales. *Value in Health : The Journal of the International Society for Pharmacoeconomics and Outcomes Research*, *19*(4), 404–412. <https://doi.org/10.1016/J.JVAL.2015.06.006>
- Trentham-Dietz, A., Alagoz, O., Chapman, C., Huang, X., Jayasekera, J., van Ravesteyn, N. T., Lee, S. J., Schechter, C. B., Yeh, J. M., Plevritis, S. K., & Mandelblatt, J. S. (2021). Reflecting on 20 years of breast cancer modeling in CISNET: Recommendations for future cancer systems modeling efforts. *PLoS Computational Biology*, *17*(6). <https://doi.org/10.1371/JOURNAL.PCBI.1009020>
- Ucar, I., Hernández, J. A., Serrano, P., & Azcorra, A. (2018). Design and Analysis of 5G Scenarios with simmer: An R Package for Fast DES Prototyping. *IEEE Communications Magazine*, *56*(11), 145–151. <https://doi.org/10.1109/MCOM.2018.1700960>
- van Haperen, V. R. (2018). *Factsheet Bevolkingsonderzoek Borstkanker*. <https://www.rivm.nl/sites/default/files/2019-02/Factsheet%20borstkanker%202018%20DEF.pdf>
- van Luijt, P. A., Heijnsdijk, E. A. M., & de Koning, H. J. (2017). Cost-effectiveness of the Norwegian breast cancer screening program. *International Journal of Cancer*, *140*(4), 833–840. <https://doi.org/10.1002/IJC.30513>
- Why “life expectancy” is a misleading summary of survival | Understanding Uncertainty. (n.d.). Retrieved April 10, 2023, from <https://understandinguncertainty.org/why-life-expectancy-misleading-summary-survival>
- Wright, S., Gray, E., Rogers, G., Donten, A., Hainsworth, R., & Payne, K. (2022). *MANC-RISK-SCREEN Version 1.0* (1.0). University of Manchester. <https://research.manchester.ac.uk/en/datasets/manc-risk-screen-version-10>
- Zorginstituut Nederland. (2016). *Guideline for economic evaluations in healthcare*.

Appendix IX – AdViSHE questionnaire

Vemer P, Corro Ramos I, Van Vroom GAK, Al MJ, Feenstra TL and the AdViSHE Study Group "AdViSHE: A Validation-Assessment Tool of Health-Economic Models for Decision Makers and Model Users" *Pharmacoeconomics* (2016) 34:349–361. DOI: 10.1007/s40273-015-0327-2

AdViSHE

Assessment of the Validation Status of Health-Economic decision models

AdViSHE is a questionnaire that modellers can complete to report on the efforts performed to improve the validation status of their health-economic (HE) decision model. It is not intended to replace validation by model users but rather to inform the direction of validation efforts and to provide a baseline for replication of the results. In addition to using it after a model is finished, the modellers can use AdViSHE to guide validation efforts during the modelling process.

The modellers are asked to comment on the validation efforts performed while building the underlying HE decision model and afterwards. Many of the questions simply refer to the model documentation. AdViSHE is divided into five parts, each covering an aspect of validation:

- Part A: Validation of the conceptual model (2 questions)
- Part B: Input data validation (2 questions)
- Part C: Validation of the computerized model (4 questions)
- Part D: Operational validation (4 questions)
- Part E: Other validation techniques (1 question)

No final validation score is calculated, as the assessment of the answers and the overall validation effort is left to the model users. It is assumed that the model has been built according to prevailing modelling and reporting guidelines. For instance, the model builders would presumably adhere to the ISPOR-SMDM[†] Modeling Good Research Practices (Caro et al., 2010) and/or CHEERS[†] Statement (Husereau et al., 2013). Some questions may not be applicable to a particular model. If this is the case, the model builder should take the opt-out option and provide a justification of why this item is not deemed applicable.

Part A: Validation of the conceptual model (2 questions)

Part A discusses techniques for validating the conceptual model. A conceptual model describes the underlying system (e.g., progression of disease) using a mathematical, logical, verbal, or graphical representation. Please indicate where the conceptual model and its underlying assumptions are described and justified.

The conceptual model is described and justified in chapter 4: Methods. The model consists of one patient level model with different trajectories. Disease progression is determined through a separate tumor growth model. Diagnosis is determined through a separate tumor detection model, which is related to the MISCAN model. The staging model is also derived from literature. All models have logical, verbal and graphical representation in the paper. Wherever advanced mathematics are used, the use is explained in detail.

