umcg university of groningen faculty of mathematics and natural sciences life sciences

Assessment of the benefits and cost-effectiveness of implementing systematic breast cancer screening in the urban Chinese population.

Chris de Jonge (s1634259)

Under supervision of:
Dr. M.J.W. Greuter [1]
Prof. G.H. de Bock [2]

14th June 2015

Abstract

Currently, there is no systematic breast cancer screening implemented for the urban Chinese population. However, as the incidence of breast cancer is rising, implementing screening might be a cost efficient way of reducing breast cancer mortality. In this study, the SiMRiSc breast cancer screening model has been adapted for the Chinese urban population, with the aim to assess the benefits and cost-effectiveness of implementing systematic breast cancer screening in the urban Chinese population. Literature research was conducted to deduce relevant model input parameters for the Chinese ethnicity. Also, the SiMRiSc model was modified and improved, and new tumor growth and survival models have been implemented. Then, the model was validated internally and externally. The results of different screening scenarios that have been simulated include the number of cancers detected and the incremental costs per life-year saved. These simulations suggest that the mortality can be reduced by implementing screening, as tumors will be detected at a significantly smaller size than without screening, giving an improved survival chance for women. Starting early with screening, at age 30, seems to be cost-effective for the urban Chinese population, resulting in a cost of €2781 per life year gained. The robustness and uncertainty of the simulations have been analyzed by a sensitivity analysis, showing that the tumor volume doubling time is the greatest contributor to the uncertainty in the results, with the self-detection diameter second. Running simulations with 100000 women is deemed the minimum population size to keep statistical variance limited. Because SiMRiSc does not add pseudo-disease to the population, it is doubtful that overdiagnosis can be simulated with SiMRiSc. A suggestion has been given to implement stagnating tumors to the tumor growth model to allow for the simulation of overdiagnosed tumors.

[1] University Medical Center Groningen, Department of Radiology, Groningen , The Netherlands
[2] University Medical Center Groningen, Department of Epidemiology, Groningen , The Netherlands

Preface

The work presented in this master's thesis is the result of a research at the department of Epidemiology of the University Medical Center Groningen, and is part of the Biomedical Engineering curriculum at the University of Groningen. This thesis is primarily aimed at my supervisors, but it is also a thorough description of the SiMRiSc breast cancer screening simulation model and the changes I've made to the software, so can be considered as a manual and guideline for future users of SiMRiSc.

Many changes were made to the SiMRiSc code, up to the point where it can be called 'SiMRiSc 1.5'. While the underlying idea of the simulation has not changed, the code was completely overhauled including many bug- and performance-fixes. Still I am still far from satisfied with the quality of the code, and many parts of the program are still 'spaghetti code'. My recommendation is that this is to be re-written by somebody with a stronger programming background to make the SiMRiSc codebase maintainable, less prone to bugs, more user-friendly and useable for a long time in the future.

The newly implemented tumor growth model, including the introduction of a 'self-detected tumor distribution' and a new tumor survival model consist of novel work, partly based on the work of James S. Michaelson, Ph.D. at the university Department of Pathology, Massachusetts General Hospital. Especially these parts I am very proud of. These parts might also have further applications in cancer epidemiology as the methods described might not only be applicable to breast cancer.

I am grateful to prof. de Bock of the department of Epidemiology of the University Medical Center Groningen for providing a position at the UMCG to perform my master project. Also I would like to thank my primary supervisor, Dr. M. Greuter, for his supervision and positive feedback.

Groningen,

June 14, 2015

Table of contents

1. General Introduction 3
1.1. Motivation 3
1.2. Research aim 3
1.3. The SiMRiSc model 3
1.4. Ethics 4
1.5. Outline 5
2. Implementing SiMRiSc for the urban Chinese population 6
2.1. Introduction 6
2.2. Materials and Methods 6
2.3. Results 6
2.4. Tumor growth model 13
2.5. Tumor survival 20
2.6. Discussion 26
2.7. Conclusion 26
3. External validation 27
3.1. Introduction 27
3.2. Material and methods 27
3.3. Results 27
3.4. Discussion 30
3.5. Conclusion 31
4. Scenario simulations 32
4.1. Introduction 32
4.2. Material and methods 32
4.3. Results 32
4.4. Discussion 33
4.5. Conclusion 33
5. Sensitivity analysis 34
5.1. Introduction 34
5.2. Material and methods 34
5.3. Results 35
5.4. Discussion 37
5.5. Conclusion 38
6. Overdiagnosis 39
6.1. Introduction 39

6.2. Material and methods 39
6.3. Results 39
6.4. Discussion 39
6.5. Conclusion 40
7. General Discussion 41
7.1. Further work 41
7.2. Conclusions 41
8. Bibliography 43

1. General Introduction

1.1. Motivation

Systematic mammographic breast cancer screening is implemented in most Western countries as a measure to reduce mortality. Diagnosing a tumor early and at a smaller size than when self-detected increases the survival chances for the women. However, the cost-effectiveness of systematic screening and the optimal screening strategy depends on factors like breast-cancer incidence and breast density.

While the incidence of breast-cancer among urban Chinese women is lower than among western women, the incidence is rising, due to a change in lifestyle, diet, birth control and other factors. [1] [2] As breast cancer cases diagnosed in Hong Kong have tripled from 1993 to 2011, similar trends can be expected for urban Chinese women. [3] Figure 1 shows the number of new cases of breast cancer detected in Hong Kong from 2003 until 2012. The majority of the new cases are in the age-group 45-64. The same can be expected for urban Chinese cities.

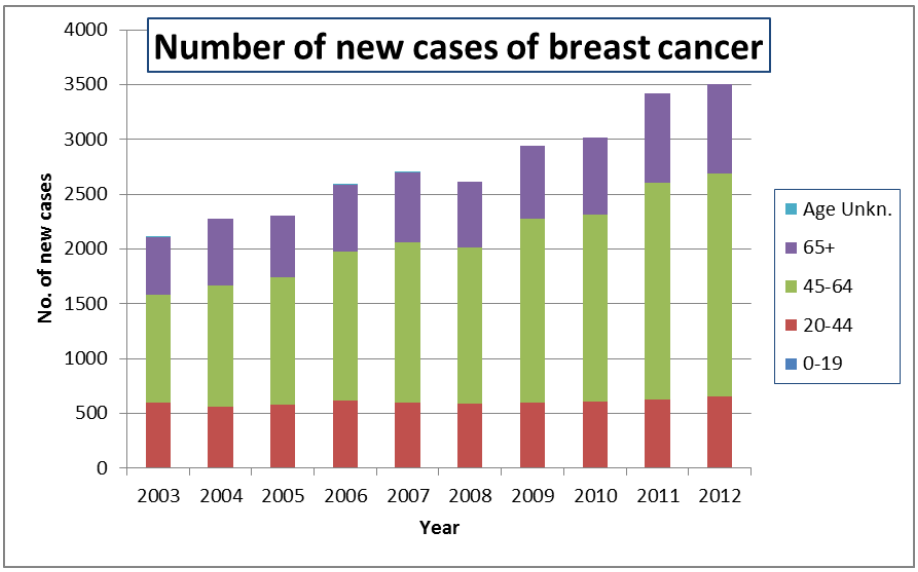


Figure 1. No. of new cases of breast cancer in Hong Kong, divided in age groups. In 2012 3508 new cases of breast cancer have been registered. [4]

1.2. Research aim

The goal of this study is to assess the benefits and cost-efficiency of implementing systematic mammographic breast cancer screening in the urban Chinese population by using the SiMRiSc breast cancer screening model.

1.3. The SiMRiSc model

The SiMRiSc breast cancer screening model is a micro simulation written in C++. The SiMRiSc model has previously been used in studies for the Western population. [5] [6] [7] [8]

In the simulation, women are simulated through their lifetime. Each woman in the simulation population is 'born' with a random chance on breast cancer, according to the breast cancer incidence curves. Also, life expectancy and breast density are determined randomly for each woman. If it is determined that a certain women will get a tumor during her lifetime, the tumor

volume doubling time and self-detection age distributions are also sampled to determine these parameters for the women.

Then, different breast-cancer screening scenarios can be simulated, taking into account the mammographic sensitivity, exposure to (diagnostic) radiation, tumor growth and other parameters. A woman will be removed from the simulation if she dies of natural causes, if a tumor grows so large that it is self-detected, or if a tumor is detected by screening. The output of the model is given in number of cancers detected with the screening, number of interval cancers and cost of the screening and treatment (Figure 2).

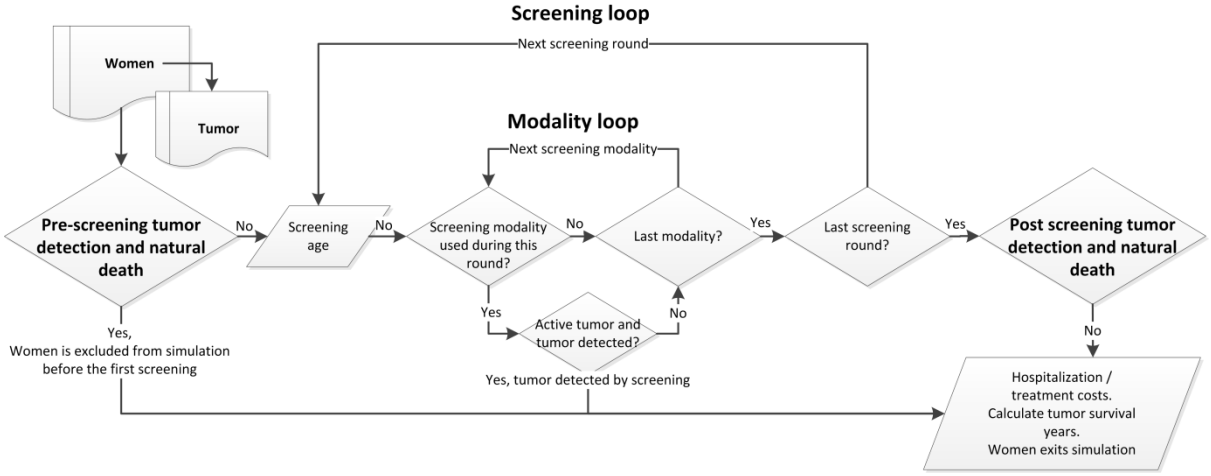


Figure 2. Simplified flow chart of the SiMRiSc model. A more detailed version is available in Appendix D.

1.4. Ethics

This research is original, unpublished, independent work by the author of this thesis. It is based on the SiMRiSc simulation software [5] [6] developed by the department of epidemiology of the University Medical Center Groningen. In this project, the SiMRiSc model was modified, improved, used with Chinese input parameters derived from literature and validated internally and externally. This project was done in collaboration with our research partner in Tianjin, China, which provided us with mammography and treatment costs data.

1.4.1. Societal impact

In Asian countries, breast cancers screening is only partly implemented and only a few studies are known on the benefit and risks of breast cancer screening in Asian women. As the breast cancer incidence is rising rapidly in the Chinese urban areas, performing systematic screening might become cost-effective for the urban population. Performing screening simulations using the SiMRiSc software can help determine whether mammographic breast cancer screening is cost effective and what are the optimal screening parameters like the start age, end age and screening interval. This can help to reduce the mortality and lead to a fair and more efficient allocation of limited healthcare budget.

1.4.2. Benefits and risks of mammographic screening

The main benefit of mammographic screening is that mammographic screening can detect breast cancers before a lump can be felt by the women. This will lead to smaller diagnosed tumors and decreased mortality of breast cancer.

However, mammographic screening can occasionally miss the detection of a tumor. This happens in approximately 13 to 35% of the cases (see paragraph 2.3.4), depending on the

breast density of the women. The cause of this is that the cancer might be obscured by dense tissue, or is in an area not easily imaged by mammography. The sensitivity of mammography is taken into account by the simulation software.

Ionizing radiation from mammography will induce new tumors, and will offset the benefits of screening. This effect is taken into account for in the simulations.

In this project, only the benefits in terms of life years saved are computed. However, even if a woman would survive both with and without screening, screening might still be able to detect a tumor when it is small enough to allow breast conserving surgery, while without screening this would not be the case. These effects are not simulated with the SiMRiSc software and the benefits of this are not visible in the analysis.

On the other hand, overdiagnosis is a side effect of screening; screening might detect cancers that would otherwise have had no influence on the women if it was undetected. This might be the case if a woman with a screen-detected tumor has a very slow growing tumor and dies of other causes before the symptoms of the tumor would present themselves. This turns the particular women into a life-long cancer patient where this would not be the case without the mammographic screening and also leads to increased healthcare costs. The health-care cost of overdiagnoses is taken into account for, but the psychological effects for the women are not.

A limited specificity of mammography will lead to false positives. The amount of false positives and the associated healthcare costs are computed by the simulation, however the psychological effect that a false positive has on women is not taken into account for when doing a cost effectiveness analysis.

1.5. Outline

The outline of this thesis is as follows. First, Chapter 2 discusses the implementation of the SiMRiSc model for the urban Chinese population, including the deduction of the SiMRiSc input parameters from literature and the implementation of new tumor growth and survival models in SiMRiSc. Then, in Chapter 3, the new model and input parameters are validated externally to hospital data from Hong Kong. In Chapter 4, different screening scenarios are simulated and assessed on cost effectiveness. The SiMRiSc model and the results are further analysed by the sensitivity analysis described in Chapter 5. Chapter 6 discusses overdiagnosis of tumors with screening. Chapter 7 presents further work, possible future improvements of SiMRiSc and the conclusions.

2. Implementing SiMRiSc for the urban Chinese population

2.1. Introduction

In this chapter, the adaptation and implementation of the SiMRiSc model for the urban Chinese population is presented.

2.2. Materials and Methods

The adaptation of SiMRiSc for the urban Chinese population is done by finding and inserting appropriate input parameters from literature. These values are mostly ethnicity specific, so are different from those that have been used for previous studies with SiMRiSc of the Western population. In addition, an error estimate for the parameters has to be calculated, as this is required for the sensitivity analysis (Chapter 5).

The literature study was performed by looking up relevant data from China. Literature was hard to find, and for most parameters only one or two relevant sources could be found. Most detailed and best documented information was found for Hong Kong, so has been used as a source. If literature could not be found for China, other Southeast Asian sources were used, like from Japan.

Then, to confirm that the model is working as expected, each of the input parameters is validated internally by comparing the output of SiMRiSc to the input data. Because this was not the case for the 'tumor growth model', this part of SiMRiSc has been completely re-written to represent reality. The new version of the tumor growth model is described in paragraph 2.4 and makes use of 5 new input parameters: Lower threshold of mammography (diameter), tumor volume doubling time (mean and standard deviation) and a self-detection diameter (mean and standard deviation). Also, a new tumor survival model has been implemented into SiMRiSc, as is described in paragraph 2.5.

2.3. Results

Table 1 lists the parameters and their corresponding values that have been used for the implementation of the SiMRiSc model the urban Chinese population. These parameters and the deduction for the Southeast Asian ethnicity and/or literature source(s) of their values will be described in subsequent paragraphs of this chapter.

Table 1. Parameters and values of the SiMRiSc model, with 1σ confidence intervals.

Parameter			Value			
Life expectancy			See Figure 3			
Breast cancer incidence rate	Lifetime risk		8.35% ± 0.53%			
	Mean age		57.73 years ± 0.552			
	Spread		17.11 year ± 0.048			
Tumor induction model	Dose per mammogram		3 mSv			
	Probability of tumor induction		0.51 ± 0.32			
Tumor growth model	Limit of clinical detection (diameter)		5 mm			
	Tumor volume doubling time	μ_{DT}	5.159 ± 0.102			
		σ_{DT}	0.98			
	Self-detection diameter	μ_{SD}	2.92 ± 0.043			
		σ_{SD}	0.66			
Tumor survival model	a		0.00004475 ± 0.000004392			
	b		1.85867 ± 0.0420			
	c		-0.271 ± 0.0101			
	d		2.0167 ± 0.0366			
Distribution of breast densities	BI-RADS density score		1	2	3	4
	Age < 50		11.80%	17.70%	51.00%	19.60%
	Age > 50		24.30%	31.90%	36.60%	7.10%
Mammographic sensitivity			87%	84%	73%	65%

The cost parameters used by the SiMRiSc model have been provided by the research partner in Tianjin, China and are listed in Table 2.

Table 2. The cost parameters used by the SiMRiSc model, assuming 0.13 Euro per Yuan.

Parameter		Costs in Yuan	Costs in €
Mammogram		200	€26
Treatment of tumor	<20mm	41063	€5356
	20-50mm	52427	€6838
	>50mm	57735	€7531
Biopsy costs		1200	€157

2.3.1. Life expectancy

Hong Kong has one of the highest female life expectancies of the world. The life expectancy of the female population is used to determine the natural death age of every women in the simulation population. Age and sex specific mortality rates have been published by the census and statistics department of Hong Kong. [9] Over 42000 deaths were reported in 2011 and used to generate this data, giving a very small statistical error over this data.

The cumulative age-specific mortality rate is shown in Figure 3, and compared to the data that has been used for SiMRiSc simulations on Dutch population. [10] It is to be noted that the women in Hong-Kong have a much higher life expectancy.

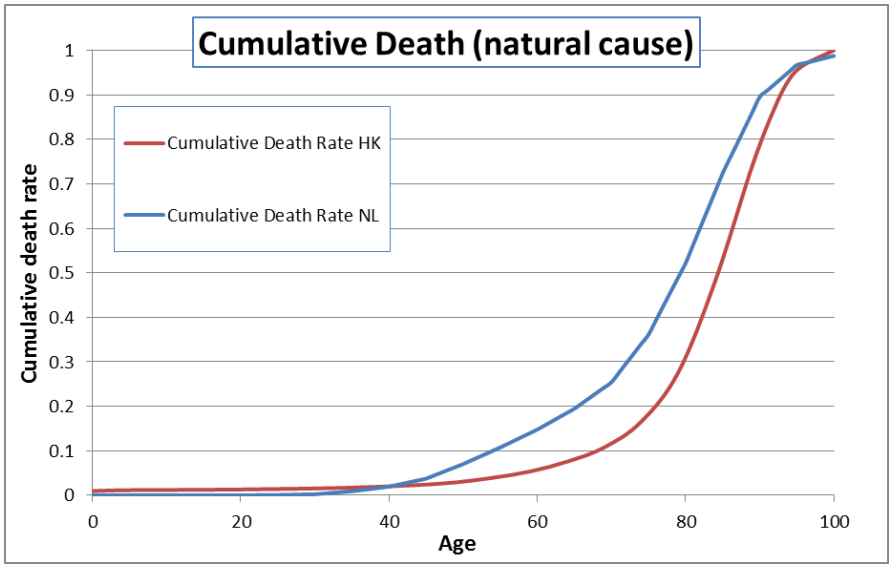


Figure 3. The cumulative death rate of females in Hong Kong, as published by the Hong Kong Department Census and Statistics, compared to data used for Dutch simulations.

2.3.2. Breast cancer incidence rate

The breast cancer incidence is used in SiMRiSc to 'assign' tumors to women from the population at a certain age, according to the breast cancer incidence probability. The breast cancer incidence rate that is implemented in SiMRiSc assumes a normal distributed breast cancer incidence as function of age. The probability density function is a Gaussian function normalized to the lifetime risk of developing breast cancer and has the form of:

$$p(a) = \frac{\mathrm{f}}{\sigma\sqrt{2\pi}} \mathrm{e}^{\frac{-1}{2}(\frac{\mathrm{a}-\mu}{\sigma})^2} \tag{1}$$

with the parameters shown in Table 3.

Table 3. Breast cancer risk parameters.

Parameter	Description
p	Risk of acquiring breast cancer during that year
a	Women age
f	Lifetime risk
σ	Spread
μ	Mean

The cumulative risk describes the total risk the women has to acquire breast cancer up until a certain age, and is the integral of the probability density function:

$$I_{cumulative}(a) = \int_0^a \frac{f}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}(\frac{a-\mu}{\sigma})^2} da \tag{2}$$

Calculating the indefinite integral of this yields the lifetime risk.

The values of f (total lifetime risk), σ (the spread) and μ (the mean) have been deduced from the breast cancer incidence rate for Hong Kong published by the Hong Kong Cancer Registry of the Hong Kong Hospital Authority. [4] Table 4 shows the calculated cumulative risk from the published data.

Table 4. Breast cancer risk (only invasive tumors) in Hong Kong in 2012 over 3508 registered cases. [3] At age 70 the risk is 1 in 16.1. The total lifetime risk is 1 in 12, or about 8.4%.

Age	Cumulative risks (one in)	Cumulative risk [%] and 1σ error estimate
≤ 30	772.2	0.130 ± 0.022
≤ 35	238.4	0.420 ± 0.043
≤ 40	102.2	0.978 ± 0.072
≤ 45	55.2	1.813 ± 0.11
≤ 50	37.6	2.658 ± 0.14
≤ 55	28.3	3.534 ± 0.18
≤ 60	22.3	4.492 ± 0.23
≤ 65	18.4	5.441 ± 0.29
≤ 70	16.1	6.221 ± 0.35
lifetime	12	8.355 ± 0.53

The error interval has been calculated using counting statistics, assuming a count rate consistent with a Poisson distribution, using the number of cases detected in that age group. Figure 4 shows this data in graphical form, and it can be seen that it follows roughly the expected 'S-shape' of an integrated Normal curve.

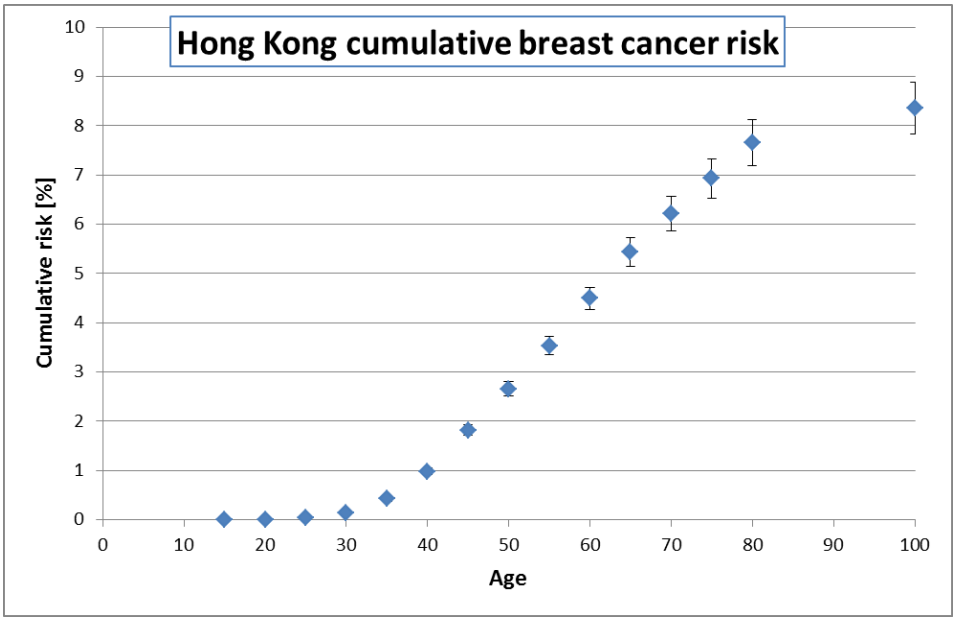


Figure 4. The cumulative breast cancer risk in Hong Kong, with error bars (1 σ).

The differential of the cumulative incidence described the incidence rate, and this data roughly follows the expected normal distribution. This principle was used to curve-fit a Normal curve to the data, to determine the required parameters of Table 3 to be used in SiMRiSc.