1

[†]: ISPOR: International Society For Pharmacoeconomics and Outcomes Research, SMDM: Society for Medical Decision making, CHEERS: Consolidated Health Economic Evaluation Reporting Standards

A1/ Face validity testing (conceptual model): Have experts been asked to judge the appropriateness of the conceptual model?

If yes, please provide information on the following aspects:

- Who are these experts?
- What is your justification for considering them experts?
- To what extent do they agree that the conceptual model is appropriate?

If no, please indicate why not.

Yes

- Xavier Pouwels
 - PHd in health-economic modelling
 - Appropriate and told me what to look out for
- Talitha Ahmed
 - Statistician at Accenture
 - Appropriate, helped with statistics

Aspects to judge include: appropriateness to represent the underlying clinical process/disease (disease stages, physiological processes, etc.); and appropriateness for economic evaluation (comparators, perspective, costs covered, etc.).

A2/ Cross validity testing (conceptual model): Has this model been compared to other conceptual models found in the literature or clinical textbooks?

If yes, please indicate where this comparison is reported.

If no, please indicate why not.

Conceptual model was based on models in the CISNET repository and the MANC SCREEN Model. Comparison reported in methods and validation sections.

Part B: Input data validation (2 questions)

Part B discusses techniques to validate the data serving as input in the model. These techniques are applicable to all types of models commonly used in HE modelling.

Please indicate where the description and justification of the following aspects are given:

- search strategy;
- data sources, including descriptive statistics;
- reasons for inclusion of these data sources;
- reasons for exclusion of other available data sources;
- assumptions that have been made to assign values to parameters for which no data was available;
- distributions and parameters to represent uncertainty;
- data adjustments: mathematical transformations (e.g., logarithms, squares); treatment of outliers; treatment of missing data; data synthesis (indirect treatment comparison, network meta-analysis); calibration; etc.

Search strategy: systematic review was performed, data partly pulled & aggregated from those papers. Others from public, reliable government sources

Data sources: All provided where statistics are presented in the paper and in program code

Reason for data source inclusion: data sources deemed valid have been included. Often from recommended & peer reviewed papers or governmental data.

Reason for exclusion of other sources: only reliable sources were used. It was often easier to get more information from one source or linked sources, as data correlation between different input variables makes it difficult to exchange parts of this data with data from other sources

Assumptions that have been made: All data has been available in some way from sources.

Assumptions regarding data have been explained at points where the data was used.

Distributions and parameters to represent uncertainty: extensive distribution fitting was done to various data sources, taking the best-fitting and logically sound distribution each time. For PSA, means of these distributions can be varied upon. All this is explained in the paper where distributions are first introduced.

Data adjustments: all adjustments are explained in the paper, most in chapter 4: methods

B1/ Face validity testing (input data): Have experts been asked to judge the appropriateness of the input data?

If yes, please provide information on the following aspects:

- Who are these experts?
- What is your justification for considering them experts?
- To what extent do they agree that appropriate data has been used?

If no, please indicate why not.

All input data was based on reputable literature, all sourced in the thesis. Various experts consulted throughout to see if they agree with decision making. Including health scientists, an MD, and statisticians.

Aspects to judge may include but are not limited to: potential for bias; generalizability to the target population; availability of alternative data sources; any adjustments made to the data.

B2/ Model fit testing: When input parameters are based on regression models, have statistical tests been performed?

If yes, please indicate where the description, the justification and the outcomes of these tests are reported.
If no, please indicate why not.

Yes, all input data that has been fitted and resulted in output data that has undergone statistical testing to prove *likelihood* of similarity of distributions.

Examples of regression models include but are not limited to: disease progression based on survival curves; risk profiles using regression analysis on a cohort; local cost estimates based on multi-level models; meta-regression; quality-of-life weights estimated using discrete choice analysis; mapping of disease-specific quality-of-life weights to utility values.

Examples of tests include but are not limited to: comparing model fit parameters (R^2 , Akaike information criterion (AIC), Bayesian information criterion (BIC)); comparing alternative model specifications (covariates, distributional assumptions); comparing alternative distributions for survival curves (Weibull, lognormal, logit); testing the numerical stability of the outcomes (sufficient number of iterations); testing the convergence of the regression model; visually testing model fit and/or regression residuals.

Part C: Validation of the computerized model (4 questions)

Part C discusses various techniques for validating the model as it is implemented in a software program. If there are any differences between the conceptual model (Part A) and the final computerized model, please indicate where these differences are reported and justified.

The conceptual model is a fair representation of the actual computerized model. There are no significant changes between the conceptual framework and the actual computerized implementation of the models. The theoretical framework is translated into code to the best of the authors abilities, and has, as far as can reasonably be assumed, not led to drastic differences in the models.

C1/ External review: Has the computerized model been examined by modelling experts?

If yes, please provide information on the following aspects:

- Who are these experts?
- What is your justification for considering them experts?
- Can these experts be qualified as independent?
- Please indicate where the results of this review are reported, including a discussion of any unresolved issues.

If no, please indicate why not.

Examined by fellow students, Accenture colleagues and PhDs at various points.

Aspects to judge may include but are not limited to: absence of apparent bugs; logical code structure optimized for speed and accuracy; appropriate translation of the conceptual model.

C2/ Extreme value testing: Has the model been run for specific, extreme sets of parameter values in order to detect any coding errors?

If yes, please indicate where these tests and their outcomes are reported.

If no, please indicate why not.

Yes. PSA has been performed and is listed under results.

Examples include but are not limited to: zero and extremely high (background) mortality; extremely beneficial, extremely detrimental, or no treatment effect; zero or extremely high treatment or healthcare costs.

C3/ Testing of traces: Have patients been tracked through the model to determine whether its logic is correct?

If yes, please indicate where these tests and their outcomes are reported.
If no, please indicate why not.

Yes, patients have been fully traced through the simulation model. This was done at small sample sizes. For example, simulating only 10 patients already gives 88000 data points that can be checked.

Patients' performance in the model is also monitored by the graphs at the end of each simulation. Oddities would become visible there as well.

In cohort models, this would involve listing the number of patients in each disease stage at one, several, or all time points (e.g., Markov traces). In individual patient simulation models, this would involve following several patients throughout their natural disease progression.

C4/ Unit testing: Have individual sub-modules of the computerized model been tested?

If yes, please provide information on the following aspects:

- Was a protocol that describes the tests, criteria, and acceptance norms defined beforehand?
- Please indicate where these tests and their outcomes are reported.

If no, please indicate why not.

Yes

The underlying sub-modules (tumor growth model, tumor detection model, staging model) have all been tested separately from the large simulation and results were in line with that of literature.

Various trajectories can also be turned off without hampering the overall patient flow. All trajectories are modelled individually and linked at the end. These 7 sub trajectories can be found in the appendix as well, with explanation of what happens.

Examples include but are not limited to: turning sub-modules of the program on and off; altering global parameters; testing messages (e.g., warning against illegal or illogical inputs), drop-down menus, named areas, switches, labelling, formulas and macros; removing redundant elements.

Part D: Operational validation (4 questions)

Part D discusses techniques used to validate the model outcomes.

D1/ Face validity testing (model outcomes): Have experts been asked to judge the appropriateness of the model outcomes?

If yes, please provide information on the following aspects:

- Who are these experts?
- What is your justification for considering them experts?
- To what extent did they conclude that the model outcomes are reasonable?

If no, please indicate why not.

Various fellow students & Accenture colleagues asked to judge model outcome graphs.

Experts: Accenture – Talitha, UT – Xavier, Erwin & Ton

Outcomes may include but are not limited to: (quality-adjusted) life years; deaths; hospitalizations; total costs.

D2/ Cross validation testing (model outcomes): Have the model outcomes been compared to the outcomes of other models that address similar problems?

If yes, please provide information on the following aspects:

- Are these comparisons based on published outcomes only, or did you have access to the alternative model?
- Can the differences in outcomes between your model and other models be explained?
- Please indicate where this comparison is reported, including a discussion of the comparability with your model.

If no, please indicate why not.

Outcomes compared.

Medical outcomes similar: reduction in late stage tumors accomplished, increase in total cancers seen.
Utility outcomes different: inclusion of more utility values in this model leading to non-significant differences in total QoL overall.
Cost outcomes explicably different due to inclusion of more costs.

Other models may include models that describe the same disease, the same intervention, and/or the same population.

D3/ Validation against outcomes using alternative input data: Have the model outcomes been compared to the outcomes obtained when using alternative input data? If yes, please indicate where these tests and their outcomes are reported. If no, please indicate why not.

This has been done as part of the probabilistic sensitivity analysis, where testing with extreme values led to new insights. This is in the thesis.

Alternative input data can be obtained by using different literature sources or datasets, but can also be constructed by splitting the original data set in two parts, and using one part to calculate the model outcomes and the other part to validate against.