The peak of the incidence is the mean of the normal curve and is at 57.7 years. With the mean and total lifetime risk known, only one parameter (the standard deviation) of the normal distribution is unknown, and is found by making a fit. Curves were also fitted to points of the upper and lower confidence limits, yielding the upper and lower boundaries for the Normal-curve parameters. The resulting parameters are listed in Table 5 and compared to the parameters that have been found for the Netherlands by Zahn. [8] It can be concluded

that the lifetime risk on breast cancer is much lower for Hong Kong than for the Netherlands, and on average women in Hong Kong get breast cancer at an earlier age.

Table 5. Parameters for the Hong Kong breast cancer incidence, with 1σ confidence interval and compared to the parameters of the Netherlands for reference.

	HK	NL
f (lifetime risk)	0.0835 ±0.0053	0.21
σ (the standard deviation)	17.11 ±0.048	20
μ (the mean)	57.73 ±0.552	69.5

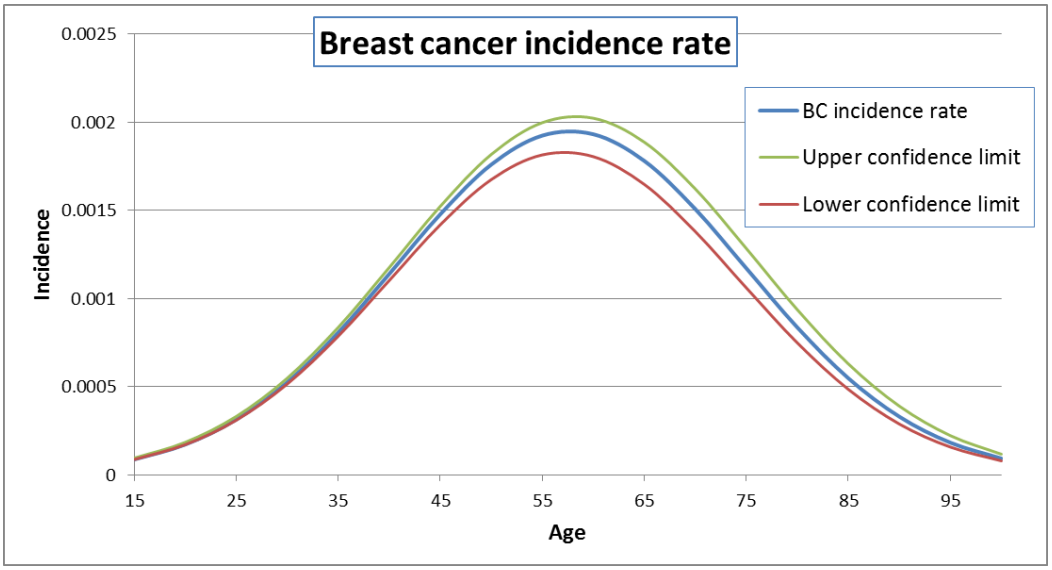


Figure 5. Normal curves of the Hong Kong breast cancer incidence rate with 1σ upper and lower confidence limits.

By taking the integral of the normal curves with the found parameters, the fitted data can be compared to the original data from Figure 4. At later ages the Gaussian fit is good, but at earlier ages there is a deviation due to the 'side-lobe' of the Gaussian function. This will lead to a small increased breast cancer incidence at earlier ages in the simulation.

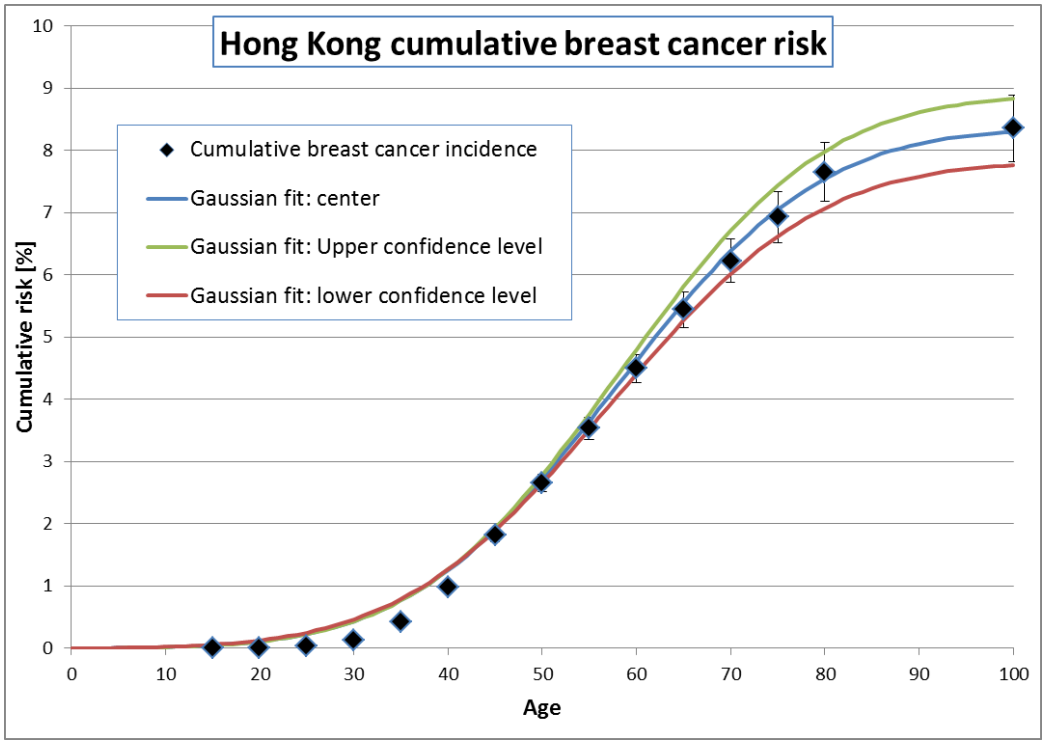


Figure 6. Hong Kong cumulative breast cancer risk; fitted data compared to the input data.

2.3.3. Radiation risk

At low doses, statistical limitations make it difficult to evaluate cancer risk. The BEIR VII committee has developed risk estimates for exposure to low-dose, low-LET radiation to humans (mammography screening falls inside this regime), and concluded that a linear no-threshold model is valid (Figure 7). Even a small amount of radiation can induce cancer, thus tumor induction due to mammographic radiation dose has to be accounted for in SiMRiSc. The tumor induction parameters are listed in Table 6 and these parameters have been taken from previous simulations [5] as no changes are expected for different ethnicities.

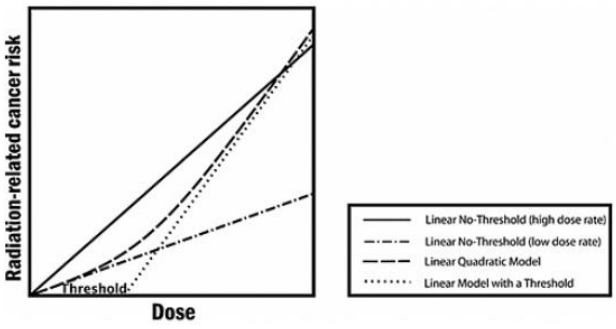


Figure 7. Different models for tumor induction. According to the BEIR VII report [11], the linear no-threshold model is valid.

Table 6. The tumor induction parameters used in the SiMRiSc model. [5]

Dose per mammogram	3 mSv	± 2
Probability of tumor induction	0.51	± 0.32

When screening annually from age 30 to age 70 (41 doses of 3mSv), the tumor risk increases from 8.3% to 9.0% (Figure 8).

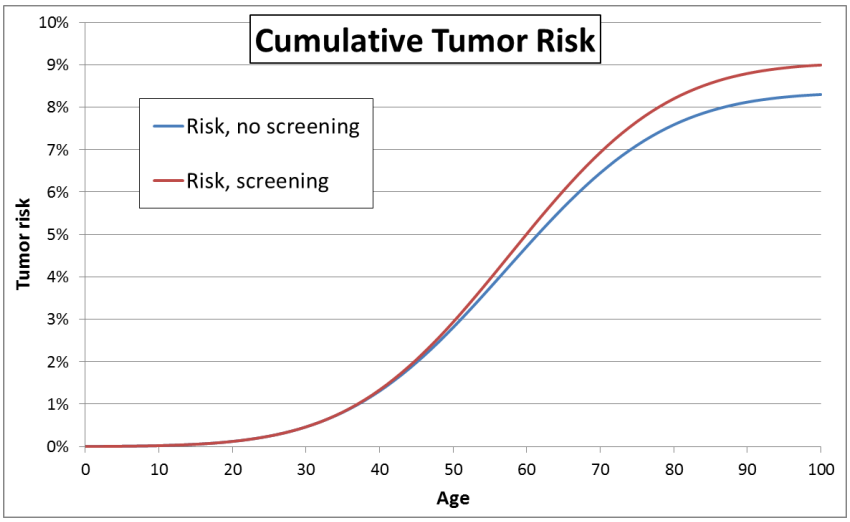


Figure 8. The cumulative tumor risk, with and without screening.

2.3.4. Breast density, sensitivity and limit of mammographic detection

The mammographic sensitivity determines the fraction of false negatives. Over 10% of the tumors are missed by the mammographic screening.

Mammographic sensitivity decreases with increasing breast density; dense tissue can obscure the tumor. During the life of a woman, breast will become less dense. Sensitivity and age are thus related; younger women have a lower sensitivity. Breasts of Asian women are denser than western women, making the sensitivity also lower. [12]

Table 7. Mammographic sensitivity for Asian women according to age and breast-density. [12]

Distribution of breast densities	BI-RADS density	1	2	3	4
	Age <50	11.80%	17.70%	51.00%	19.60%
	Age >50	24.30%	31.90%	36.60%	7.10%
Mammography sensitivity		87%	84%	73%	65%

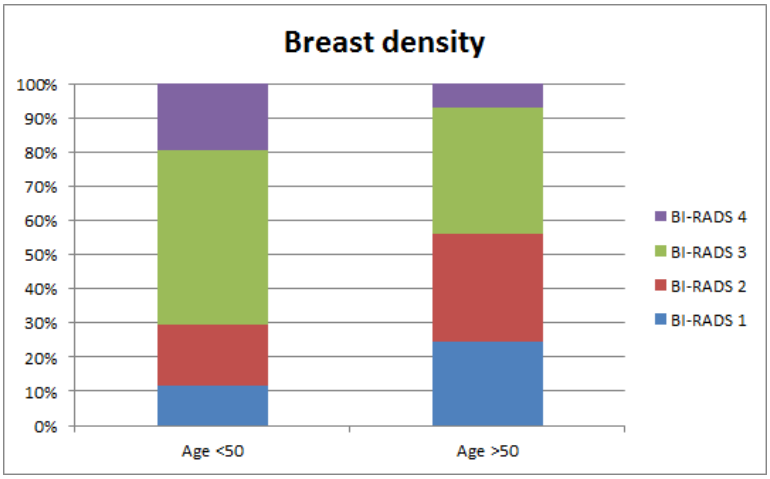


Figure 9. Breast density distribution of Asian women according to age and breast density.

The median size at which breast tumors become detectable by mammography has been reported at 7mm by Michaelson et. al. [13] This limit is influenced by physical and human factors, and thus may vary from hospital to hospital. Because the mammographic sensitivity also depends on breast density; mammography is less capable of detecting invasive breast cancers in women with dense breasts. Therefore, mammography may be capable of detecting invasive breast cancers at slightly smaller sizes in older women than in younger women.

The spread on the clinical detection limit is reported asymmetric around the mean: while 7mm tumors are found approximately 50% of the time, 5mm tumors are found approximately 35% of the time, 10mm tumors are found approximately 65% of the time and 15-mm tumors are found approximately 80% of the time. Tumors larger than 30mm appear never to be missed (Table 8).

Table 8. Sensitivity according to tumor diameter.

Tumor diameter	Found approximately
5mm	35%
7mm	50%
10mm	65%
15mm	80%
30mm	100%

In the current implementation of SiMRiSc, the limit of mammographic detection and the sensitivity are implemented as separate fixed parameters. 5mm (slightly below the median reported by Michaelson et al.) has been used as the limit of mammographic detection. In reality however, the sensitivity is related to tumor size and breast density. A smaller tumor simply has a lower sensitivity associated to it. While the sensitivity will approach 0 rapidly for tumors <5mm diameter, there is no 'hard' limit of mammographic detection. It is possible for a future improvement of SiMRiSc this is implemented according to the work of Michaelson et al. [13], and sensitivity becomes one continuous function of tumor size and breast density, removing the need for the concept of a 'hard' limit of mammographic detection.

2.4. Tumor growth model

A new tumor growth model has been implemented in SiMRiSc that is based on exponential tumor growth. A tumor volume doubling time (TVDT) and self-detected tumor size are sampled for each woman randomly from log-normal distributions. From the TVDT and self-detected tumor size, the detectable preclinical phase (the timeframe in which the tumor is potentially detectable with mammography) for each women is calculated.

2.4.1. Doubling time and detectable preclinical phase

The tumor growth model implemented in SiMRiSc is based on the exponential tumor growth model (Figure 11) proposed by Collins et al. [14] Assuming spherical tumors, the tumor volume V_1 at time t_1 starting at volume V_0 at time t_0 can be calculated as:

$$V_1 = V_0 \cdot 2^{(t_1 - t_0)/t_D} \quad (3)$$

The tumor volume doubling time (t_D) is expressed by [15]:

$$t_D = \frac{(t_1 - t_0) \cdot ln(2)}{ln(V_1/V_0)} \tag{4}$$

where V_0 and V_1 are the tumor volumes at times t_0 and t_1 respectively. The time it takes for a tumor to grow from V_0 to V_1 can be calculated by solving equation 3 for $(t_1 - t_0)$:

$$(t_1 - t_0) = \log_2(\frac{V_1}{V_0}) \cdot t_D = \frac{ln(V_1/V_0)}{ln(2)} \cdot t_D \tag{5}$$

The 'detectable preclinical phase' is defined as the time period in which the cancer is potentially detectable on the mammogram (Figure 10); starting from when the tumor is large enough to be detected by mammography (limit of clinical detection) until the time that the tumor is found by the women through 'self-detection' because the tumor has grown so large that it is detected by palpation or symptoms.

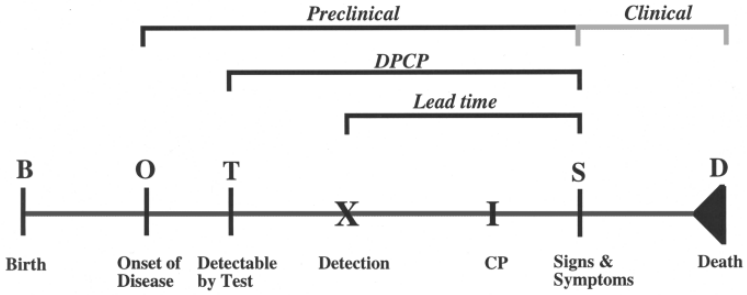


Figure 10. Timeline showing different phases of the tumor during a person's lifetime. The detectable preclinical phase (DPCP) is the time between the lower threshold of detection and the time of self-detection because of symptoms. CP is the critical point (eg. metastasis) of the tumor. Screening and detection need to have happened before this point to be effective. Figure from Obuchowski et al. [16]

Assumed is that invasive tumors with a diameter larger than 5mm may be detected on the mammogram. Then, for example for a woman that would self-detect a tumor at a diameter of 20mm, the detectable preclinical phase can be calculated by using $V_0 = 65mm^3$ (equivalent of the volume of a 5mm diameter sphere) and $V_1 = 4189mm^3$ (equivalent of the volume of a 20mm diameter sphere).

If a woman would self-detect a tumor at a diameter $D_{sd} = 20mm$ and the detectable preclinical phase t_{PC} is defined in years and the doubling time is defined in days, there is a linear relationship between the tumor volume doubling time and detectable preclinical phase with constant a:

$$t_{PC} = \frac{ln\,(4188.79/65.45)}{ln(2) \cdot 365} t_D = a \cdot t_D \approx 0.01644\, t_D \tag{6}$$

Using this relationship, the detectable preclinical phase t_{PC} for $t_D = 174\, days$ and $D_{sd} = 20mm$ is 2.86 year (Figure 11).

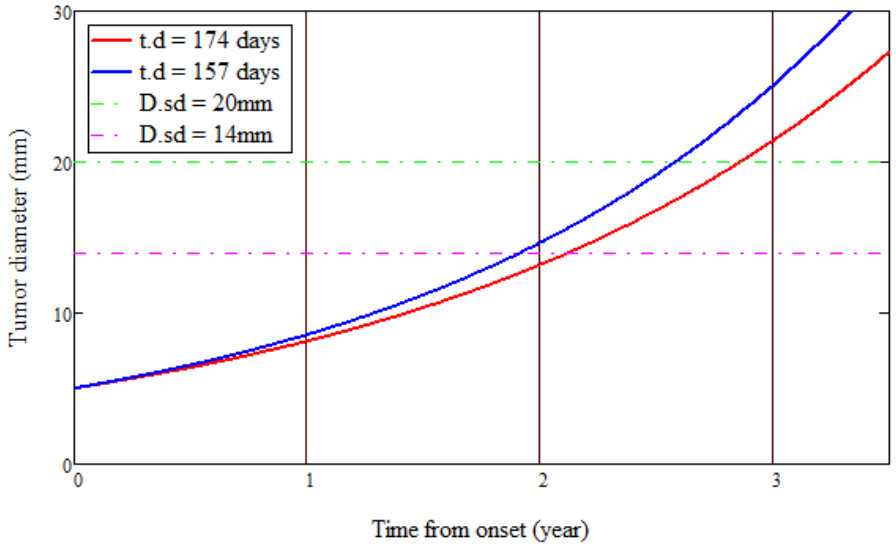


Figure 11. The exponential tumor growth model. The tumor diameter is calculated from the volume assuming spherical tumors. The detectable preclinical phase is the time it takes for a tumor to grow from 5mm diameter to the self-detect diameter. This is, for example, approximately 2.86 year for t_D = 174 days and D_{sd} = 20mm, or approximately 1.92 year for t_D = 154 days and D_{sd} = 14mm.

However, the parameters t_D and D_{sd} will be different for each women, and are sampled in SiMRiSc from distributions, as will be described in the next 3 paragraphs.

2.4.2. Doubling time distribution

Some literature studies reporting data on breast tumor volume doubling times are listed in Table 9. The data from Peer et al [17] [18] has been used in previous studies to verify the SiMRiSc model for the Dutch population. For this research, data from a Japanese study (Kuroishi et al) [19] is used as this data is most relevant for the urban Chinese ethnicity.

Table 9. A number of studies with estimates of the breast tumor volume doubling time. [20]

Reference	Tumor Volume Doubling Time
Peer et al [17]	Geometric mean 80 days (age < 50)
Peer et al [17]	Geometric mean 157 days (age 50–70)
Peer et al [17]	Geometric mean 188 days (age > 70)
Michaelson et al [20]	Median 130 days
Fournier et al [21]	Geometric mean 225 days
Lundgren [22]	Mean 211 days
Spratt et al [23]	Median 260 days
Kuroishi et al [19]	Geometric mean 174 days

According to Kuroishi et al. [19] and Greuter et al. [5] the tumor volume doubling time exhibits approximately a log-normal distribution (Figure 12).

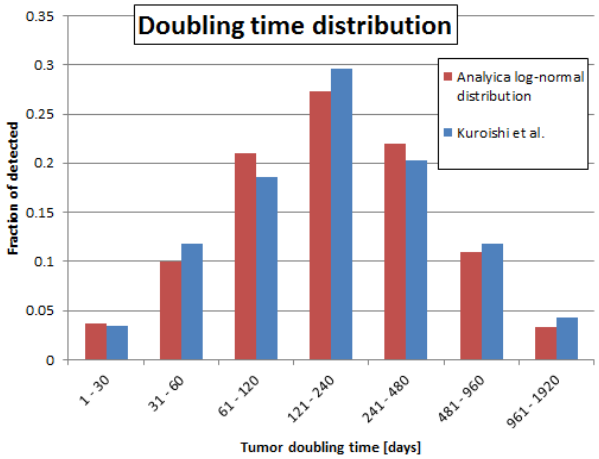


Figure 12. Comparing Kuroishi et al to the binned analytical log-normal distribution with μ_{DT} = 5.16 and σ_{DT} = 0.98.

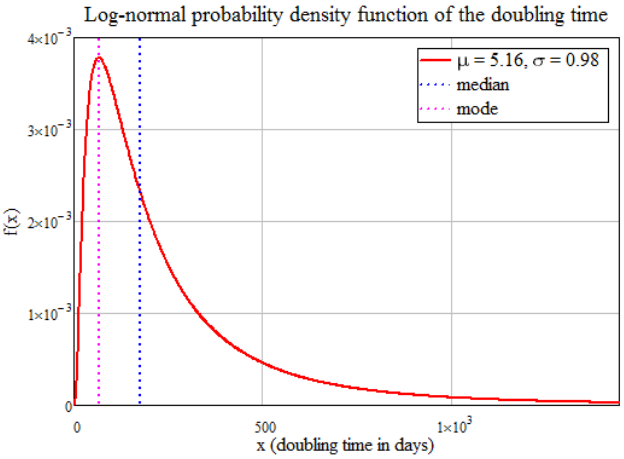


Figure 13. Log-normal probability density function f(x) of the tumor volume doubling time, with μ = 5.16 and σ = 0.98.

The probability density function of the log-normal distribution (Figure 13) is defined as:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma}\right)^2} \tag{7}$$

where μ is the (arithmetic) mean and median of $ln(x)$, and σ is the standard deviation of $ln(x)$. The relation between μ and the geometric mean m_g of the non-logarithmic data is:

$$m_g = e^{\mu} \tag{8}$$

$$\mu = ln(m_g) \tag{9}$$

According to Kuroishi et al [19], the geometrical mean of the tumor volume doubling time $m_{gDT} = 174$ days. The mean μ_{DT} of $\ln(t_D) = 5.16$ and the standard deviation σ_{DT} of $\ln(t_D) = 0.98$.

The data of the histogram provided by Kuroishi et al. can be compared to the analytical log-normal distribution by binning the probability density function (Figure 12). Binning is done by integrating the probability density function over the start a and end b of the interval:

$$P(a,b) = \int_a^b \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma}\right)^2} dx \tag{10}$$

2.4.3. Detectable preclinical phase distribution

Because of the linear relationship between the tumor volume doubling time and the detectable preclinical phase, the detectable preclinical phase will also exhibit a log-normal distribution. If the geometric mean of the tumor volume doubling time is m_{gDT}, then the geometric mean of the detectable preclinical phase m_{gPC} can be expressed as:

$$m_{gPC} = a \cdot m_{gDT} \tag{11}$$

where a is the constant between t_{PC} and t_D from equation 6, assuming a fixed self-detect diameter of 20mm (this is not the case, as will be described in the next paragraph).

Then, by substituting into equation 9:

$$\mu_{PC} = l\,n(m_{gPC}) = ln(a \cdot m_{gDT}) = ln(m_{gDT}) + l\,n(a) = \mu_{DT} + l\,n(a) \quad (12)$$

Thus, there is an offset of $ln(0.01644) \approx -4.11$ between μ_{DT} and μ_{PC}. The geometric mean of the detectable preclinical phase using the doubling time mean reported by Kuroishi et al. is 2.86 year, resulting in $\mu_{PC} = 1.05$. Because the spread of the distribution does not change by the offset, the standard deviation of $\ln(t_D)$ is the same as the standard deviation of $\ln(t_{PC})$:

$$\sigma_{PC} = \sigma_{DT} \quad (13)$$

2.4.4. Self-detection size distribution

A study that determines the differences in breast cancer presentation was performed in the Hong Kong Sanatorium and Hospital centre was performed by Leung et al. [24] In this study, of the 702 patients with newly diagnosed primary breast cancer 80% of these new tumors were detected by self-discovery, while screening accounted for only 8% of the detected tumors. Self-detected tumors were found to be significantly larger than screen-detected tumors.

A fit to and the distribution of self-detected tumor sizes exhibits approximately a log-normal distribution (Figure 14) with LN parameters $\mu_{SD} = 2.92$ and $\sigma_{SD} = 0.66$. This distribution has been implemented into the SiMRiSc tumor growth model, and for every women with a tumor, a self-detect diameter will be sampled from this distribution using formula (17) and is used to calculate the detectable preclinical phase for that women through formula (6).

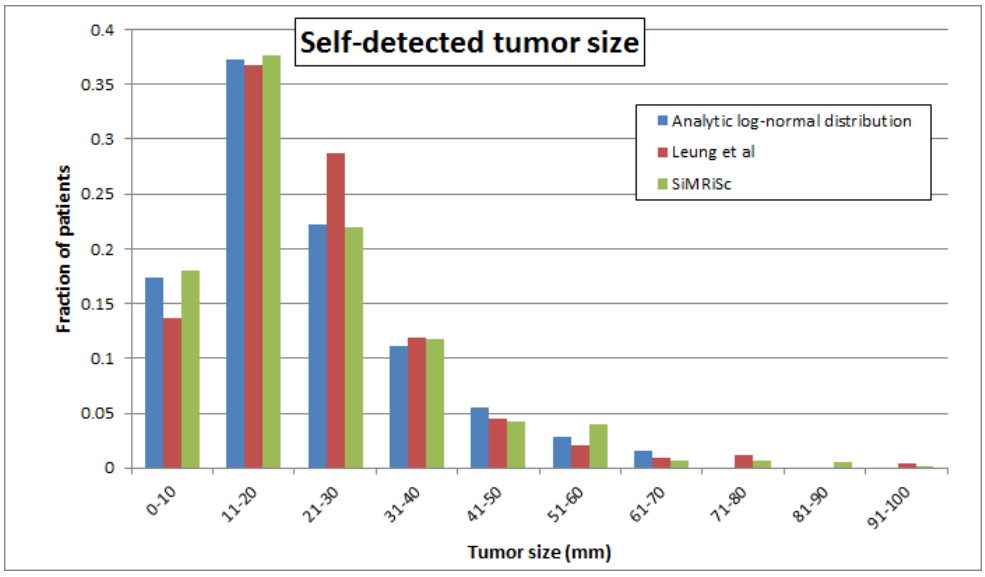


Figure 14. Log-normal fit of the self-detected tumor sizes, with LN parameters μ=2.92 and σ=0.658. The literature source data, the analytic log-normal fit and the results generated by the SiMRiSc model confirm a good fit.

2.4.5. SiMRiSc parameters and spread

The population size of the study of Kuroishi $N = 114$ women. The standard error of the mean (95% confidence interval) is calculated using:

$$SEM = 1.96 * \sqrt{\frac{\mu^2}{N}} = 31.9\ days \tag{14}$$

Giving:

$$UCL = 173.6 + 31.9 = 205.5\ days \tag{15}$$
$$LCL = 173.6 - 31.9 = 141.7\ days \tag{16}$$

for the upper and lower confidence levels of the mean of the tumor volume doubling time. These numbers are nearly the same as those reported by Kuroishi et al. in Table II. The spread in log-normal space can then be calculated by taking the natural logarithm of the UCL and LCL, resulting in Table 10 as the parameters.

The same principle is applied to the self-detected tumor size parameter. With the larger sample size of Leung et al (586 cases) the resulting confidence interval is 0.084.

Table 10. Parameters of the tumor growth model used in SiMRiSc.

Limit of clinical detection (diameter)		5 mm
Tumor volume doubling time	μ_{DT}	5.159 ± 0.2
	σ_{DT}	0.98
Self-detection diameter	μ_{SD}	2.92 ± 0.084
	σ_{SD}	0.658

Note that the sensitivity analysis of SiMRiSc requires the spread in 1σ (~68%) interval figures, so the standard error has to be divided by 1.96.

2.4.6. Internal validation

The SiMRiSc model generates a tumor volume doubling time and a self-detect diameter for every women with a tumor by sampling from their corresponding log-normal distributions through [5]:

$$t_{DT} = e^{\mu_{DT} + Z \cdot \sigma_{DT}} \tag{17}$$
$$D_{SD} = e^{\mu_{SD} + Z \cdot \sigma_{SD}} \tag{18}$$

where Z are standard normal variables generated by a Box-Muller transform of a uniformly distributed random number. If we would set the spread σ_{DT} to 0 and use a fixed self-detection diameter, the algorithm generates only detectable preclinical phases of the geometric mean value, as expected, since $\mu = ln(m_g)$ and $e^{\ln(m_g)} = m_g$.

The results of 1000 women generated by the SiMRiSc model show good agreement to the results to the (transformed to detectable preclinical phase) data of Kuroishi et al and to a binned log-normal distribution with $\mu_{DT} = 1.05$ and $\sigma_{DT} = 0.98$ (Figure 15) when using a fixed self-detect diameter of 20mm. When introducing the spread on the self-detection diameter, there is a clear shift towards shorter detectable preclinical phases, as can be expected as there will be women who will self-detect their tumor when it is still small (<10mm).

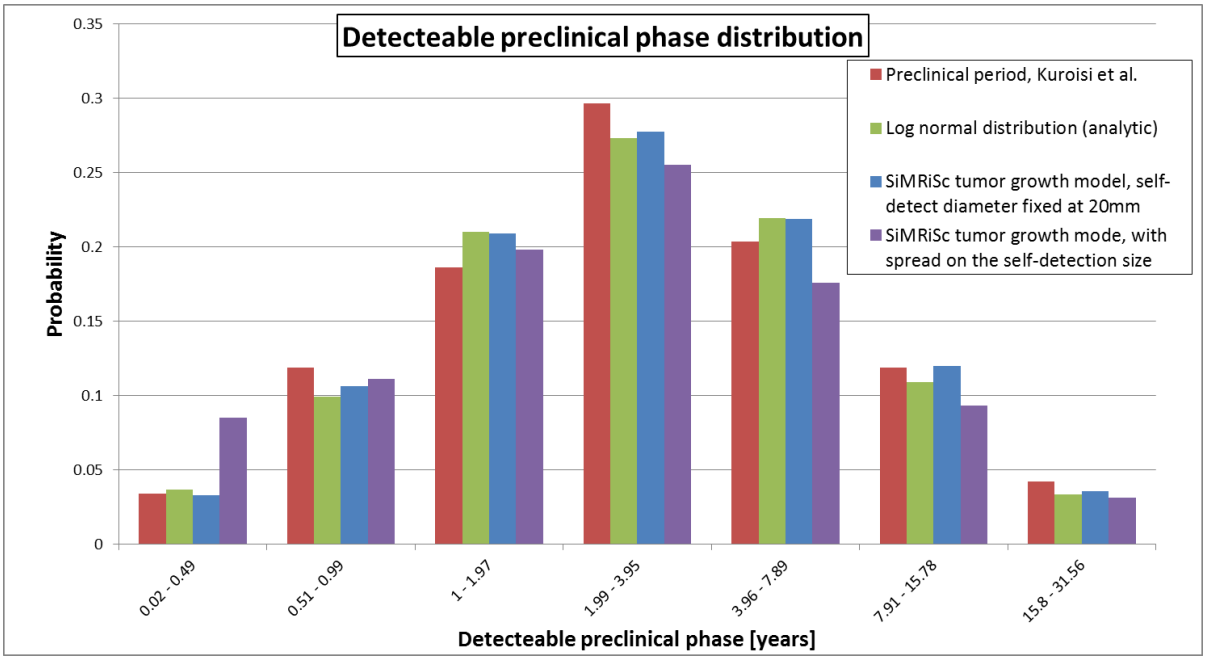


Figure 15. Comparing the detectable preclinical phases generated by SiMRiSc to the (transformed to detectable preclinical phase) data of Kuroishi et al and to a binned log-normal distribution with μ = 1.05 and σ = 0.98.

What is especially to be noted from these results is the wide distribution of the detectable preclinical phase. There are a fair amount of women with very short or very long detectable preclinical phases; ~1/4 women that develop a tumor will have a detectable preclinical phase > 4 year, and ~1/20 women will have a detectable preclinical phase even longer than 10 years.

Figure 16 shows the detectable preclinical phases created by the new tumor growth model compared to the method used by Peer et al. [17] [18] to estimate the detectable preclinical phase.

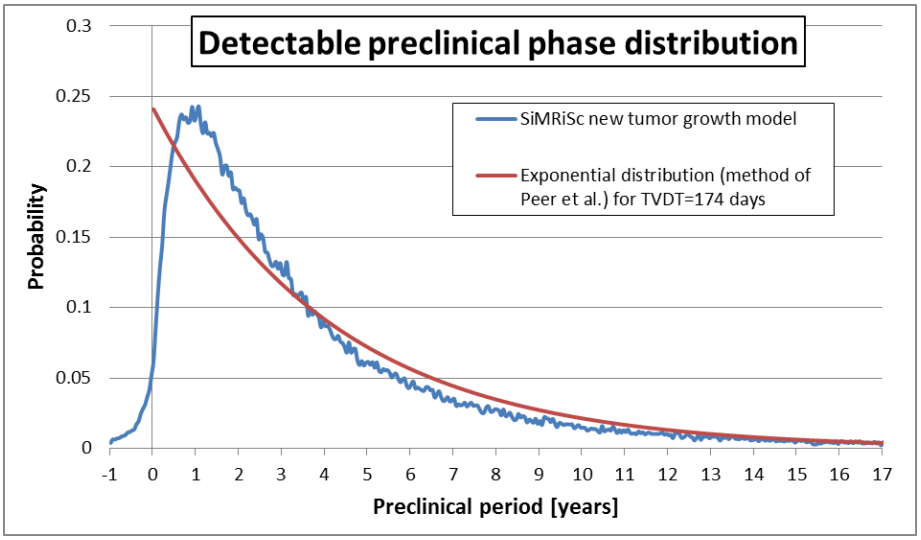


Figure 16. Comparing the results of the detectable preclinical phase distribution of 1000 SiMRiSc women with the distribution using the method of Peer et al. Note that a negative detectable preclinical phase is actually possible if a women self-detects a tumor when smaller than <5mm (limit of clinical detection).

2.5. Tumor survival

Once a tumor is found in SiMRiSc by screening or self-detection, the survival years for that woman is calculated. Survival chance was calculated in previous SiMRiSc versions as a function of tumor size grouped in only 3 categories; tumors grouped in categories of 0-20mm, 20-50mm and >50mm diameter had an associated survival chance. To generate more accurate results, a new breast cancer survival chance has been modelled as a 2D function of tumor diameter at diagnosis versus years after diagnosis. This was done by extending the method of Michaelson et al. [25] for different times after diagnosis to obtain a continuous curve for all times after diagnosis. Kaplan-Meijer data of the Van Nuys Breast Center [25] was used for this approach and a 2D tumor survival chart was obtained (Figure 17).

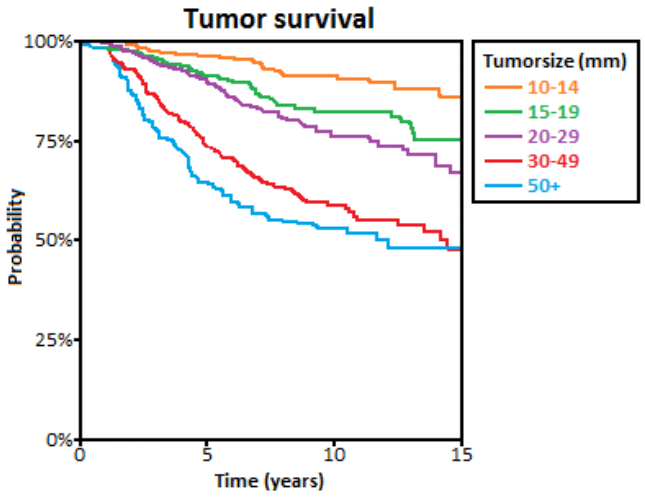


Figure 17. Kaplan-Meier survival curves from the Van Nuys population [25] have been used as the source for the Tumor Survival model.

It is possible to calculate survival chance as a function of tumor diameter only because of the fact that the fraction of women with lethal metastatic disease has been shown to have a strong correlation to primary tumor diameter. [26] Survival chance and tumor diameter are thus directly related. Also, survival data showed no difference in survival of women with clinically detected or mammographic detected tumors of the same size.

Table 11 and Figure 18 shows the raw data of the Van Nuys population with the standard error (95% interval) on this data calculated using equation (19):

$$SE = 1.96\sqrt{\frac{p(1-p)}{N}} \tag{19}$$

where N is the number of patients in that group.

Table 11. Survival rates of the Van Nuys study.

Tumor size (mm)	Median size (mm)	No. patients	Fraction surviving				
			15 years	12.5 years	10 years	7.5 year	5 years
10-14	12	248	86%±4.3%	88%±4%	91%±3.6%	93%±3.2%	96%±2.4%
15-19	17	222	72%±5.9%	81%±5.2%	82%±5.1%	86%±4.6%	91%±3.8%
20-29	25	318	67%±5.2%	74%±4.8%	76%±4.7%	82%±4.2%	89%±3.4%
30-49	39	222	46%±6.6%	54%±6.6%	58%±6.5%	64%±6.3%	74%±5.8%

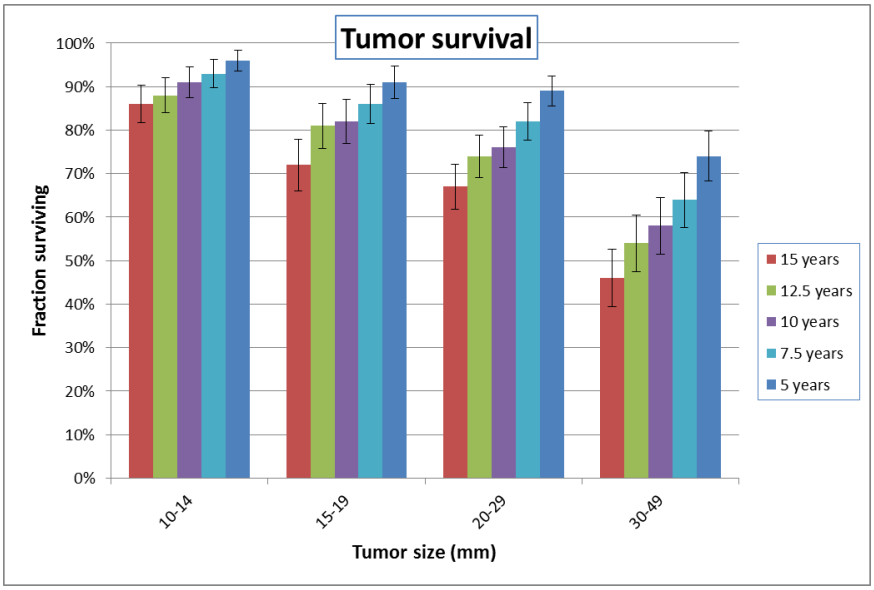


Figure 18. Survival rates of the Van Nuys study in graphical form.

Michaelson et al. [25] has shown the correlation between tumor size and survival can be described by equation (9):

$$\mathrm{F(D)} = \mathrm{e}^{-\mathrm{QD^Z}} \tag{20}$$

where F is the survival chance at 15 years after diagnosis and D is the tumor diameter. Q and Z are constants, and are reported to be 0.0061 and 1.3276 respectively for Van Nuys data (for 15 years after diagnosis) and are found through curve fitting using equation (21).

$$-\ln(F) = QD^Z \tag{21}$$

If the same process that Michaelson performed for 15 years after diagnosis is repeated for 5,7.5,10 and 12.5 years on the Van Nuys data (Figure 19) fits of equation (21) can be made to obtain Q and Z parameters at different times after diagnosis (Table 12). This can also be done for the curves of the upper and lower confidence level of the data to obtain the uncertainty over Q and Z.

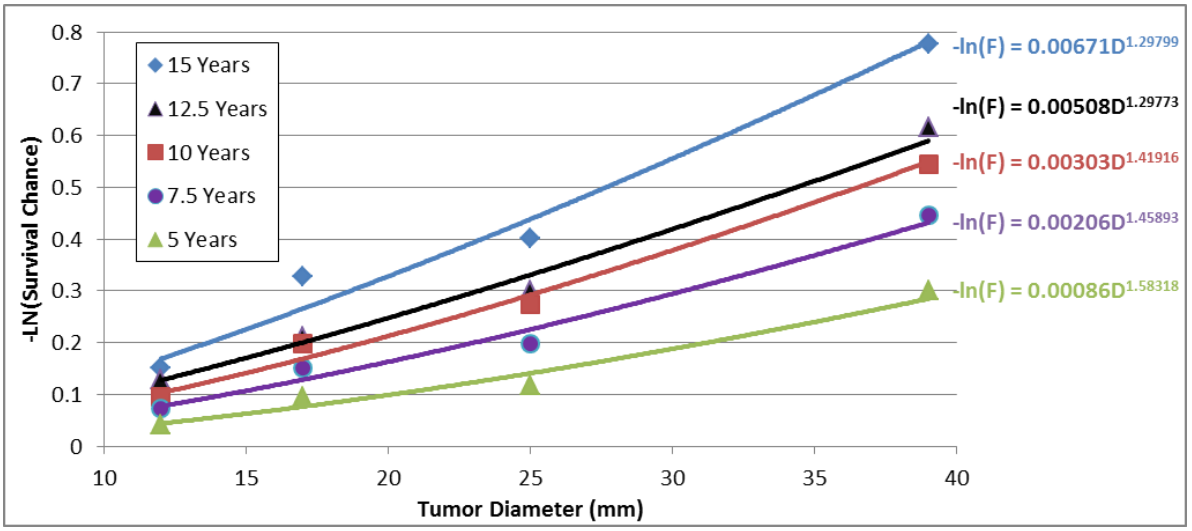


Figure 19. Fitting Q and Z parameters for different times after diagnosis through formula (21).

Table 12. Q and Z parameters for the Van Nuys hospital data, for different times after tumor diagnosis.

	5 years	7.5 years	10 years	12.5 years	15 years
Q	0.00086 ±0.00110	0.00206 ±0.00195	0.00303 ±0.00246	0.00508 ±0.00353	0.00671 ±0.00411
Z	1.583 ±0.376	1.459 ±0.238	1.419 ±0.196	1.298 ±0.150	1.298 ±0.134

Q and Z can then be made into a continuous function (as a function of time after tumor diagnosis) by fitting a power function for Q and a logarithmic function for Z:

$$Q(T) = aT^b \tag{22}$$
$$Z(T) = c \ln T + d \tag{23}$$

where a, b, c and d are constants (Figure 20, Figure 21) and have the values:

a	$4.475 \cdot 10^{-5} \pm 4.392 \cdot 10^{-5}$
b	1.85867 ± 0.420
c	-0.271 ± 0.101
d	2.0167 ± 0.366

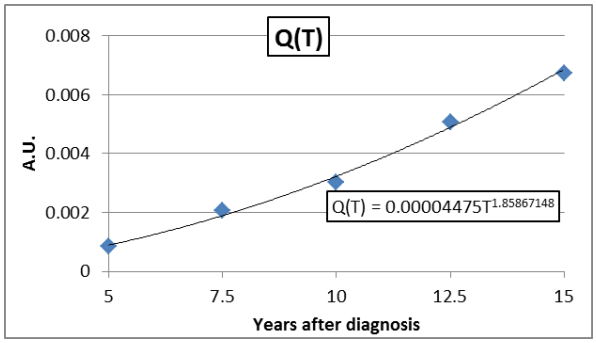


Figure 20. Fitting Q to a continuous power function.

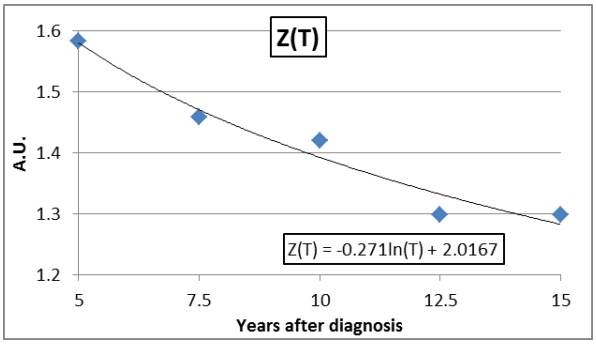


Figure 21. Fitting Z to a continuous logarithmic function.

By substituting (22) and (23) back into equation (20), continuous 2D survival rate curves as function of tumor diameter at diagnosis and years after diagnosis can be created (equation (24), Figure 22):

$$F(D,T) = e^{-Q(T)D^{Z(T)}} \tag{24}$$

where D is the tumor diameter (at diagnosis) and T the time after diagnosis. This equation is solved numerically in SiMRiSc by using a bisection algorithm for the survival years.

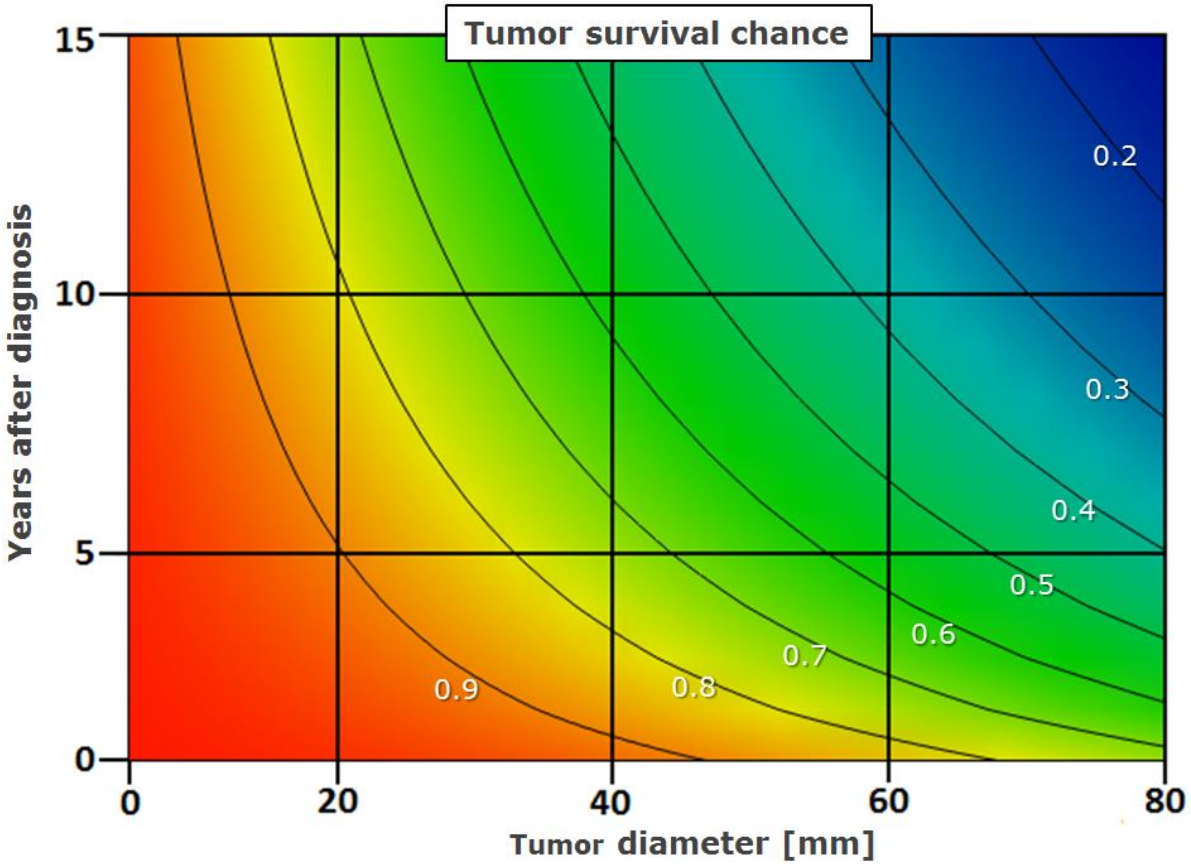


Figure 22. 2D Survival chance as function of years after diagnosis and primary tumor diameter at the moment of diagnosis, generated by formula (24).