D4/ Validation against empirical data: Have the model outcomes been compared to empirical data?

If yes, please provide information on the following aspects:

- Are these comparisons based on summary statistics, or patient-level datasets?
- Have you been able to explain any difference between the model outcomes and empirical data?
- Please indicate where this comparison is reported.

If no, please indicate why not.

D4.A/ Comparison against the data sources on which the model is based (dependent validation).

Yes, continuously

Compared with graphs and kpis literature outcomes

D4.B/ Comparison against a data source that was not used to build the model (independent validation).

Yes, continuously

Compared with graphs and kpis to IKNL figures and CBS data

Part E: Other validation techniques (1 question)

E1/ Other validation techniques: Have any other validation techniques been performed?

If yes, indicate where the application and outcomes are reported, or else provide a short summary here.

Walk throughs performed by following patients in computerized model to see if conceptual framework is followed.

Double programming for lots of different aspects and functions between R and Python.

Naïve benchmarking: what should be the expected outcome when changing a parameter: often done with excel to check some outputs, e.g. with the PSA.

Examples of other validation techniques: structured "walk-throughs" (guiding others through the conceptual model or computerized program step-by-step); naive benchmarking ("back-of-the-envelope" calculations); heterogeneity tests; double programming (two model developers program components independently and/or the model is programmed in two different software packages to determine if the same results are obtained).

Appendix X – 2K Factorial Experiment Design

2k Factorial design for screening strategies

Screening interval: 2 years

Compute time estimate

100 patient runtime: 15 sec
 processors -1: 15
 pats per exp 10000
 Experiments 256
time: 25600
mins 426.6667
hours 7.111111

Exp No	40-45	45-50	50-55	55-60	60-65	65-70	70-75	75-80
1	-	-	-	-	-	-	-	-
2	+	-	-	-	-	-	-	-
3	-	+	-	-	-	-	-	-
4	+	+	-	-	-	-	-	-
5	-	-	+	-	-	-	-	-
6	+	-	+	-	-	-	-	-
7	-	+	+	-	-	-	-	-
8	+	+	+	-	-	-	-	-
9	-	-	-	+	-	-	-	-
10	+	-	-	+	-	-	-	-
11	-	+	-	+	-	-	-	-
12	+	+	-	+	-	-	-	-
13	-	-	+	+	-	-	-	-
14	+	-	+	+	-	-	-	-
15	-	+	+	+	-	-	-	-
16	+	+	+	+	-	-	-	-
17	-	-	-	-	+	-	-	-
18	+	-	-	-	+	-	-	-
19	-	+	-	-	+	-	-	-
20	+	+	-	-	+	-	-	-
21	-	-	+	-	+	-	-	-
22	+	-	+	-	+	-	-	-
23	-	+	+	-	+	-	-	-
24	+	+	+	-	+	-	-	-
25	-	-	-	+	+	-	-	-
26	+	-	-	+	+	-	-	-
27	-	+	-	+	+	-	-	-
28	+	+	-	+	+	-	-	-
29	-	-	+	+	+	-	-	-
30	+	-	+	+	+	-	-	-
31	-	+	+	+	+	-	-	-
32	+	+	+	+	+	-	-	-
33	-	-	-	-	-	+	-	-
34	+	-	-	-	-	+	-	-
35	-	+	-	-	-	+	-	-
36	+	+	-	-	-	+	-	-
37	-	-	+	-	-	+	-	-
38	+	-	+	-	-	+	-	-
39	-	+	+	-	-	+	-	-
40	+	+	+	-	-	+	-	-
41	-	-	-	+	-	+	-	-
42	+	-	-	+	-	+	-	-
43	-	+	-	+	-	+	-	-
44	+	+	-	+	-	+	-	-
45	-	-	+	+	-	+	-	-
46	+	-	+	+	-	+	-	-
47	-	+	+	+	-	+	-	-
48	+	+	+	+	-	+	-	-
49	-	-	-	-	+	+	-	-
50	+	-	-	-	+	+	-	-
51	-	+	-	-	+	+	-	-
52	+	+	-	-	+	+	-	-
53	-	-	+	-	+	+	-	-
54	+	-	+	-	+	+	-	-
55	-	+	+	-	+	+	-	-