2.5.1. Validation of the tumor survival model

Internal Validation

The results of equation (20) and (24) can be compared to see if there is good correspondence after the fitting process (Figure 23).

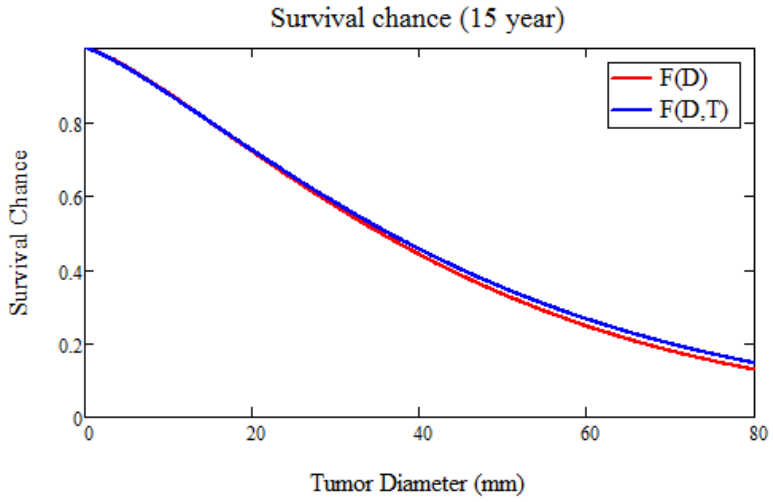


Figure 23. 15 year survival chance as function of tumor diameter. Comparing equation (20) with (24) (with T=15). There is only a very small deviation at large tumor diameters.

The survival rate curves can be internally validated by comparing it to the Kaplan-Meier survival curves and to the discrete data points of the Van Nuys hospital (Figure 24, Figure 25, Table 13).

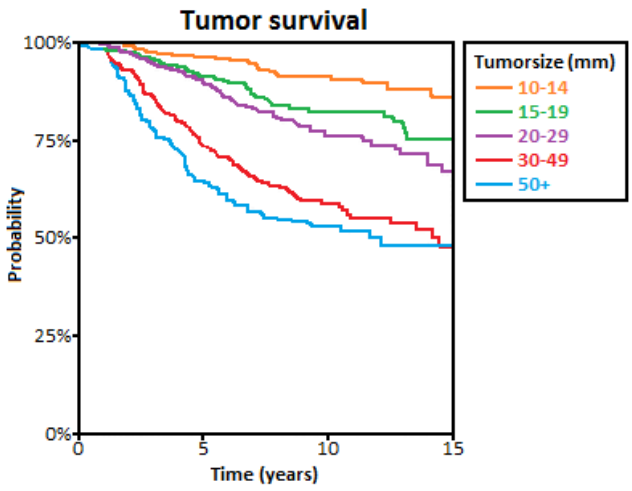


Figure 24. Kaplan-Meier survival curves by the Van Nuys hospital.

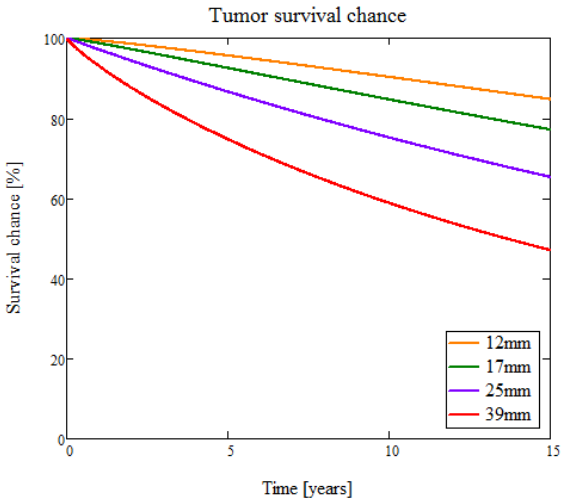


Figure 25. Survival rates generated by formula (24). This gives a good fit compared to the Kaplan-Meier curves by the Van Nuys hospital.

Table 13. Comparison of tumor survival rates 15 years after diagnosis of the model with Van Nuys data.

Tumor survival rate (15 years)	Van Nuys	Model
12mm	88%	84.7%
17mm	72%	77.1%
25mm	67%	65.3%
39mm	46%	47%

External validation

The survival curves have been compared to 4 different external sources, of which 2 are from Hong Kong (King et al. [27] and Kwong et al. [28]) and 2 from western countries (Tabar et al. [29, 25] and the CBS [8]). Tabar et al. and King et al. provide only data points for discrete tumor diameters / stages and diameters after cancer diagnosis, while the Kwong et al. and the CBS sources provide continuous curves.

Table 14. Comparison of tumor survival rates 13.3 years after diagnosis between data published by Tabar et al. [29, 25] (Sweden) and the model.

Tumor survival rate (13.3 years)	Tabar et al.	Model
12mm	87%	86.6%
17mm	80%	79.6%
25mm	55%	68.5%
39mm	44%	50.7%

Breast cancer is typically categorized by tumor stage. Stage I breast cancer are tumors with a size less than two centimeters and has not spread to surrounding lymph nodes or outside the breast. Stage II tumors are 2 to 5cm in diameter, and Stage III tumors are larger than 5cm. Because the model does not account for positive nodes and/or metastasis, tumors with a diameter of 10mm are simply compared to stage I, 35mm to stage II and 60mm to stage III.

Table 15 and Figure 26 until Figure 29 show the comparison between the various sources and the model data.

Table 15. Comparison between tumor survival rates 5 years after diagnosis, reported by King et al. [27] (Hong Kong) and the tumor survival model.

Tumor survival rate (5 years)	King et al.	Model
Stage I / 10mm	100%	96.7%
Stage II / 35mm	78.8%	78.2%
Stage III / 60mm	59.7%	56.2%

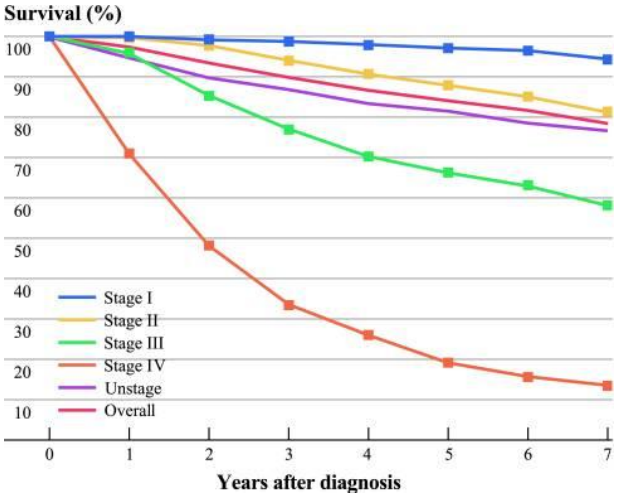


Figure 26. Tumor survival data of Hong Kong by Kwong et al. [28]

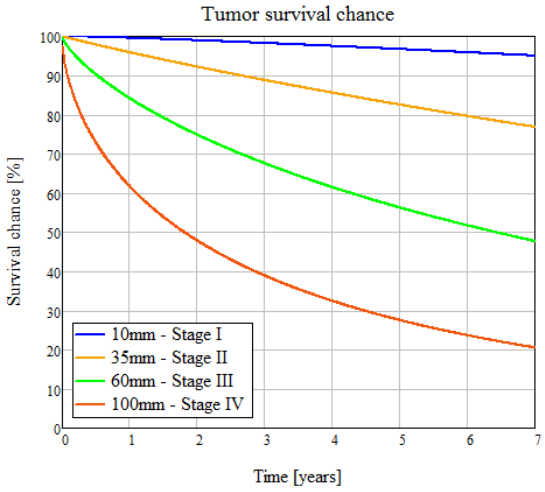


Figure 27. Tumor survival rates of various tumor sizes by the model.

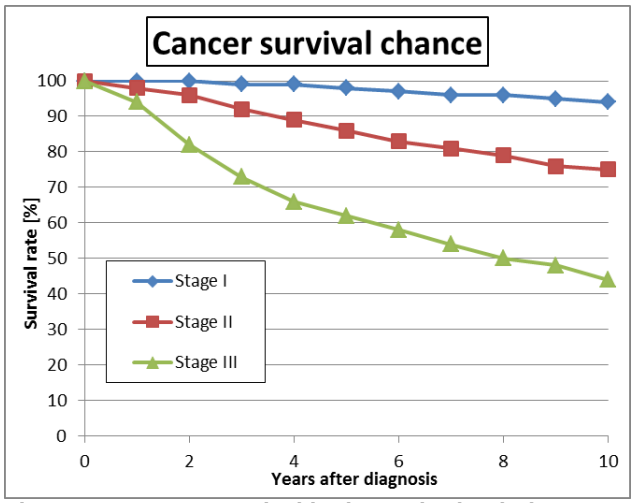


Figure 28. Cancer survival in the Netherlands (18 years and older). Source: CBS. As used by Z. Zhan and E.J Postema for Dutch simulations using SiMRiSc. [8] [10]

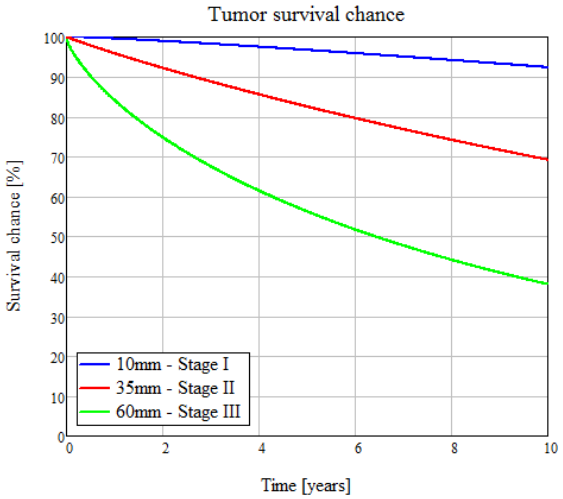


Figure 29. Tumor survival rates of various tumor sizes by the model.

2.6. Discussion

The SiMRiSc model was modified and changed significantly. New tumor growth and survival models have been implemented in SiMRiSc. The growth model matches conceptually better with reality, and now women will self-detect tumors in SiMRiSc according to a distribution. The preclinical period is now back-calculated from the tumor growth speed and self-detect diameter for each woman individually. The tumor survival model now is a fully continuous function. The tumor survival model was validated to multiple sources, both Western and from Hong-Kong. It can be concluded that the new survival curves can be used for SiMRiSc simulations of both urban-Chinese and Western populations.

2.7. Conclusion

The SiMRiSc model has been adapted for the urban Chinese population. Literature and values for the ethnicity specific input parameters have been found and deduced, including error-estimates for the variables based on the population size used in the literature. The different parts of SiMRiSc have succeeded the internal validation. The next step is the external validation of the model, as is presented in the next chapter.

3. External validation

3.1. Introduction

To verify the correct functioning and interaction between the different parts of the model, the 'model-as-a-whole' is validated externally by comparing the results of the model to independent (from the input data) sources/data.

3.2. Material and methods

The results of the SiMRiSc model are validated externally against 2 studies performed in Hong Kong: a screening trial performed in the Hong Kong Kwong Wah Hospital (KWH) by Lui et al. [30] in Chapter 3.3.1 and to 702 consecutive patients referred to the Hong Kong Sanatorium and Hospital by Leung et al. [24] in Chapter 3.3.2.

3.3. Results

3.3.1. Kwong Wah Hospital

During this 5 year study (1998-2002), 46637 screening mammograms have been performed in the KWH, leading to the detecting 232 cases of breast cancer, of which 160 cases of invasive breast cancer. The total detection rate was thus 3.43 per 1000 mammograms.

Furthermore, in this screening program:

- The screening interval is 2 years
- 2 interval cancers have been detected (in total)
- 25% of the detected cancers of which the pathology was available were found to be large (>2cm), 47% small (<2cm) and 28% were DCIS (DCIS is excluded for this validation).

Table 16 and Table 17 list the data for this screening program according to age and screening round. The patients with DCIS have been subtracted. Unfortunately there is no data available combining both tables.

Table 16. Number of woman attending for mammographic screening and the number of cancers detected in the KWH during the screening program in the period of 1998 to 2002.

Age*	Mammograms	Tumors detected	Detection rate per 1000 mammograms
40-44	6678	45	6.68
45-49	15516	50	3.20
50-54	13301	34	2.54
55-59	6559	16	2.42
60-64	2304	8	3.44
>65	2091	7	3.44
Total	**46449**	**160**	**3.43**

* Note there were also 188 mammograms performed in age-group 35-39, but all women in this age-group where considered to be in a high risk group due to a positive family history and have been excluded from our validation.

Table 17. Number of tumors detected according to the screening round.

Screening round	Mammograms	Tumors detected	Detection rate per 1000 mammograms (95% CI)
First	29028	121	4.16
Second	11236	32	2.82
Third	4772	7	1.51
Fourth	1388	5	3.63
Fifth or higher	213	1	3.38
Total	46637*	230	**3.55**
Total, ex. 1st round	17609*	45	**2.54***

* Including the 188 mammograms performed in age-group 35-39.

A SiMRiSc simulation to compare to this data is set up with a start population of 250000 women to suppress the statistical error as much as possible in the results. Screening is performed in 4 rounds with an interval of 2 years, as the number of women having more than 4 rounds is very low in the KWH study. Note that only ~1/6 women of the KWH trial would go through to the third round or higher, while in the SiMRiSc model all women go through all the screening rounds.

First round

Table 18 shows the detection rates comparing only the first screening round of the SiMRiSc simulation and the KWH study.

Table 18. First round screening, comparison between the KWH trial and the SiMRiSc results.

	Mammograms	Tumors detected	Detection rate per 1000 mammograms
SiMRiSc	242121	1337	**5.52**
Kwong Wah Hospital	29028	121	**4.16**

Subsequent rounds

The results of the subsequent rounds are shown in Table 19. A weighted average of the subsequent rounds in SiMRiSc is compared to the weighted average of the subsequent rounds of the KWH study.

Table 19. Subsequent round screening, comparison between the KWH trial and the SiMRiSc results.

	Mammograms	Tumors detected	Detection rate per 1000 mammograms
SiMRiSc	715055	1918	**2.68**
Kwong Wah Hospital	17609	45	**2.54**

Total study

The results of the total studies (first and subsequent rounds combined) are listed in Table 20.

Table 20. Total screening results, comparison between the KWH trial and the SiMRiSc results.

	Mammograms	Tumors detected	Detection rate per 1000 mammograms
SiMRiSc	957176	3255	**3.40**
Kwong Wah Hospital	46637	166	**3.55**

3.3.2. Hong Kong Sanatorium and Hospital

A biannual screening scenario from age 40 to 70 is setup in SiMRiSc. The tumor size distribution generated by SiMRiSc is shown in Figure 30, showing the difference between the self- and screen-detected tumors. The self-detected tumor distribution shown here is the same as shown in Figure 14.

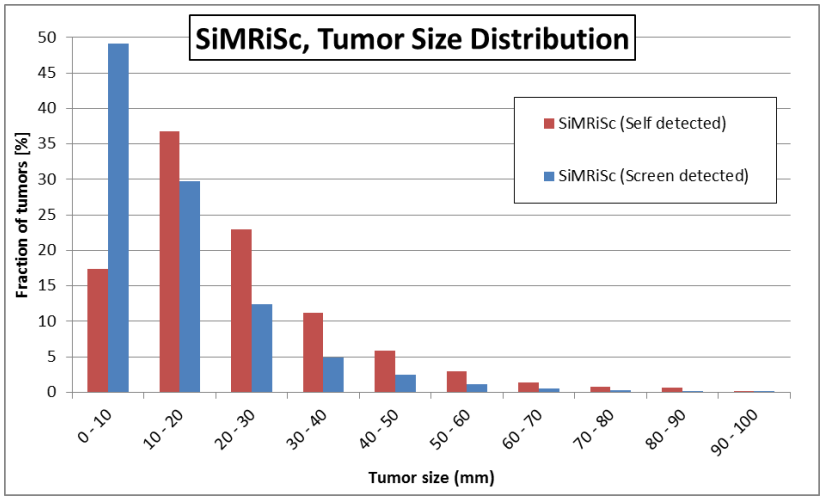


Figure 30. The size distribution of found tumors by self-detection only or with the screening program implemented. Self-detected tumors are significantly larger than screen detected tumors.

Figure 31 shows the same self- and screen detected tumor size distribution according to the study of Leung et al. [24]

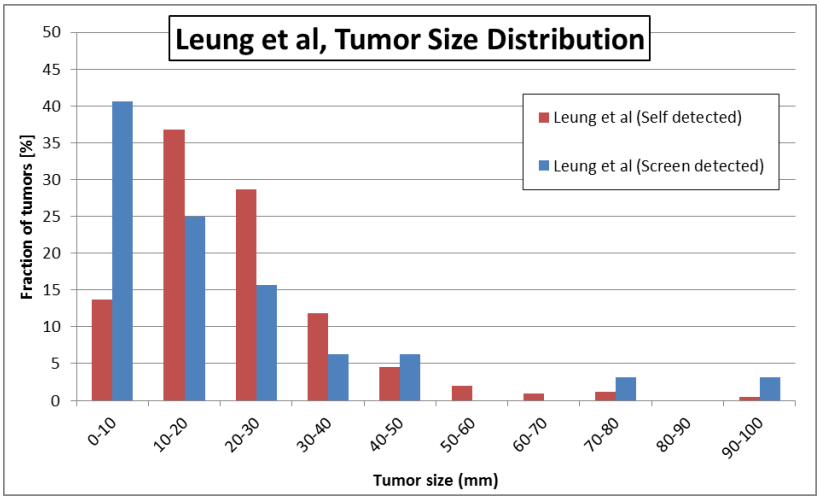


Figure 31. Distribution of tumor sizes of mammographic detected and self-detected tumors, in Hong Kong. [24]

Figure 32 and Figure 33 show the direct comparison between the data of Leung et al and the simulation results.

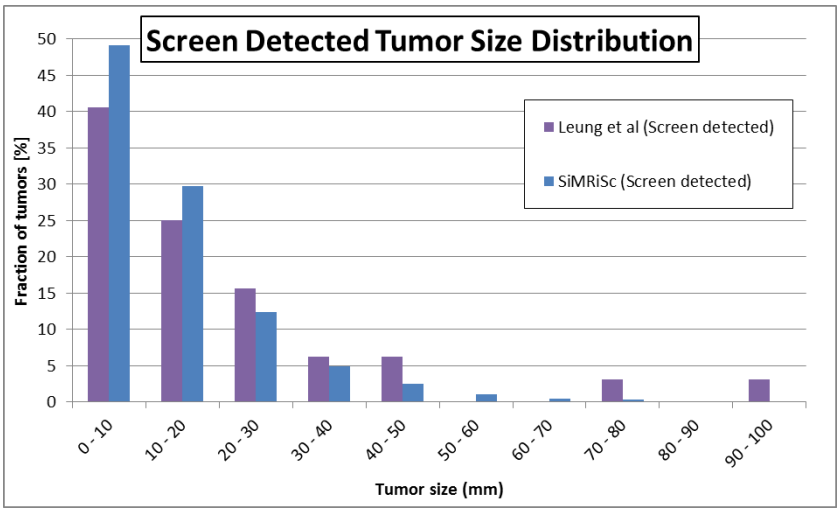


Figure 32. Screen detected tumor size distribution, comparing Leung et al with the simulation results.

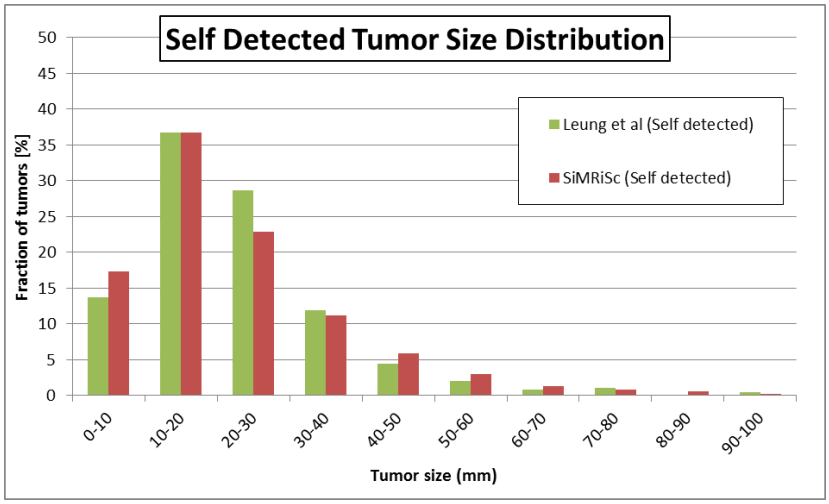


Figure 33. Self-detected tumor size distribution, comparing Leung et al with the simulation results.

3.4. Discussion

3.4.1. Kwong Wah Hospital

Both in the SiMRiSc simulation and the KWH trial the detection rate in the first round is higher than the subsequent rounds. This can be explained by due to the fact that there are a fair amount of women with slow growing tumors (Figure 12), where the tumor might have been present for multiple years. These tumors have a good chance of being detected in the first screening round. The detection rates of the first and subsequent rounds of SiMRiSc simulation and the KWH trial show good agreement.

3.4.2. Hong Kong Sanatorium and Hospital

Assessing the self-detected versus the screen-detected tumors clearly shows the effect of screening, and tumors are being found early at smaller size. Taking the self-detected tumors into account only, the majority of the detected tumors are larger than 10mm. When implementing the screening program, there is a great shift towards smaller tumors.

This same shift is also visible when comparing the self-detected tumor sizes with mammographic detected tumors from data published by Leung et al. The tumor size distributions are very similar for the simulation and the literature data. Mind that these mammographic detected tumors are not detected by a systematic screening program, so cannot be compared to the outcome of the SiMRiSc in a quantitative way.

3.5. Conclusion

The SiMRiSc model behaves as expected. In the first round, the detection ratio is higher than subsequent rounds because slow growing tumors are being detected. This effect is also visible in the screening trial performed at the Kwong Wah Hospital in Hong Kong. The detection rate between the simulations and this trial study seem comparable. Tumor size distributions of self- and screen detected tumors are also very comparable, between the SiMRiSc simulations and patients referred to the Hong Kong Sanatorium and Hospital. Simulations with SiMRiSc are accurately enough to allow predictive scenario simulations.

4. Scenario simulations

4.1. Introduction

In this chapter, the results of different screening scenarios simulations are presented and the effect of the screening in terms of years of life saved (YOLS), costs (in € per year of life saved), number of detected and interval cancers, false positives and tumor size are evaluated.

4.2. Material and methods

The screening scenarios are setup using the input parameters from chapter 2. Each scenario uses the same input population of 100.000 women. This way, the exact influence of the different screening scenarios can be compared for each woman in the simulation.

The screening scenarios are varied by changing the start age of screening between 30 and 50 years, varying the end age between 60 and 70 and using annual or biennial screening intervals. A scenario without screening is included for reference.

The YOLS are calculated by subtracting the sum of the death age of all women without screening from the sum of the death age of all women with screening (Equation (25)). The incremented costs of the scenario are calculated by subtracting the sum of all costs without screening from the sum of all costs with screening (Equation (26)). The average costs per life-year saved (€/YOLS) can then be calculated by dividing these.

$$YOLS = \sum a_{death(with\ screening)} - \sum a_{death(no\ screening)} \quad (25)$$

$$Increased\ costs = \sum woman\ costs_{with\ screening} - \sum woman\ costs_{no\ screening} \quad (26)$$

4.3. Results

Table 21 shows the results of the scenario simulations.

Table 21. Simulated screening scenarios and results, sorted to increasing costs per life-year saved.

Start age	End age	Interval (years)	Tumors Detected *	Tumor deaths *	YOLS *	€/YOLS **
No Screening			0	33.11	-	-
40	**60**	**2**	29.13	26.55	128.98	2655
30	**60**	**2**	35.28	25.36	180.22	2781
40	**70**	**2**	38.87	25.06	145.79	3329
50	**70**	**2**	28.74	27.63	81.56	3961
40	**60**	**1**	34.33	25.14	163.50	4073
30	**60**	**1**	43.08	23.78	226.76	4372
40	**70**	**1**	46.98	23.09	187.19	5104
50	**70**	**1**	34.17	26.08	106.77	5889

* Per 1000 women ** Assuming €0.13/Yuan

4.4. Discussion

When implementing the '40-60-2' scenario (biennial screening starting at 40 and ending at age 60), the mortality of breast cancer would go down from 33.11 per 1000 women to 26.55 per 1000 women, with 128.98 years of life saved per 1000 women. The increased healthcare costs for this are €2655 per life year saved, the most cost effective of all tested scenarios.

Then, when starting screening for 10 years earlier, at age 30, the mortality further decreases to 25.36 per 1000 women, but the gained life years go up significantly (to 180.22) while the costs per won life year goes up only marginally (to €2781). Starting with screening earlier seems like a sensible thing to do.

When screening from 40 until age 70 (screening for 10 years longer compared to the 40-60 scenario) results in a lower breast cancer mortality than the '30-60' scenario, but the life-years saved is lower; this can be explained by the fact that a tumor found at earlier age results in more life year saved than a tumor found at later age.

The 50-70-2 scenario (like implemented in the Netherlands) has only disadvantages compared to the scenarios discussed above. This can partly be explained by the earlier peak-age of the breast-cancer incidence in the Chinese urban population; the incidence peaks over 10 year earlier (see paragraph 2.3.2) and also because of much increased overdiagnosis (according to 'definition 2' of chapter 6) when screening at later ages.

Changing the screening interval for the various scenarios from 2 years to 1 year significantly reduces the cost effectiveness, but also the won life-years increase with about 25%.

4.5. Conclusion

Breast cancer mortality is simulated to go down from 33.11 tumor death per 1000 born women to 26.55 when implementing a biennial screening scenario starting from age 40 until age 70, giving a 129 years of life saved per 1000 women at a costs of €2655 per YOLS. Starting screening 10 years earlier, at age 30 further increases the YOLS to 180 years per 1000 women, while increasing the costs slightly to €2781 per YOLS. Starting screening earlier seems like a sensible thing to do. Changing the screening interval from 2 years to 1 year decreases the mortality with about 25%, but reduces the cost effectiveness to over €4000 per YOLS. Shortening the screening interval really comes down to willingness to spend.

5. Sensitivity analysis

5.1. Introduction

The SiMRiSc model and the results are studied by analyzing the propagation of uncertainty of the input parameters on the results by performing a sensitivity analysis. This will show the variance in the results, and give a good indication of which of the input variables contribute most to the uncertainty of the output. Uncertainty in the output in SiMRiSc can be caused by 2 things (Figure 34):

- Uncertainty in the input population.
- Uncertainty in the input parameters.

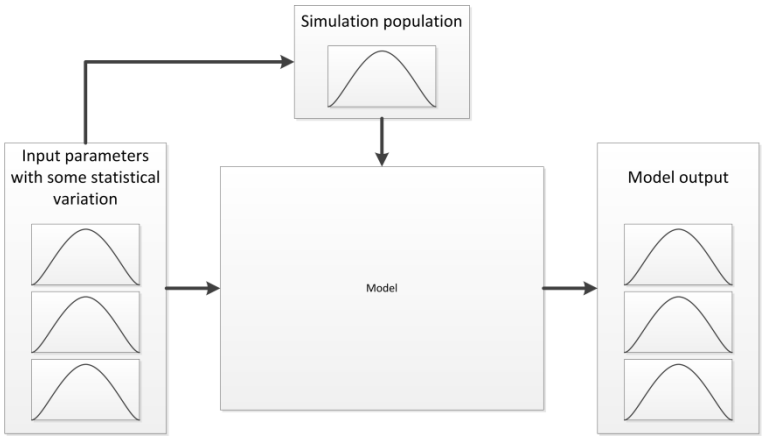


Figure 34. Input variables will have uncertainty, due to finite sample size used in the study. The simulation population will also add uncertainty, due to a finite number of women in the simulation.

5.2. Material and methods

The sensitivity analysis is performed by running (iterating) the same scenario 100 times, each time using a different seed of the random number generator, so a different input population or set of random sampled input parameters is taken. For each iteration, a simulation without screening using the same input parameters/population has to be run, in order to provide the baseline to calculate equations (25) and (26). The output is a plot of the years of life saved versus additional costs of each iteration, visualized in a scatter plot. From the scatter plots, the CEAC (Cost effectiveness acceptability curves) can be created, which show the variation of the output.

3 different sensitivity analyses are run:

- The first sensitivity analysis is performed for different SiMRiSc population sizes, by using a different seed for each iteration, effectively ‘picking’ a random population of N each time, showing the variance in the output because of the statistical uncertainty of a finite population size. The 40-60-2 scenario is used for this analysis.
- A sensitivity analysis is run by varying one of the input parameters each iteration, by sampling this input parameter from the distribution according to the uncertainty described in Chapter 2 on that input parameter. The other input parameters are kept constant at their nominal value, and the same input population of 100.000 is used. This way, the contribution of only that input variable on the variance of the output is shown. Again the 40-60-2 scenario is used.

- Finally, the sensitivity analysis is run for all the scenarios of Chapter 4. Again, the same input population of 100.000 is used for each iteration, but now all parameters are randomly sampled, showing the total variance in the cost effectiveness of each scenario due to the uncertainty of all input parameters.

5.3. Results

5.3.1. Population size

Figure 35 shows CEAC curves for the 40-60-2 scenario using different population sizes of 10000, 25000, 50000 and 100000 women per iteration. Scatter plots are attached as Appendix C.

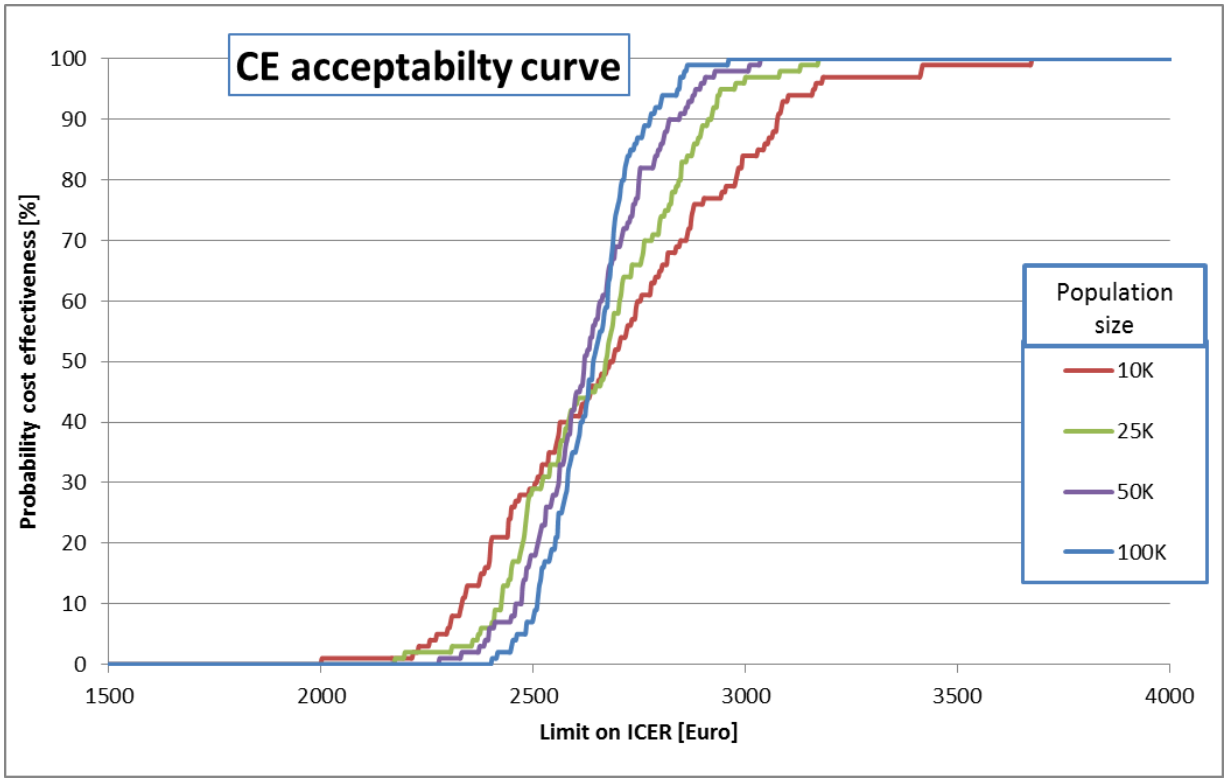


Figure 35. 100 simulations using a random population of various sizes, using the 40-60-2 scenario.

5.3.2. Model input parameters

Figure 36 shows the SEAC curves when varying the input parameters. See Appendix B for ICER scatter plots and CEAC curves of all the input parameters.

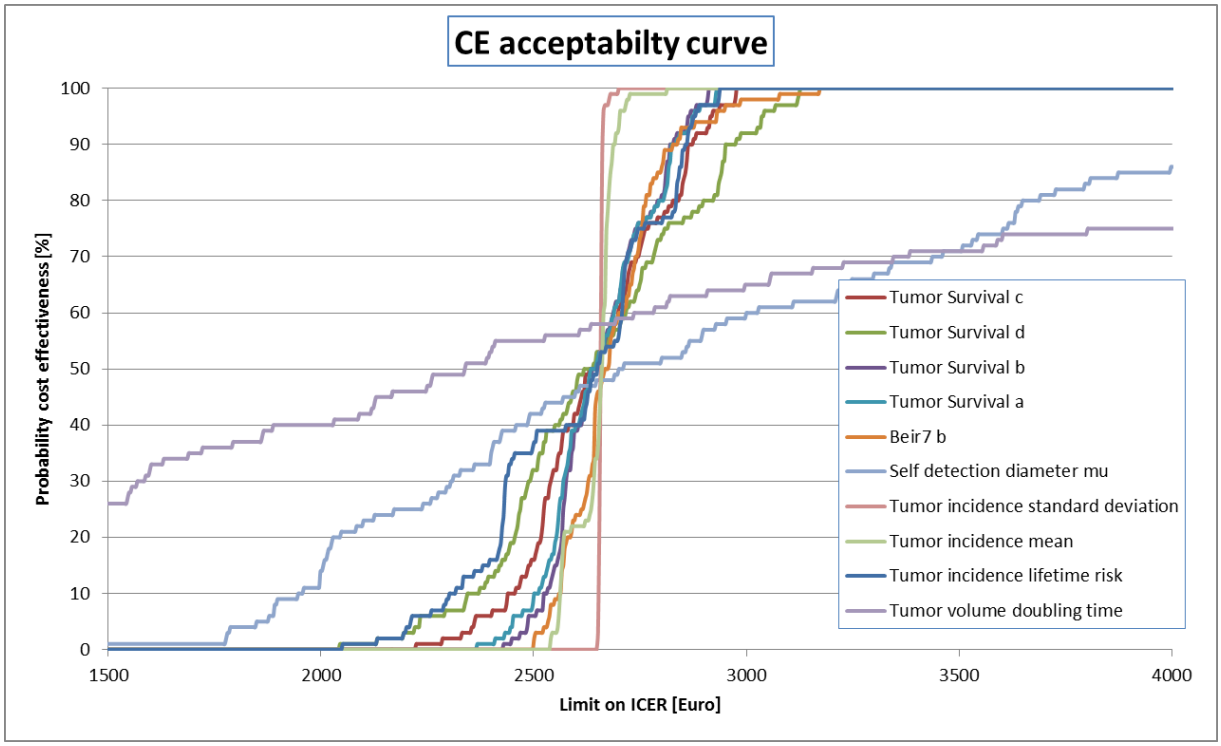


Figure 36. CEAC curves variating only single input parameters, using the 40-60-2 scenario.

5.3.3. Scenarios

Figure 37 shows the CEAC curves of the scenarios of Chapter 4.

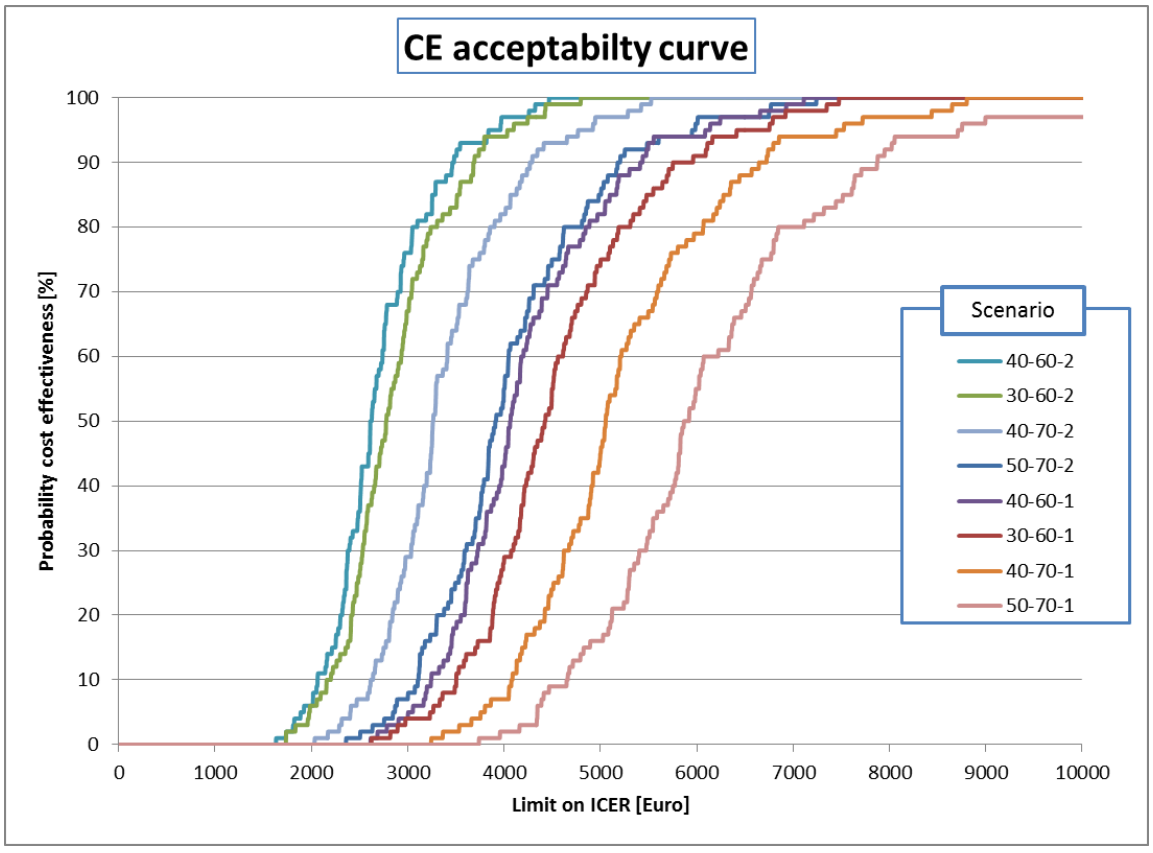


Figure 37. Cost effectiveness acceptability curves of different screening scenarios. 100 simulations using a population of 100.000 women per scenario.

The spread on the costs per life-year saved can also be visualized as a histogram. Figure 38 shows the histogram for the 40-60-2 scenario. Scatter plots, CEAC curves and histograms for all scenarios are in Appendix A.

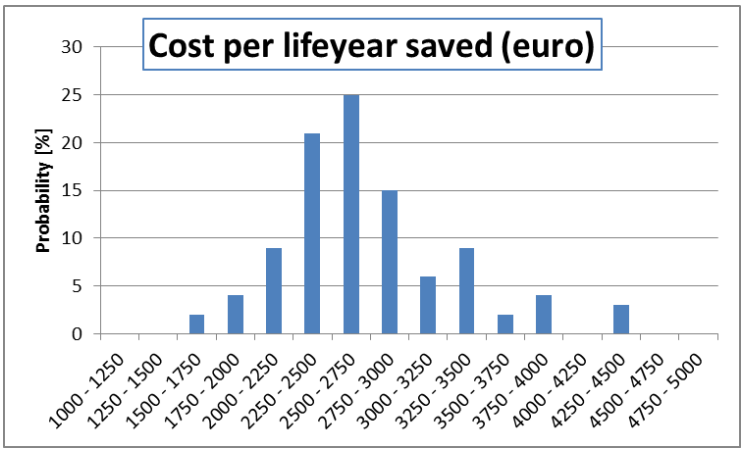


Figure 38. Uncertainty on the costs per life-year saved for the 40-60-2 scenario.

5.4. Discussion

As expected, using a smaller simulation population gives a larger variance of the populations (Figure 35). Using a larger population per simulation decreases this statistical spread, but also increases the computation time. When simulating with a population size of 100.000 women keeps, the uncertainty because of the population is about the same magnitude as the spread because one of the input parameters (see next paragraph), while the simulation time is still reasonable (~10 minutes for 100 iterations).

Performing a sensitivity analysis over the individual input parameters is helpful to identify the input parameters that have the largest contribution to the uncertainty in the output. Extra attention could be given to these input variables (perhaps by finding literature sources with less uncertainty for these variables) to decrease the uncertainty in the output. The tumor volume doubling time and self-detection diameter are shown to have the greatest influence on the variance of the output. The tumor induction (Beir 7) and mean age of the tumor incidence have almost no influence on the outcome.

The 50% point of the ICER (incremental cost-effectiveness ratios) for the scenarios corresponds to the median €/YOLS for that scenario (Table 21). Rough 95% confidence interval estimation for the 40-60-2 scenario is €1900 to €3500 per year of life saved.

5.4.1. Further applications of the sensitivity analysis

The ICER scatter plots are also useful for validation of the model, and can be used to search for unexpected relationships between inputs and outputs. For example, Figure 39 displays the scatterplot for a parameter of the tumor survival model, and proves that the tumor survival model does not influence costs, only years of life saved, as expected. In principle these plots could be made for any input/output relation (for example, number of detected tumors, interval tumors, etc), and the response of these outputs could be analyzed to detect unexpected behavior, but is beyond the scope of this research.

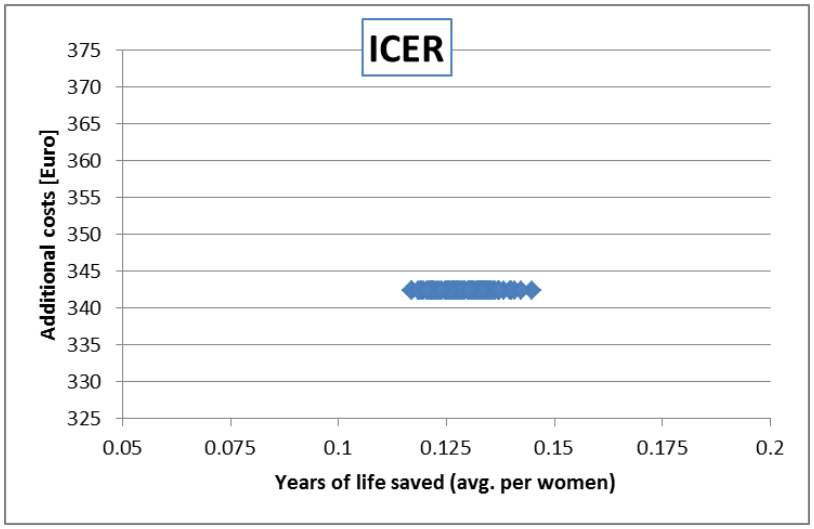


Figure 39. Scatter plot of the 'a' parameter of the tumor survival model. The tumor survival model shows no influence on the costs, only on the years of life saved.

5.5. Conclusion

A sensitivity analyses have been run to analyze the SiMRiSc model and uncertainty of the output. The tumor volume doubling time and self-detection diameter are shown to have the greatest influence on the variance of the output. In addition to the input parameters, the finite population size adds uncertainty. 100.000 women per iteration is determined as the minimal population size to be used to keep the uncertainty because of the population to about the same magnitude as the spread because one of the input parameters. ICER plots of the scenarios are created, showing the variance on the cost-effectiveness.

6. Overdiagnosis

6.1. Introduction

In case of overdiagnosis, a tumor is detected that will never cause symptoms or death during the women's life.

6.2. Material and methods

Two definitions of what is overdiagnosis and how to calculate the overdiagnosis percentage in SiMRiSc can be made:

- Over-diagnosed tumors are screen detected, but would not be self-detected without screening
- Over-diagnosed tumors are screen detected, but the women would die of another (than tumor) cause without screening

6.3. Results

Table 22 presents the overdiagnosis of the screening scenarios.

Table 22. Overdiagnosis of tumors of the various scenarios, according to definition 1 and definition 2

Start age	End age	Interval (years)	Tumors detected *	Over-diagnosed tumors (def 1) *	Over-diagnosis (def 1)	Over-diagnosed tumors (def 2) *	Over-diagnosis (def 2)
No Screening			0				
40	**60**	**2**	29.13	0.59	2.03%	13.91	48%
30	**60**	**2**	35.28	0.66	1.87%	16.1	46%
40	**70**	**2**	38.87	1.08	2.78%	20.85	54%
50	**70**	**2**	28.74	0.82	2.85%	16.68	58%
40	**60**	**1**	34.33	0.64	1.86%	16.64	48%
30	**60**	**1**	43.08	0.88	2.04%	20.29	47%
40	**70**	**1**	46.98	1.35	2.87%	25.73	55%
50	**70**	**1**	34.17	0.9	2.63%	20.17	59%

* Per 1000 women

6.4. Discussion

The 2 definitions of overdiagnosis result in vastly different numbers for the overdiagnosed tumors per scenario.

Definition 1 tells how much tumors are detected that really would have had no influence on the women's life. These numbers are quite low, as the chance is high that the women would self-detect the tumor at some point in her life when it grows larger.

But, as definition 2 shows, a large fraction of the screen-detected tumors do not yield any extra life-years gained. This can be explained by the fact that in the case where no screening is implemented, and women find cancer by means of self-detection, a large fraction of the cancers are cured or delayed long enough that the women will die of other causes. In these cases, screening can still detect tumors earlier and at a smaller size, but it does not lead to any life years gained or contribute to the effectiveness in terms of €/YOLS for the screening program. However, screening might still have other advantages in these cases, like detecting the tumor while it is small enough to allow breast conserving surgery while this might not have been the case without screening. Also to be noted is that this overdiagnosis percentage of definition 2 goes up when either the start- or end-age of

screening is later. For example, the 50-70 scenarios have a higher overdiagnosis than the 40-60 scenarios. This can be explained by the fact that when a tumor onsets at a later age, and is detected by screening at a later age, the chance of the women dying naturally before dying because of the tumor increases.

It can be concluded that a large fraction (~50%) of the screen detected tumors do not lead to life-years gained, and thus also not contribute to the cost-effectiveness (€/YOLS) of the screening.

SiMRiSc does not add 'pseudo-disease' to the population. All the tumors in SiMRiSc are 'real' tumors that would eventually cause symptoms, because the tumor-incidence curves are from an unscreened population. Very slow progressing or even regressing tumors would not be self-detected, and are thus also not added to the tumor incidence statistics. When implementing screening, some of these tumors may very well be detected and this is the actual cause of overdiagnosis (Figure 40).