56	+	+	+	-	+	+	-	-
57	-	-	-	+	+	+	-	-
58	+	-	-	+	+	+	-	-
59	-	+	-	+	+	+	-	-
60	+	+	-	+	+	+	-	-
61	-	-	+	+	+	+	-	-
62	+	-	+	+	+	+	-	-
63	-	+	+	+	+	+	-	-
64	+	+	+	+	+	+	-	-
65	-	-	-	-	-	-	+	-
66	+	-	-	-	-	-	+	-
67	-	+	-	-	-	-	+	-
68	+	+	-	-	-	-	+	-
69	-	-	+	-	-	-	+	-
70	+	-	+	-	-	-	+	-
71	-	+	+	-	-	-	+	-
72	+	+	+	-	-	-	+	-
73	-	-	-	+	-	-	+	-
74	+	-	-	+	-	-	+	-
75	-	+	-	+	-	-	+	-
76	+	+	-	+	-	-	+	-
77	-	-	+	+	-	-	+	-
78	+	-	+	+	-	-	+	-
79	-	+	+	+	-	-	+	-
80	+	+	+	+	-	-	+	-
81	-	-	-	-	+	-	+	-
82	+	-	-	-	+	-	+	-
83	-	+	-	-	+	-	+	-
84	+	+	-	-	+	-	+	-
85	-	-	+	-	+	-	+	-
86	+	-	+	-	+	-	+	-
87	-	+	+	-	+	-	+	-
88	+	+	+	-	+	-	+	-
89	-	-	-	+	+	-	+	-
90	+	-	-	+	+	-	+	-
91	-	+	-	+	+	-	+	-
92	+	+	-	+	+	-	+	-
93	-	-	+	+	+	-	+	-
94	+	-	+	+	+	-	+	-
95	-	+	+	+	+	-	+	-
96	+	+	+	+	+	-	+	-
97	-	-	-	-	-	+	+	-
98	+	-	-	-	-	+	+	-
99	-	+	-	-	-	+	+	-
100	+	+	-	-	-	+	+	-
101	-	-	+	-	-	+	+	-
102	+	-	+	-	-	+	+	-
103	-	+	+	-	-	+	+	-
104	+	+	+	-	-	+	+	-
105	-	-	-	+	-	+	+	-
106	+	-	-	+	-	+	+	-
107	-	+	-	+	-	+	+	-
108	+	+	-	+	-	+	+	-
109	-	-	+	+	-	+	+	-
110	+	-	+	+	-	+	+	-
111	-	+	+	+	-	+	+	-
112	+	+	+	+	-	+	+	-
113	-	-	-	-	+	+	+	-
114	+	-	-	-	+	+	+	-
115	-	+	-	-	+	+	+	-
116	+	+	-	-	+	+	+	-
117	-	-	+	-	+	+	+	-
118	+	-	+	-	+	+	+	-
119	-	+	+	-	+	+	+	-
120	+	+	+	-	+	+	+	-
121	-	-	-	+	+	+	+	-
122	+	-	-	+	+	+	+	-
123	-	+	-	+	+	+	+	-

124	+	+	-	+	+	+	+	-
125	-	-	+	+	+	+	+	-
126	+	-	+	+	+	+	+	-
127	-	+	+	+	+	+	+	-
128	+	+	+	+	+	+	+	-
129	-	-	-	-	-	-	-	+
130	+	-	-	-	-	-	-	+
131	-	+	-	-	-	-	-	+
132	+	+	-	-	-	-	-	+
133	-	-	+	-	-	-	-	+
134	+	-	+	-	-	-	-	+
135	-	+	+	-	-	-	-	+
136	+	+	+	-	-	-	-	+
137	-	-	-	+	-	-	-	+
138	+	-	-	+	-	-	-	+
139	-	+	-	+	-	-	-	+
140	+	+	-	+	-	-	-	+
141	-	-	+	+	-	-	-	+
142	+	-	+	+	-	-	-	+
143	-	+	+	+	-	-	-	+
144	+	+	+	+	-	-	-	+
145	-	-	-	-	+	-	-	+
146	+	-	-	-	+	-	-	+
147	-	+	-	-	+	-	-	+
148	+	+	-	-	+	-	-	+
149	-	-	+	-	+	-	-	+
150	+	-	+	-	+	-	-	+
151	-	+	+	-	+	-	-	+
152	+	+	+	-	+	-	-	+
153	-	-	-	+	+	-	-	+
154	+	-	-	+	+	-	-	+
155	-	+	-	+	+	-	-	+
156	+	+	-	+	+	-	-	+
157	-	-	+	+	+	-	-	+
158	+	-	+	+	+	-	-	+
159	-	+	+	+	+	-	-	+
160	+	+	+	+	+	-	-	+
161	-	-	-	-	-	+	-	+
162	+	-	-	-	-	+	-	+
163	-	+	-	-	-	+	-	+
164	+	+	-	-	-	+	-	+
165	-	-	+	-	-	+	-	+
166	+	-	+	-	-	+	-	+
167	-	+	+	-	-	+	-	+
168	+	+	+	-	-	+	-	+
169	-	-	-	+	-	+	-	+
170	+	-	-	+	-	+	-	+
171	-	+	-	+	-	+	-	+
172	+	+	-	+	-	+	-	+
173	-	-	+	+	-	+	-	+
174	+	-	+	+	-	+	-	+
175	-	+	+	+	-	+	-	+
176	+	+	+	+	-	+	-	+
177	-	-	-	-	+	+	-	+
178	+	-	-	-	+	+	-	+
179	-	+	-	-	+	+	-	+
180	+	+	-	-	+	+	-	+
181	-	-	+	-	+	+	-	+
182	+	-	+	-	+	+	-	+
183	-	+	+	-	+	+	-	+
184	+	+	+	-	+	+	-	+
185	-	-	-	+	+	+	-	+
186	+	-	-	+	+	+	-	+
187	-	+	-	+	+	+	-	+
188	+	+	-	+	+	+	-	+
189	-	-	+	+	+	+	-	+
190	+	-	+	+	+	+	-	+
191	-	+	+	+	+	+	-	+