These cases of screen-detected overdiagnosis will also be added to the tumor incidence statistics. One hypothesis could be that the large difference in breast cancer incidence between the western world and China are (partly) due to screening. Increased awareness of women and better healthcare in general may also lead to increased detection of tumors that would otherwise never cause symptoms.

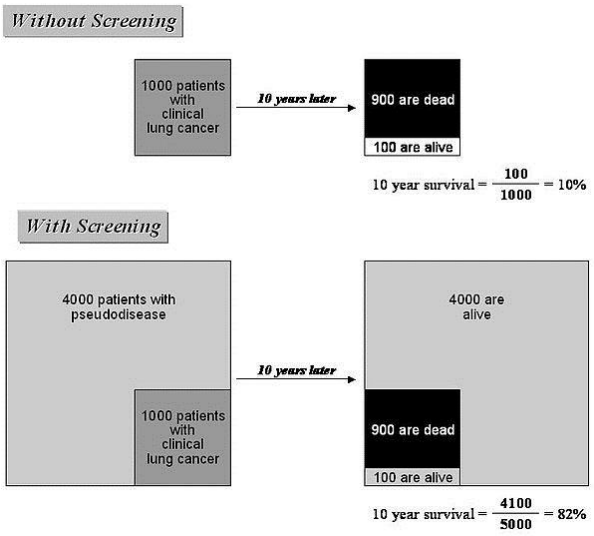


Figure 40. Exaggerated figures for the sake of the argument. Overdiagnosis may be the cause of increased tumor incidence. Also the survival rates may increase when implementing screening, but this effect may not be entirely 'true'. Figure from [31].

6.5. Conclusion

Because SiMRiSc does not add pseudo-disease to the population, it is doubtful that overdiagnosis can really be simulated with SiMRiSc. If the pseudo-disease detected tumors are treated and determine a large fraction of the hospitalization costs, it is doubtful that the cost-effectiveness analysis performed with SiMRiSc is valid. One suggestion to allow simulating overdiagnosed tumors is by implementing a Gompertz function in the tumor growth model to allow for stagnating tumors in the simulation (see chapter 7.1.2).

The actual causes and magnitude of overdiagnosis, and the influence on the screening model is beyond the scope of this research.

7. General Discussion

7.1. Further work

Some suggestions are given to improve or extend the results presented in this thesis.

7.1.1. Urban Chinese simulations

More detailed scenarios could be simulated. This study only performed very 'rough' scenario's; only varying the start- and end-age of screening with 10 years, and only looking at annual and biennial screening.

Instead of calculating the death rates, tumor incidence and detection rates per 1000 born women, these should be calculated as age-standardised rates (thus as an amount per 100.000 women alive per age-group). This would allow comparison to other studies and literature.

The sensitivity analysis could be used to analyse different output parameters, like tumor size and detection rate, to generate error bars on the data of chapter 3.

7.1.2. Possible improvements of the SiMRiSc model

Here some possible improvements of the SiMRiSc model are discussed. While the technical implementation of these suggestions is not that complicated, the difficult part will be to find accurate literature to provide input parameters for these suggested models.

Tumor growth model and overdiagnosis

The exponential growth model currently implemented in SiMRiSc is the simplest model. As this model does not allow for growth deceleration and/or take metabolic considerations into account. One model of bounded growth that could be implemented in future versions of SiMRiSc makes use of a Gompertz Function. [32] Implementing such a model could also be the first step to accurately simulate overdiagnosis, as this allows simulation of stagnating tumors that will never be self-detected or cause symptoms.

Mammographic sensitivity

Currently, the mammographic sensitivity in SiMRiSc is not a function of tumor size. It is recommended that for a future improvement of SiMRiSc this is implemented according to the work of Michaelson et al, and sensitivity becomes one continuous function of tumor size and breast density, removing the need for the concept of a 'hard' limit of mammographic detection.

7.2. Conclusions

The aim of this study was to assess the benefits and cost-efficiency of implementing systematic mammographic breast cancer screening in the urban Chinese population. This was done implementing and using the SiMRiSc breast cancer screening model.

Literature and values for the ethnicity specific input parameters have been found and deduced. The SiMRiSc model was modified and changed significantly. New tumor growth and survival models have been implemented in SiMRiSc. The growth models matches conceptually better with reality, and now women will self-detect tumors in the model according to a distribution. The preclinical period is now back-calculated from the tumor growth speed and self-detect diameter for each woman individually. The tumor survival model now is a fully continuous function. The tumor survival model was validated to multiple

sources, both Western and from Hong-Kong. It can be concluded that the new survival curves can be used for SiMRiSc simulations of both urban-Chinese and Western populations.

The SiMRiSc model as a whole was validated externally against 2 hospital studies from Hong Kong. This showed that with screening tumors are detected at a significantly smaller size than without screening, giving an improved survival chance for women.

The results of different screening scenarios that have been simulated include the number of cancers detected and the incremental costs per life-year saved. These simulations suggest that starting early with screening, at age 30, seems to be cost-effective for the urban Chinese population, resulting in a cost of €2781 per life year gained. Screening does not yield any life-year gained for a large fraction (~50%) of women with screen-detected tumors. However, the tumor will still be detected early and at a smaller size.

A sensitivity analysis has been performed and shows that the tumor volume doubling time is the greatest contributor to the uncertainty in the results, with the self-detection diameter second. Running simulations with 100000 women is deemed the minimum population size to keep statistical variance because limited.

Because SiMRiSc does not add pseudo-disease to the population, it is doubtful that overdiagnosis can really be simulated with SiMRiSc. If the pseudo-disease detected tumors are treated and determine a large fraction of the hospitalization costs, it is doubtful that the cost-effectiveness analysis performed with SiMRiSc is valid. A suggestion has been given to implement stagnating tumors to the tumor growth model to allow for simulation of overdiagnosed tumors.

8. Bibliography

[1] P. Porter, ""Westernizing" Women's Risks? Breast Cancer in Lower-Income Countries," *The new england jornal of medicine,* vol. 358, no. 3, pp. 213-216, 2008.

[2] J. Yuan, M. Yu, R. Ross, "Risk Factors for Breast Cancer in Chinese Women in Shanghai," *Cancer Research,* vol. 48, pp. 1949-1953, 1988.

[3] Hong-Kong Breast Cancer Foundation, "Local Statistics," [Online]. Available: http://www.hkbcf.org/article.php?aid=138&cid=6&lang=eng.

[4] H. K. H. Authority, "Hong Kong Cancer Registry," [Online]. Available: http://www3.ha.org.hk/cancereg/statistics.html.

[5] M. Greuter, M. Jansen-van der Weide, M. Jacobi, J. Oosterwijk, L. Jansen, M. Oudkerk, G. de Bock, "The validation of a simulation model incorporating radiation risk for mammography breast cancer screening in women with a hereditary-increased breast cancer risk," *European journal of cancer,* vol. 46, no. 3, pp. 395-504, 2010.

[6] G. de Bock, "Which screening strategy should be offered to women with BRCA1 or BRCA2 mutations? A simulation of comparative cost-effectiveness," *British Journal of Cancer,* vol. 108, pp. 1579-1586, 2013.

[7] W. Lu, "Safety and cost-effectiveness of shortening hospital follow-up after breast cancer treatment," *British Journal of Surgery,* vol. 99, pp. 1227-1233, 2012.

[8] Z. Zhan, "Validation of the SiMRiSc model for breast cancer screening in the Netherlands," 2013.

[9] C. a. S. D. -. T. G. o. t. H. K. S. A. Region, "The Mortality Trend in Hong Kong, 1981 to 2013," [Online]. Available: https://www.censtatd.gov.hk/hkstat/sub/sp160.jsp?productCode=FA100094.

[10] E. J. Postema, "Cost-effect analysis of the Dutch Breast Cancer screening programme with a lowered onset age," 2014.

[11] Committee to assess health risks from exposure to low levels of ionizing radiation, Health risks from exposure to low levels of ionizing radiation, 2006.

[12] L. Ellison-Loschmann, "Age and Ethnic Differences in Volumetric Breast Density in New Zealand Women: A Cross-Sectional Study," *PLoS ONE,* vol. 8, no. 7, p. 1, 2013.

[13] J. Michaelson, "Estimates of the Sizes at Which Breast Cancers Become Detectable on Mammographic and Clinical Grounds," *Journal of Women s Imaging,* vol. 5, no. 1, pp. 3-10, 2003.

[14] V. Collins, "Observations on growth rates of human tumors," *ARR Am J Roentgenol,* vol.

76, pp. 988-1000, 1956.

[15] M. Schwarz, "A biomathematical approach to clinical tumor growth," *Cancer,* vol. 14, pp. 1272-1294, 1961.

[16] N. Obuchowski, "Ten criteria for effective screening: their application to multislice CT screening for pulmonary and colorectal cancers," *AJR Am J Roentgenol,* vol. 176, no. 6, pp. 1357-1362, 2002.

[17] P. Peer, "Age-dependent growth rate of primary breast cancer," *Cancer,* vol. 71, no. 11, pp. 3547-3551, 1993.

[18] P. Peer, "Age-specific sensitivities of mammographic screening for breast cancer," *Br Cancer Res Treat,* vol. 38, no. 2, pp. 153-160, 1996.

[19] T. Kuroishi, et al., "Tumor Grpwth Rate and Prognosis of Breast Cancer Mainly Detected by Mass Screening," *Jpn. J. Cancer Res,* vol. 81, pp. 454-462, 1990.

[20] J. Michaelson, "Estimates of Breast Cancer Growth Rate and Sojourn Time from," *Journal of Womens's imaging,* vol. 5, pp. 11-19, 2003.

[21] D. Fournier, "Growth rate of 147 mammary carcinoma," *Cancer,* vol. 45, pp. 2198-2207, 1980.

[22] B. Lundgren, "Observations on the growth rate of breast carcinomas and its possible implications for lead time".

[23] J. Spratt, "Mammographic assessment of human breast cancer growth and duration," *Cancer,* vol. 71, pp. 2020-2026, 1993.

[24] A. Leung, "Clinicopathological correlates in a cohort of Hong Kong breast cancer patients presenting with screen-detected or symptomatic disease," *Hong Kong Med J,* vol. 13, no. 3, pp. 194-202, 2007.

[25] J. Michaelson, "Predicting the Survival of Patients with Breast Carcinoma using Tumor Size," *Cancer,* vol. 95, no. 4, pp. 713-723, 2002.

[26] J. Michaelson, "Mammographic Screening: Impact on Survival," pp. 465-471.

[27] W. King, "Analysis of 346 Chinese patients with breast cancer," *Hong Kong Med J,* vol. 2, no. 1, pp. 72-79, 1996.

[28] A. Kwong, "Breast Cancer in Hong Kong, Southern China: The First Population-Based Analysis of Epidemiological Characteristics, Stage-Specific, Cancer-Specific, and Disease-Free Survival in Breast Cancer Patients: 1997-2001," *Annals of surgical oncology,* vol. 18, no. 11, pp. 3072-3078, 2011.

[29] L. Tabar, "A new era in the diagnosis of breast cancer.," *Surg Oncol Clin North Am.,* vol.

9, pp. 233-277, 2000.

[30] C. Lui, "Opportunistic breast cancer screening in Hong Kong; a revisit of the Kwong Wah Hospital experience," vol. 13, no. 2, pp. 106-113, 2007.

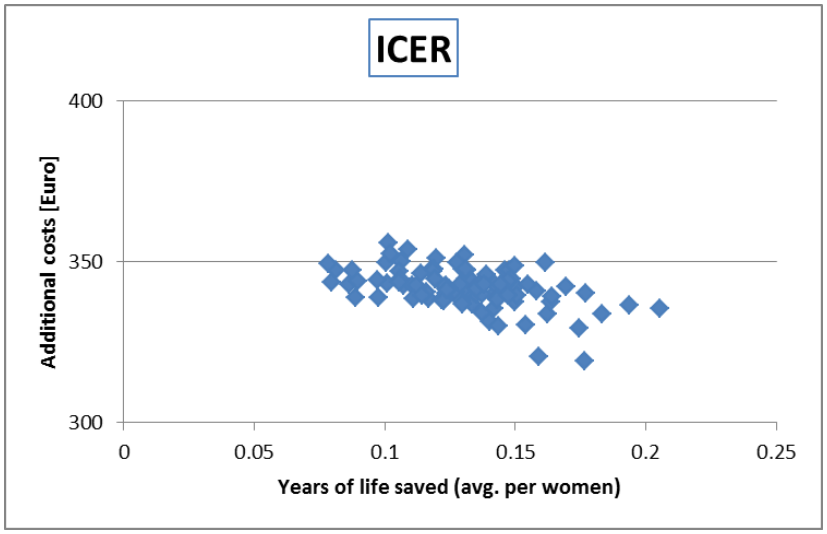
[31] [Online]. Available: https://en.wikipedia.org/wiki/Overdiagnosis.

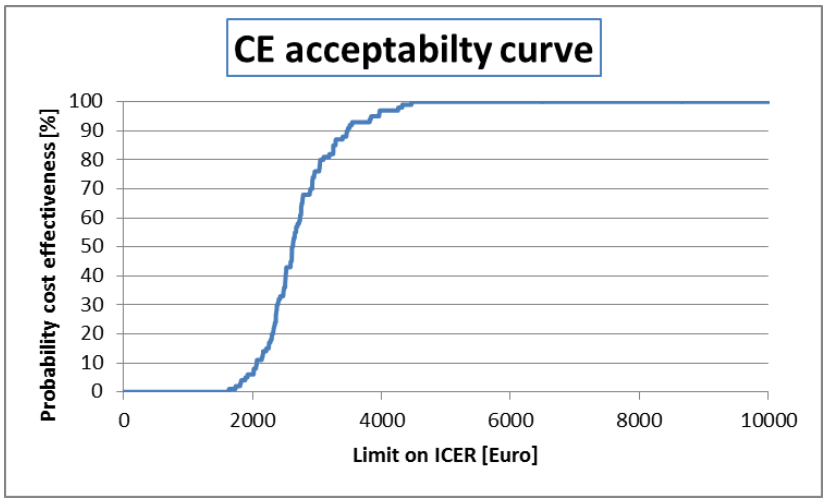
[32] D. Cameron, "Mathematical Modelleing of the Respone of Breast Cancer to Drug Therapy".

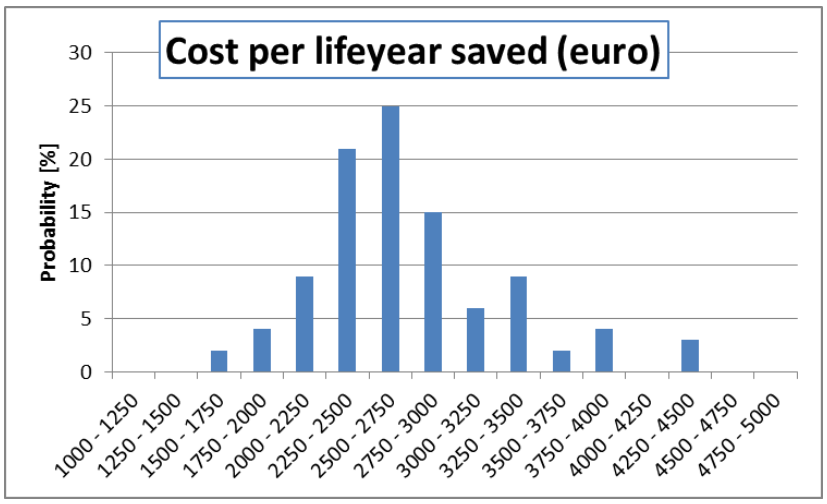
Appendixes

Appendix A. ICER and CEAC curves of scenarios

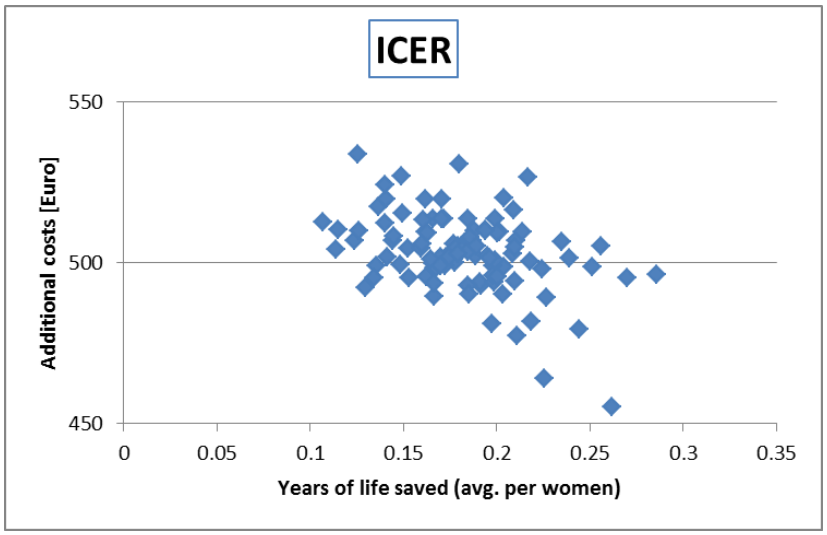
Scenario 40-60-2

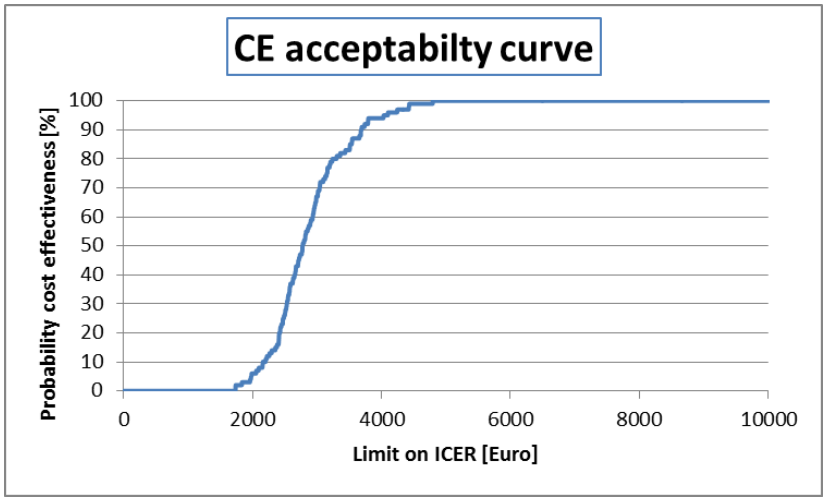


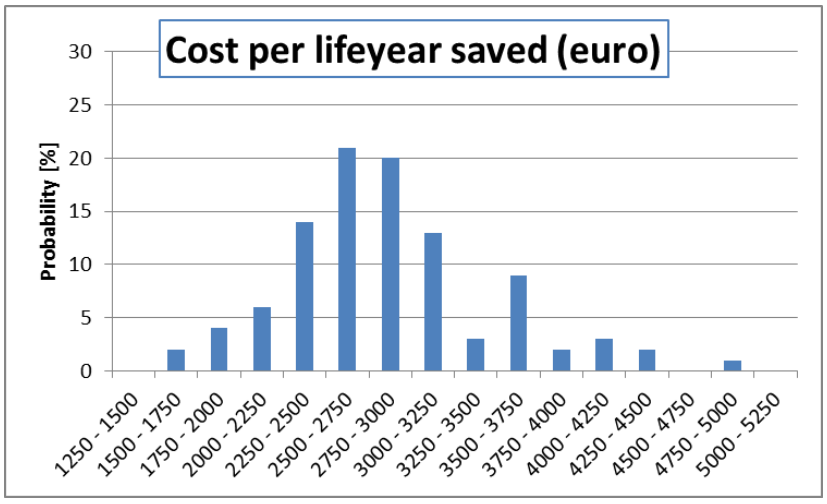




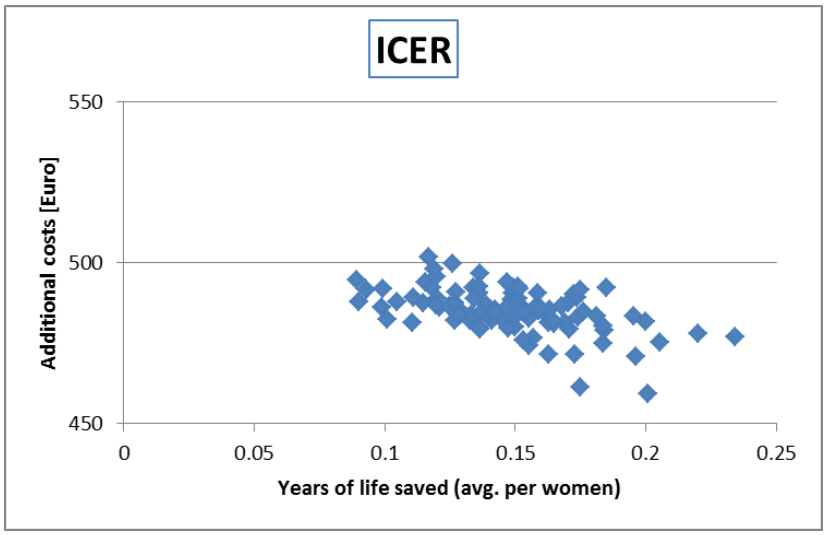
Scenario 30-60-2

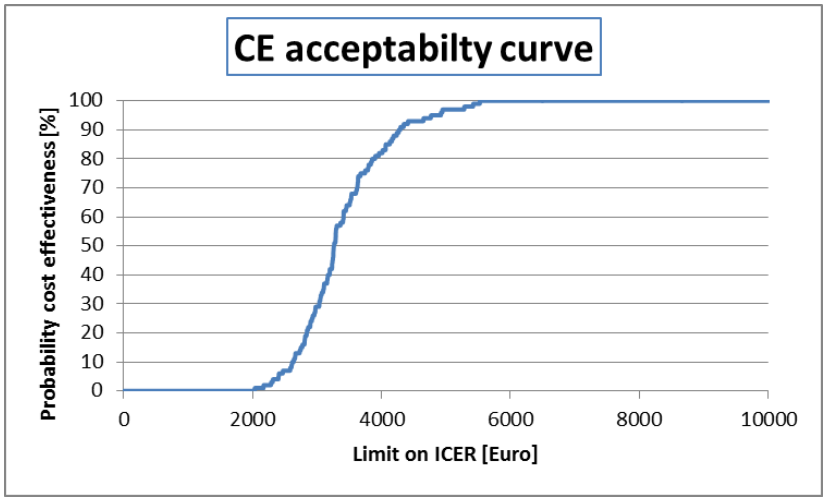


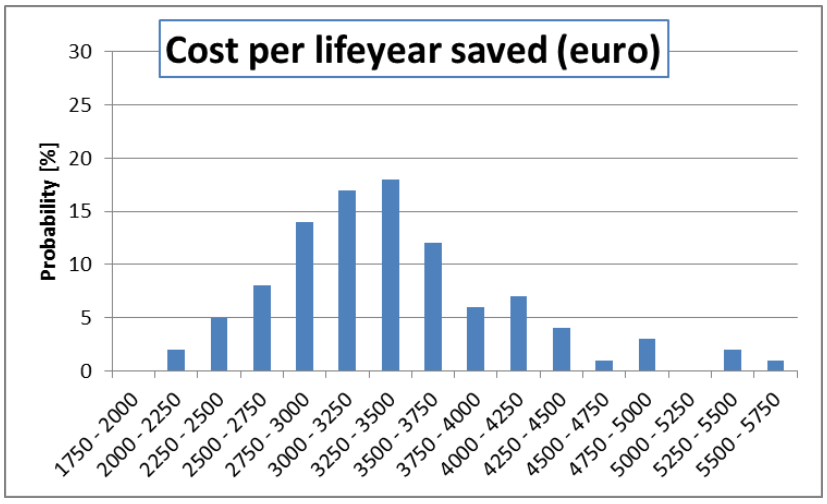




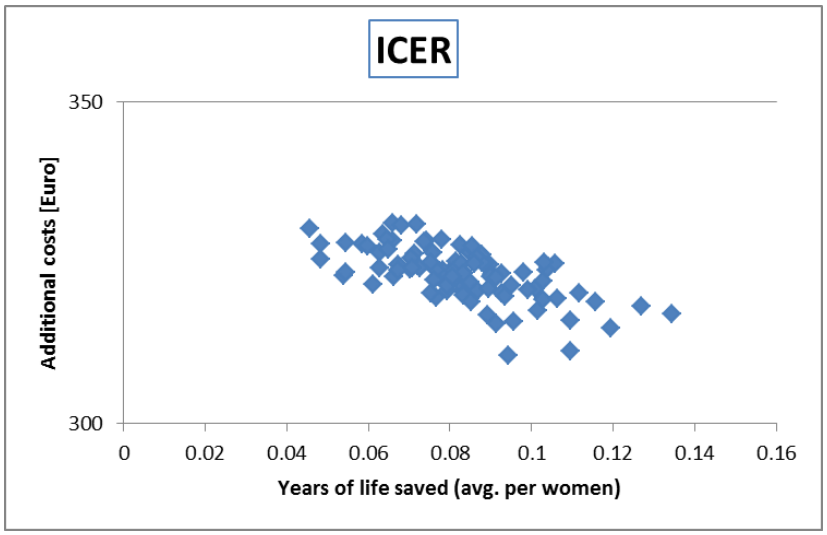
Scenario 40-70-2

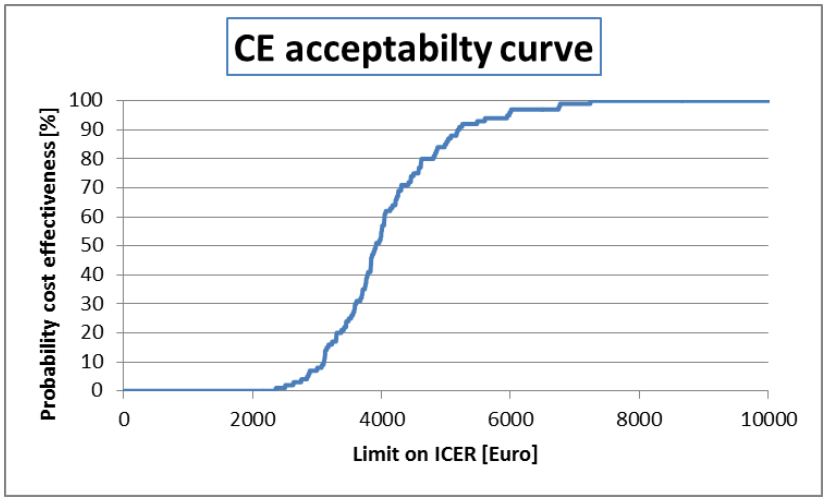


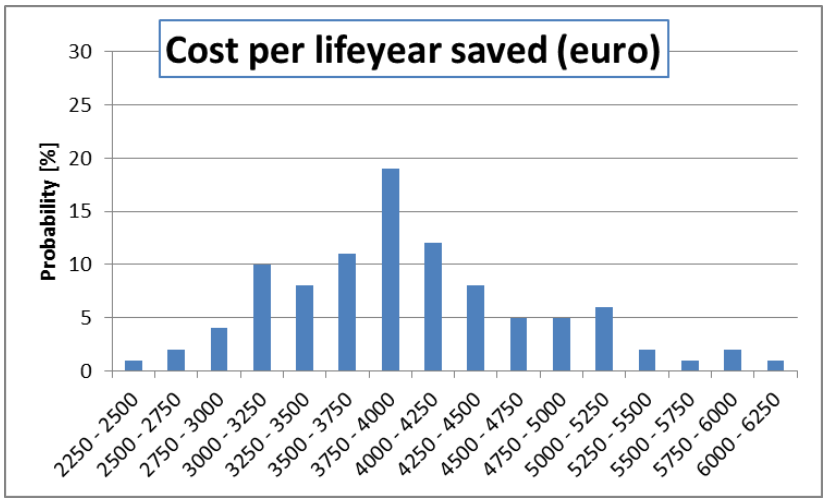




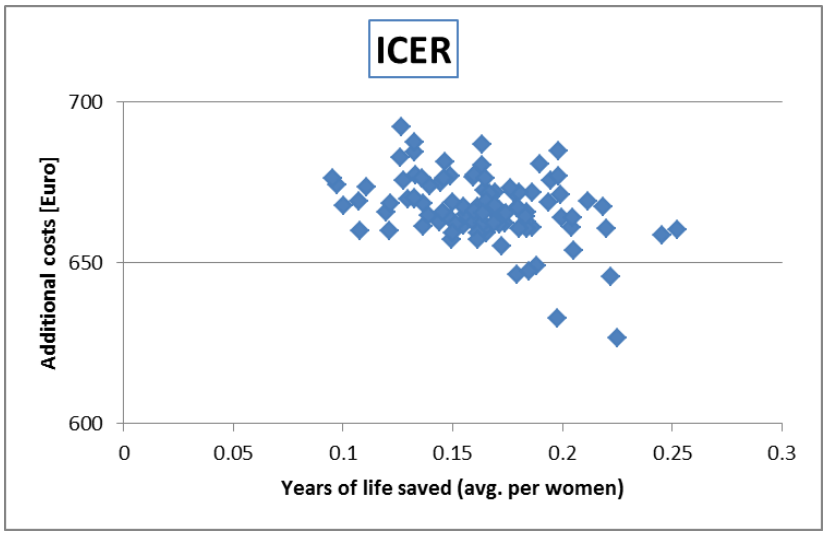
Scenario 50-70-2

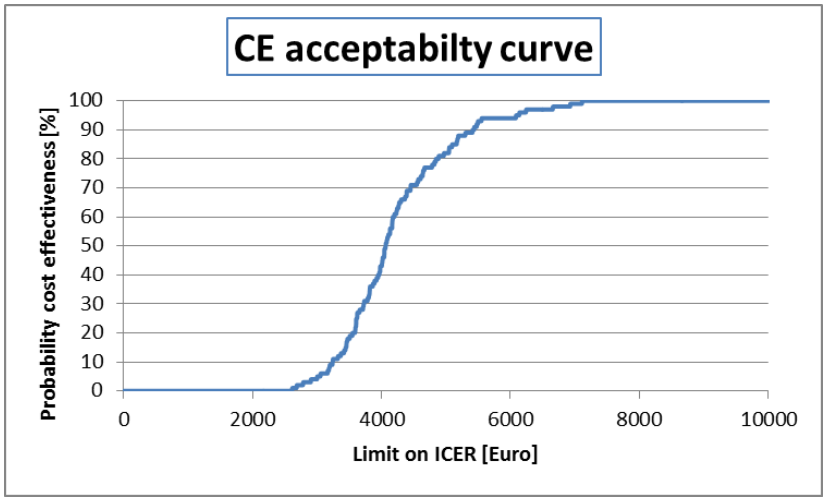


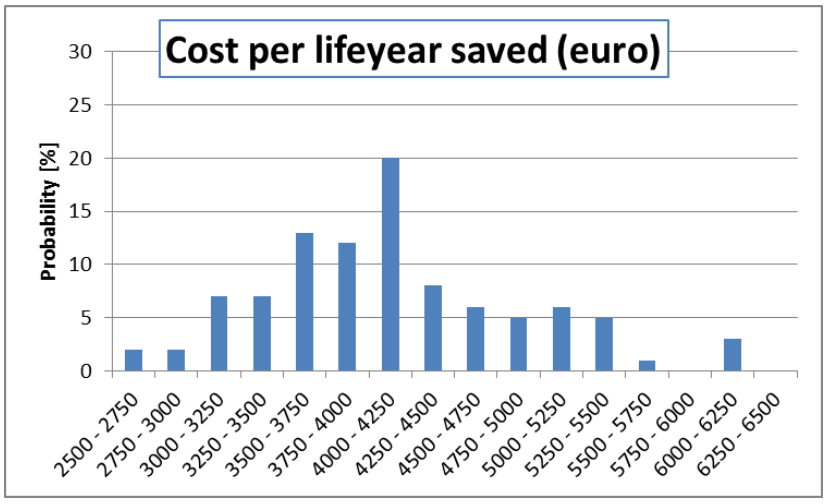




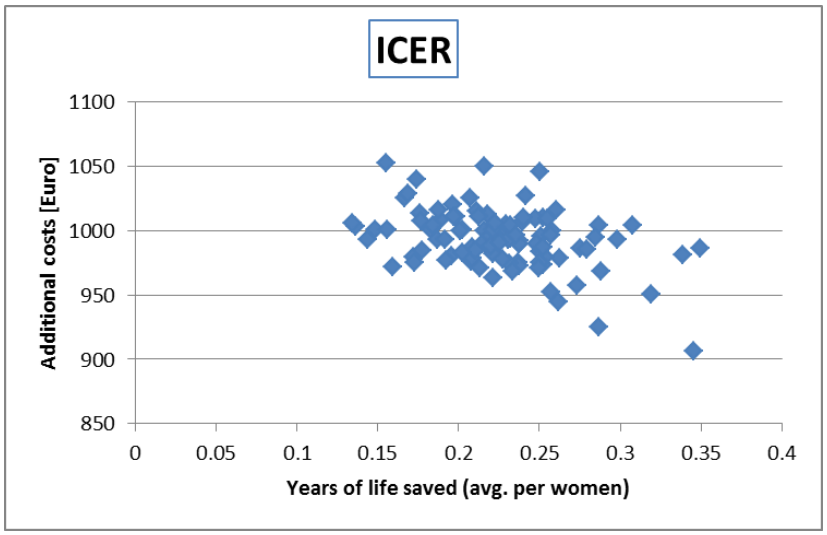
Scenario 40-60-1

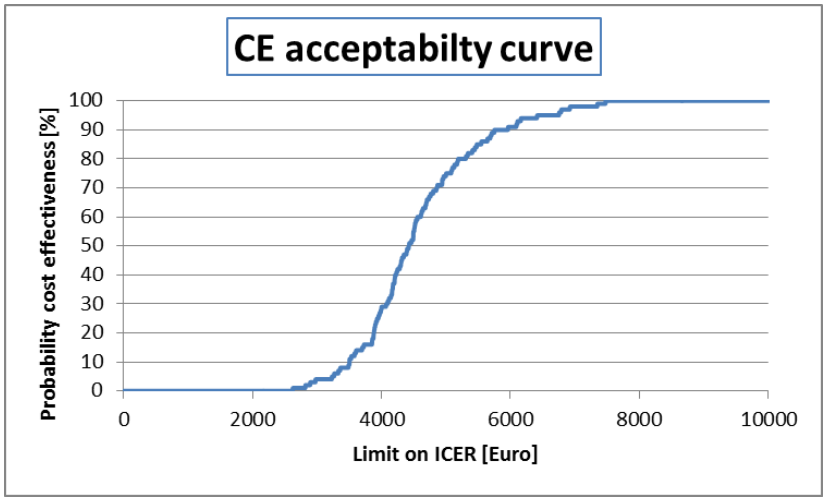


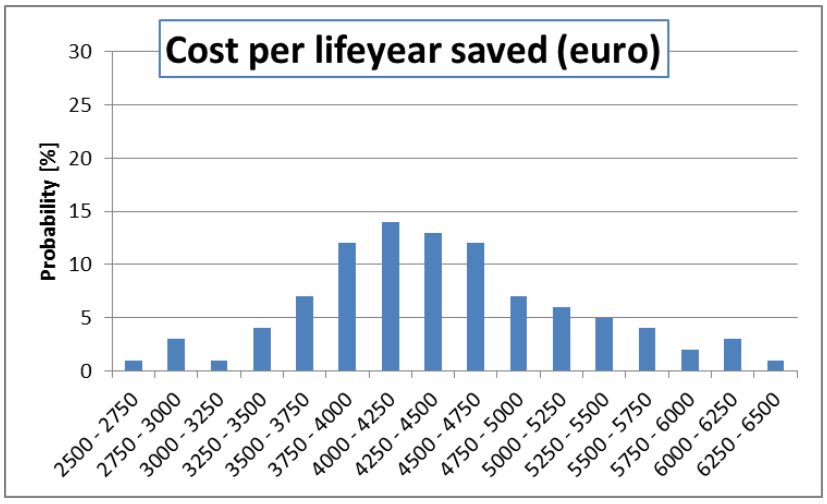




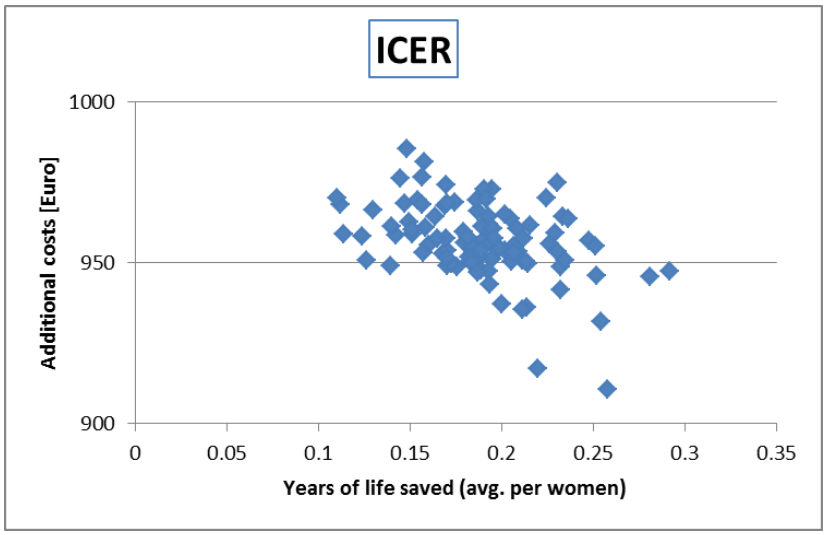
Scenario 30-60-1

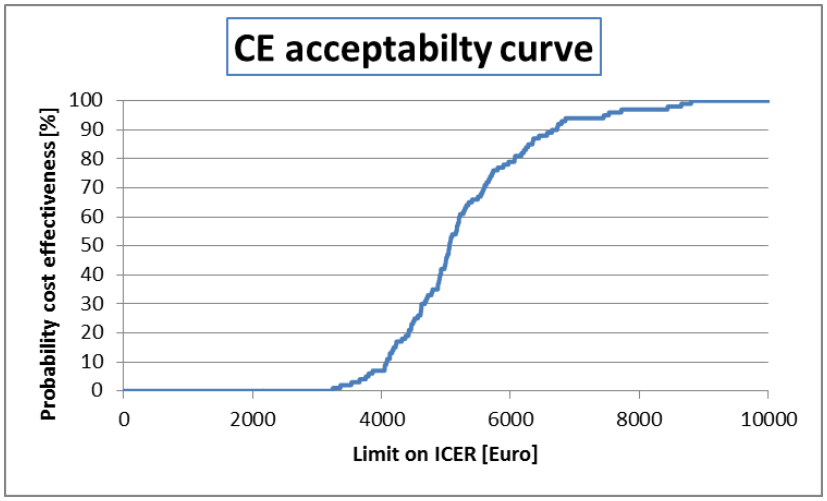


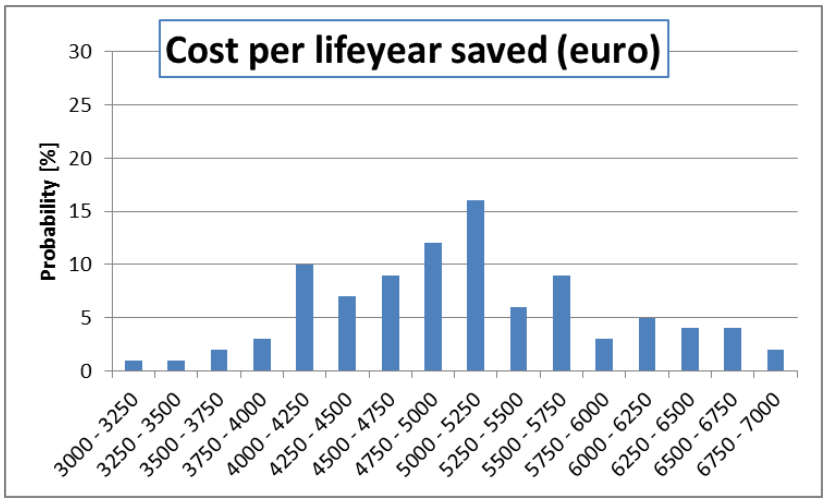




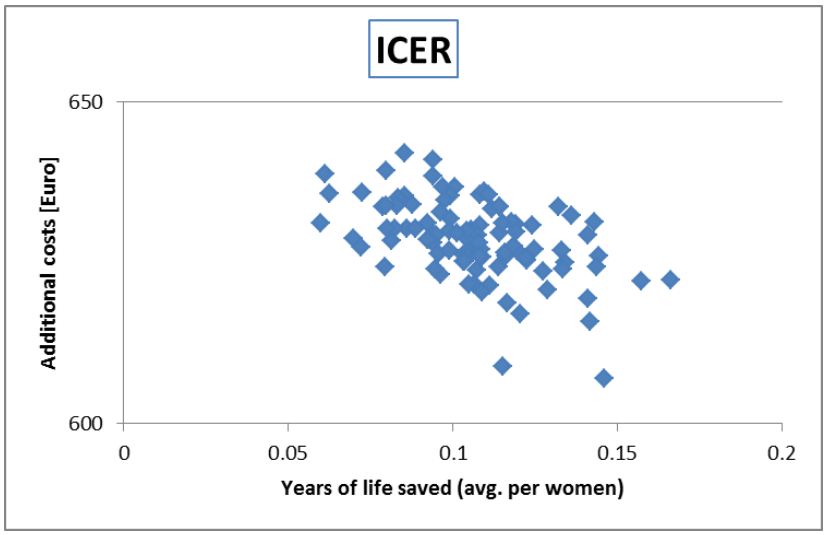
Scenario 40-70-1

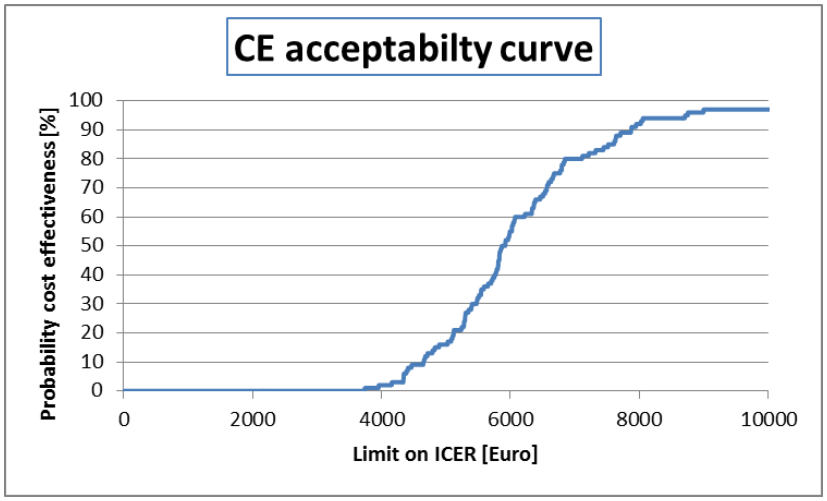


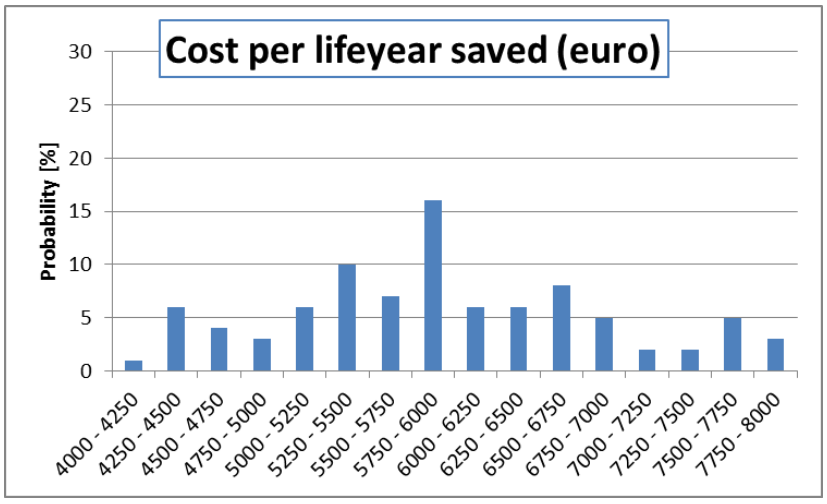




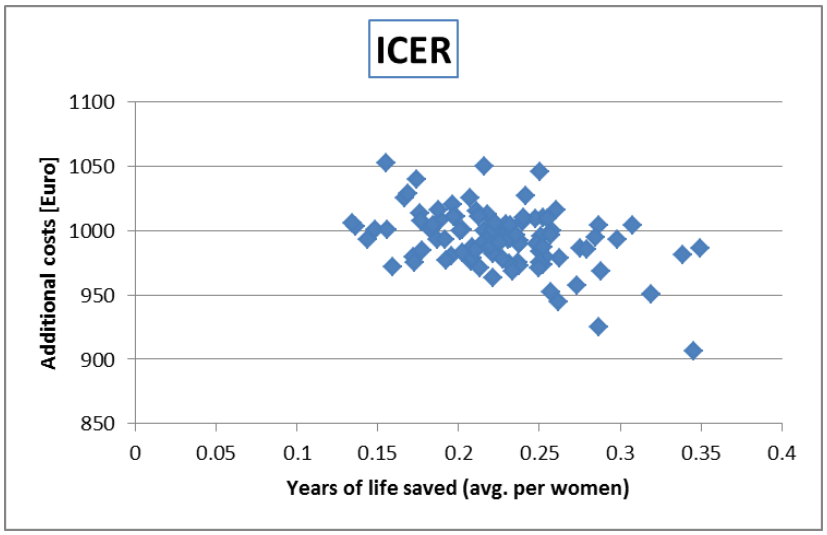
Scenario 50-70-1

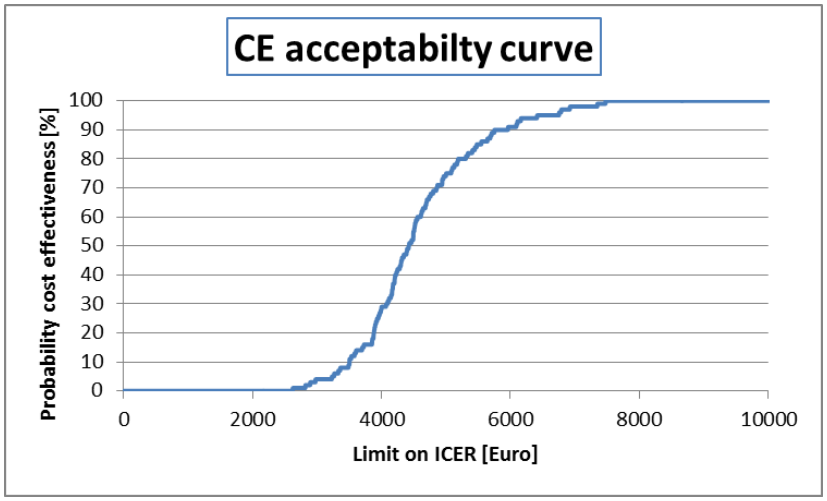


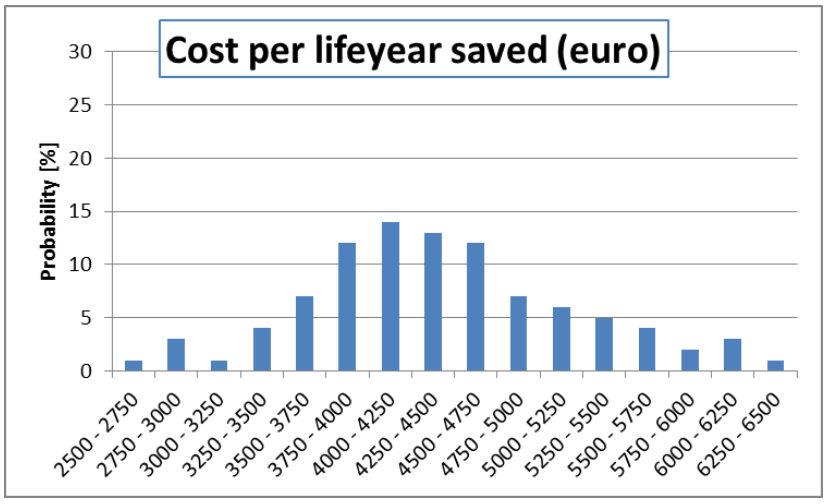




Scenario 30-60-1

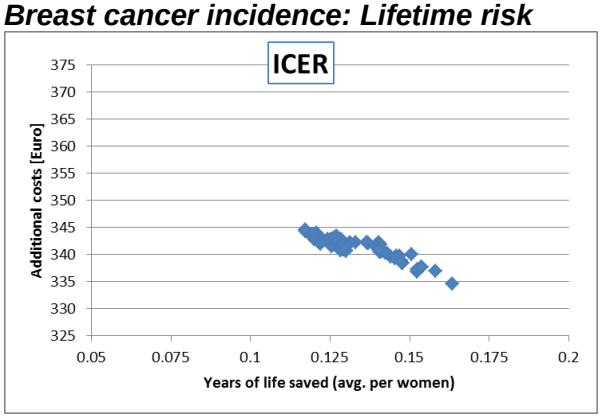


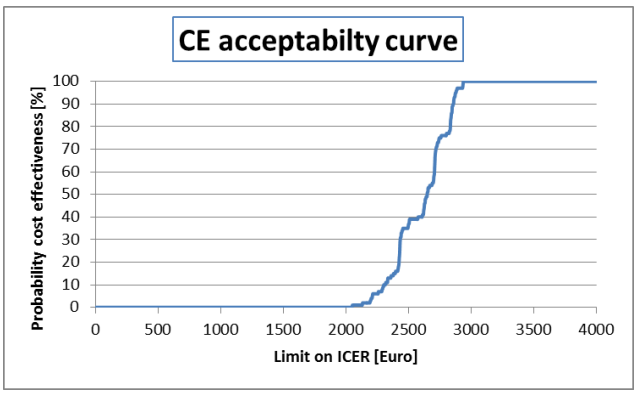




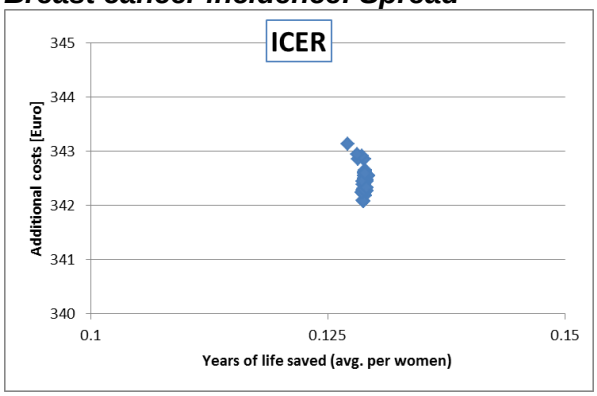
Appendix B. ICER and CEAC curves of input parameters