192	+	+	+	+	+	+	-	+
193	-	-	-	-	-	-	+	+
194	+	-	-	-	-	-	+	+
195	-	+	-	-	-	-	+	+
196	+	+	-	-	-	-	+	+
197	-	-	+	-	-	-	+	+
198	+	-	+	-	-	-	+	+
199	-	+	+	-	-	-	+	+
200	+	+	+	-	-	-	+	+
201	-	-	-	+	-	-	+	+
202	+	-	-	+	-	-	+	+
203	-	+	-	+	-	-	+	+
204	+	+	-	+	-	-	+	+
205	-	-	+	+	-	-	+	+
206	+	-	+	+	-	-	+	+
207	-	+	+	+	-	-	+	+
208	+	+	+	+	-	-	+	+
209	-	-	-	-	+	-	+	+
210	+	-	-	-	+	-	+	+
211	-	+	-	-	+	-	+	+
212	+	+	-	-	+	-	+	+
213	-	-	+	-	+	-	+	+
214	+	-	+	-	+	-	+	+
215	-	+	+	-	+	-	+	+
216	+	+	+	-	+	-	+	+
217	-	-	-	+	+	-	+	+
218	+	-	-	+	+	-	+	+
219	-	+	-	+	+	-	+	+
220	+	+	-	+	+	-	+	+
221	-	-	+	+	+	-	+	+
222	+	-	+	+	+	-	+	+
223	-	+	+	+	+	-	+	+
224	+	+	+	+	+	-	+	+
225	-	-	-	-	-	+	+	+
226	+	-	-	-	-	+	+	+
227	-	+	-	-	-	+	+	+
228	+	+	-	-	-	+	+	+
229	-	-	+	-	-	+	+	+
230	+	-	+	-	-	+	+	+
231	-	+	+	-	-	+	+	+
232	+	+	+	-	-	+	+	+
233	-	-	-	+	-	+	+	+
234	+	-	-	+	-	+	+	+
235	-	+	-	+	-	+	+	+
236	+	+	-	+	-	+	+	+
237	-	-	+	+	-	+	+	+
238	+	-	+	+	-	+	+	+
239	-	+	+	+	-	+	+	+
240	+	+	+	+	-	+	+	+
241	-	-	-	-	+	+	+	+
242	+	-	-	-	+	+	+	+
243	-	+	-	-	+	+	+	+
244	+	+	-	-	+	+	+	+
245	-	-	+	-	+	+	+	+
246	+	-	+	-	+	+	+	+
247	-	+	+	-	+	+	+	+
248	+	+	+	-	+	+	+	+
249	-	-	-	+	+	+	+	+
250	+	-	-	+	+	+	+	+
251	-	+	-	+	+	+	+	+
252	+	+	-	+	+	+	+	+
253	-	-	+	+	+	+	+	+
254	+	-	+	+	+	+	+	+
255	-	+	+	+	+	+	+	+
256	+	+	+	+	+	+	+	+

-- The end --