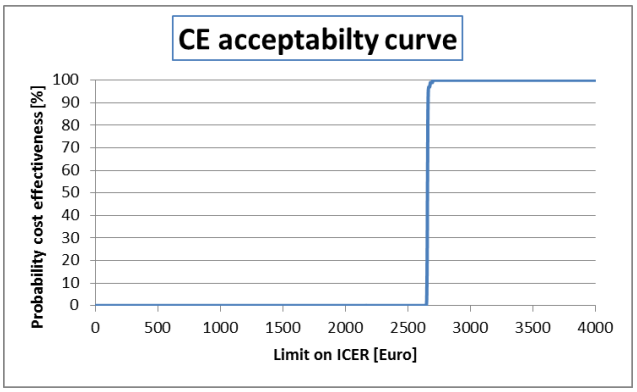
Breast cancer incidence: Lifetime risk



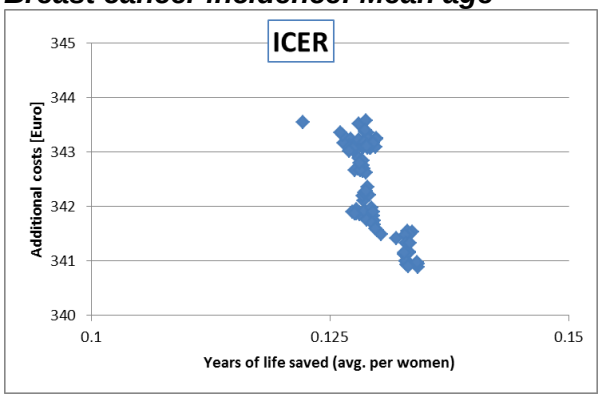


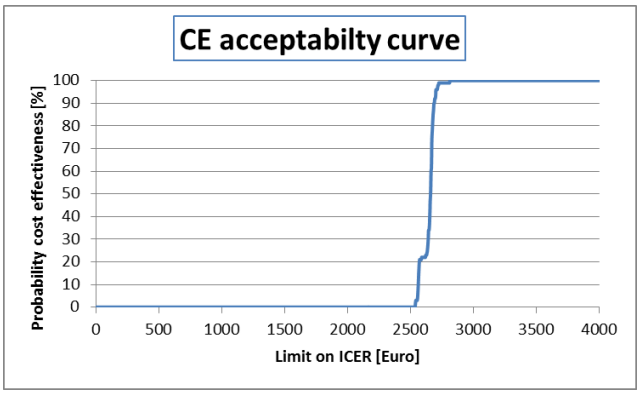
Breast cancer incidence: Spread



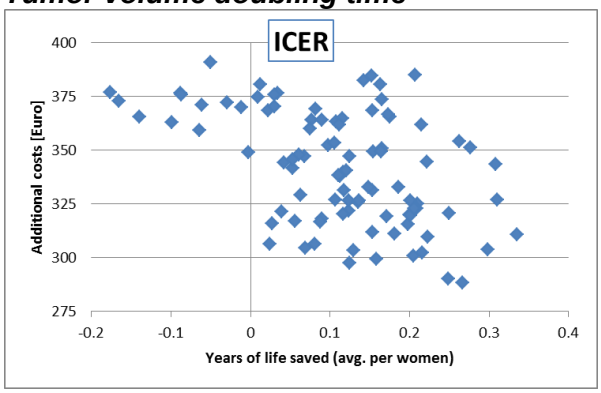


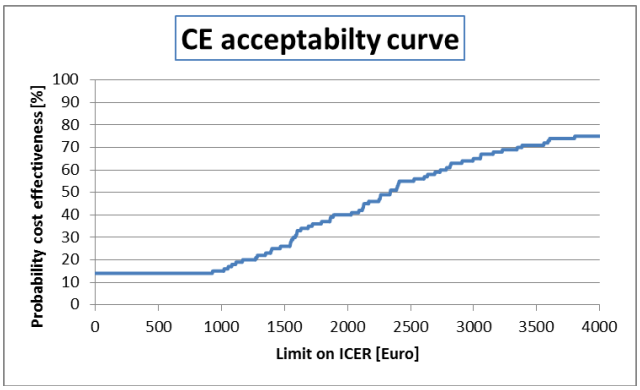
Breast cancer incidence: Mean age



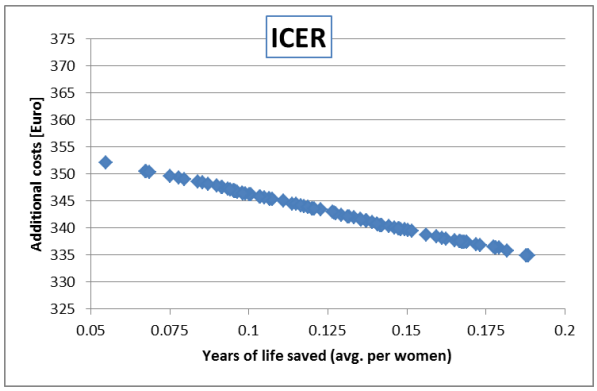


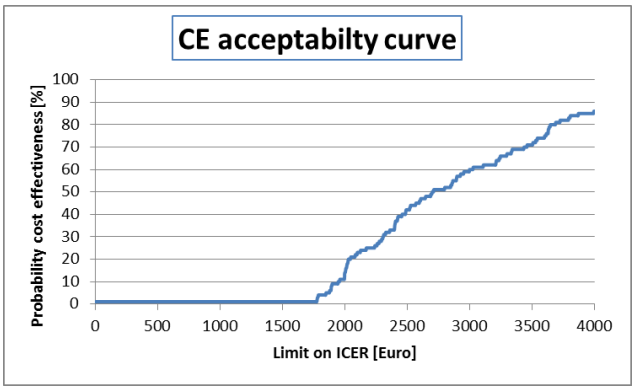
Tumor volume doubling time



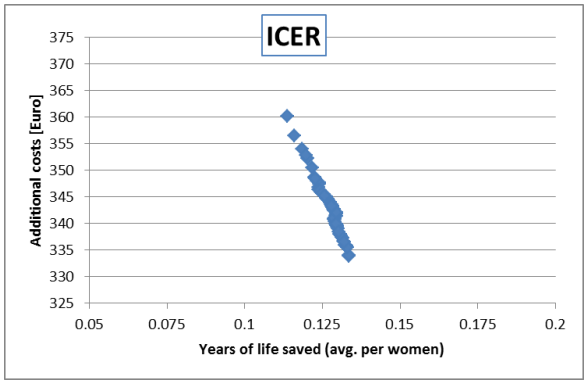


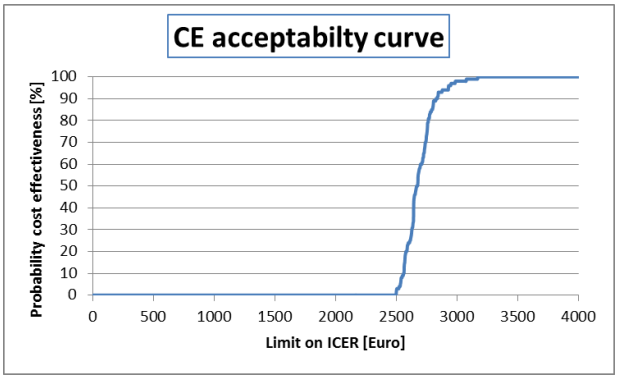
Self-detection diameter



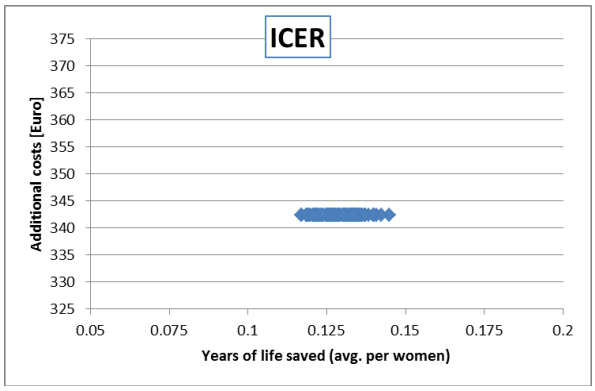


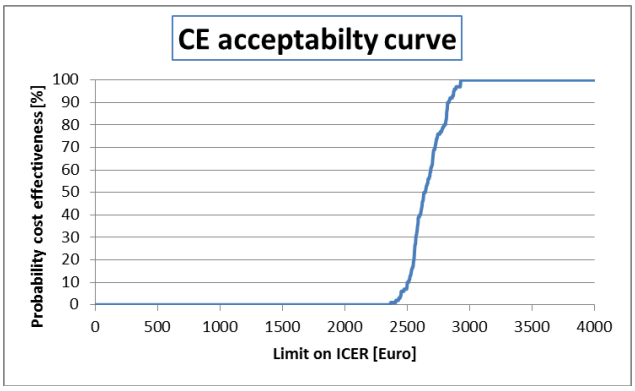
Tumor induction b



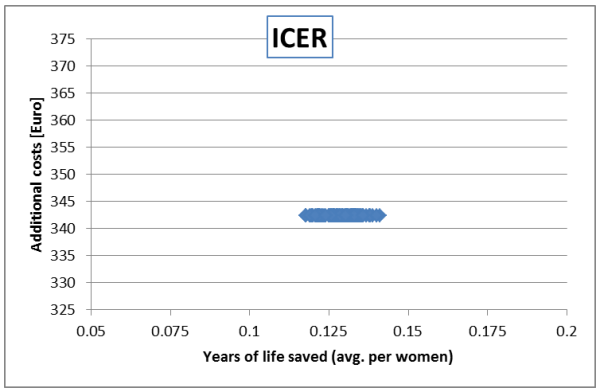


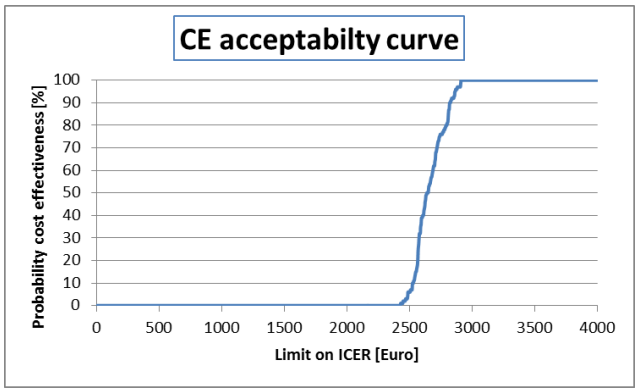
Tumor survival a



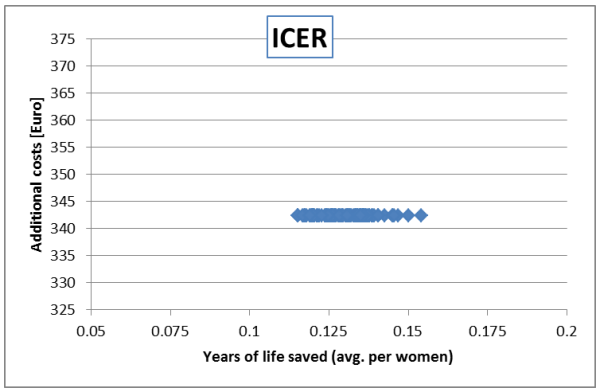


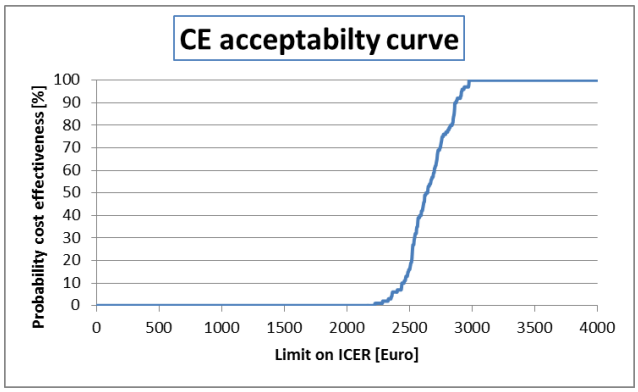
Tumor survival b



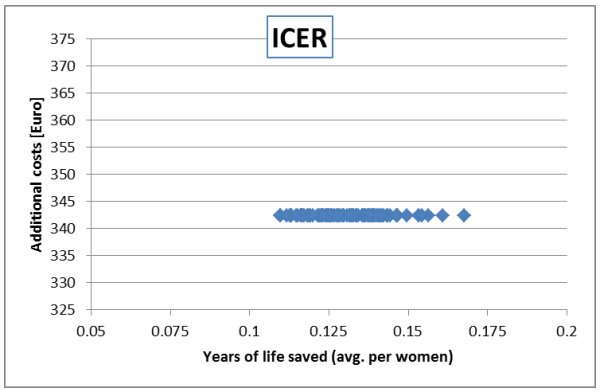


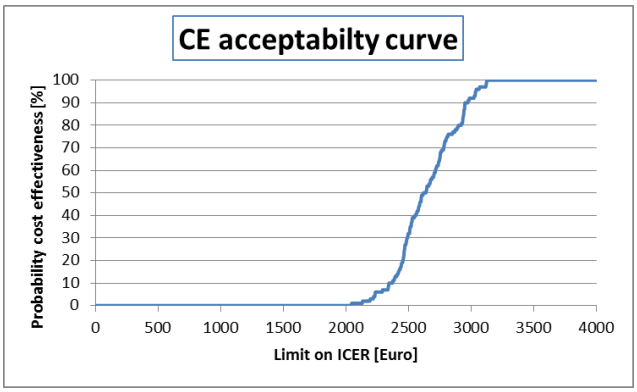
Tumor survival c





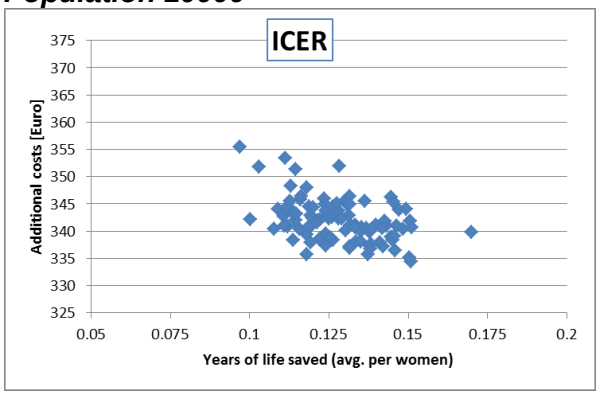
Tumor survival d

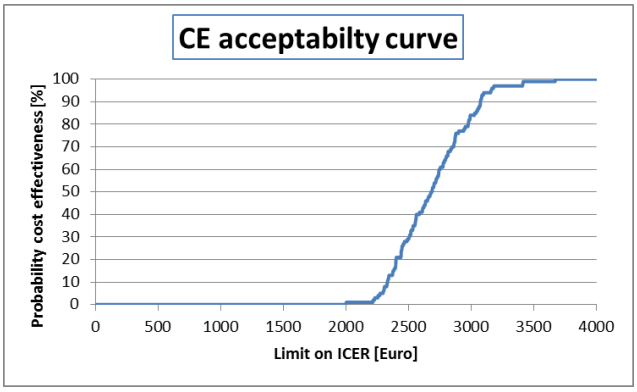




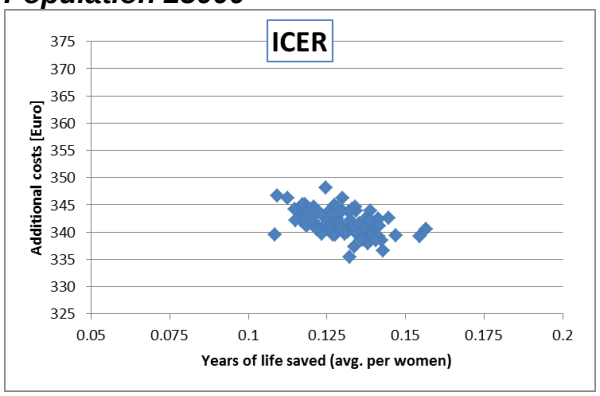
Appendix C. ICER and CEAC curves of population sizes

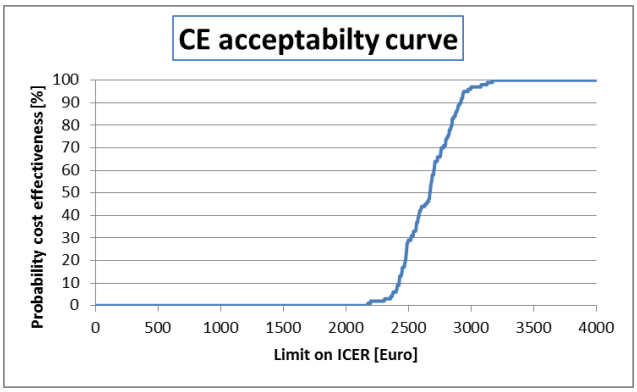
Population 10000



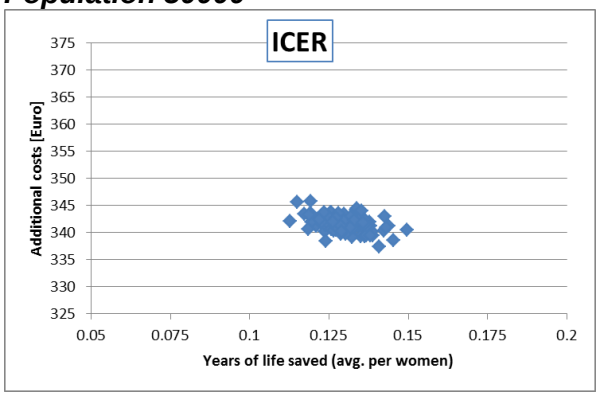


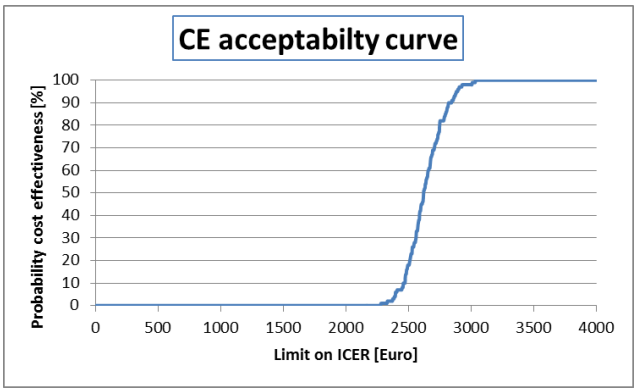
Population 25000



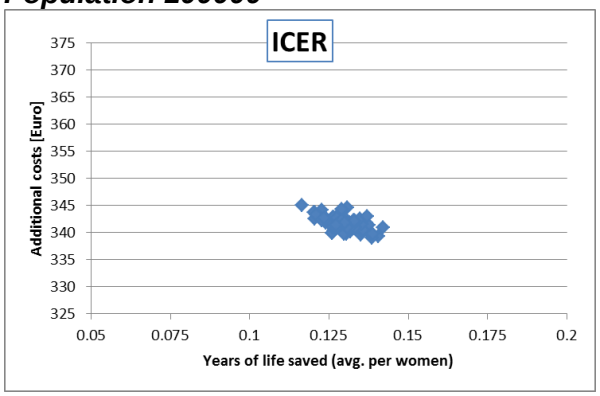


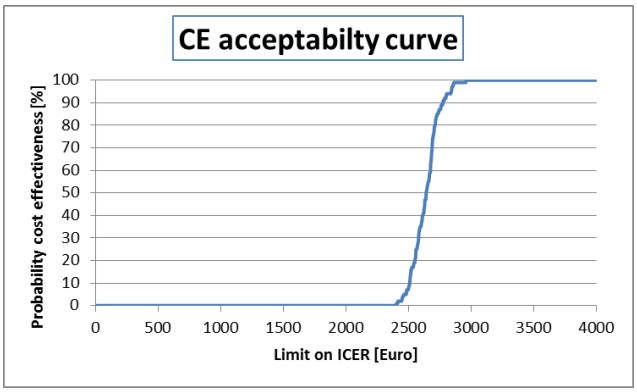
Population 50000





Population 100000





Appendix D. SiMRiSc Flowchart