Erasmus University Rotterdam
Erasmus School of Economics

Master Thesis
Econometrics & Management Science
with specialisation
Operations Research & Quantitative Logistics

A Genetic Algorithm to Personalise Stool-based Colorectal Cancer Screening

Based on measured haemoglobin concentrations

Luuk Arco van Duuren
432941

Supervisor: Dr. R. Spliet
Second assessor: Dr. T.A.B. Dollevoet

External supervisor: Dr. R.G.S. Meester

January 15, 2021

Abstract

Colorectal cancer (CRC) is one of the most frequently diagnosed cancers worldwide. To reduce CRC mortality, many countries implemented a screening programme. Such programmes come with harms and benefits. Currently, the Dutch screening strategy is population-based: it may be optimal to the population but may not be optimal to individuals. Personalised strategies may reduce the harms and increase the benefits of the screening programme even further. Literature has shown that an individual's past measured haemoglobin (Hb) concentrations, obtained by the commonly used stool-based Faecal Immunochemical Tests, predict the risk of developing CRC in the future. Therefore, screening programmes can be personalised based on these values.

This thesis presents a genetic algorithm, combined with properties from Partially Observed Markov Decision Processes (POMDPs), to find suitable, personalised screening strategies, based on someone's Hb concentrations measured in the past. The microsimulation model MISCAN-Colon was used to evaluate the costs and Quality Adjusted Life Years gained by the strategies.

Although the algorithm was stopped before convergence, the current Dutch screening strategy as well as most of the strategies advised in a recent report for the United States Preventive Task Force are dominated by the strategies found by our algorithm. This shows that personalised strategies are likely to improve population-based strategies. However, it is remarkable that all strategies, even the cheapest, prescribe to screen the full population at least once with a colonoscopy, an invasive and expensive test for which screenees need to go to a hospital. Moreover, the obtained policies do not improve when an increased number of Hb concentrations is incorporated in the decision process. Both phenomena suggest that the probability distribution used by MISCAN to simulate Hb concentrations needs further calibration.

Contents

1 Introduction **1**

2 Background Knowledge **3**

2.1 Colorectal cancer and colorectal cancer screening 3

2.2 MISCAN-Colon microsimulation model 4

3 Problem Description **7**

3.1 Colorectal cancer as an unobserved stochastic process 7

3.2 Screening strategies 8

3.3 Costs and effects 8

3.4 Objectives and Pareto dominance 9

3.5 The cost-effective frontier 10

4 Literature Review **12**

5 Methodology **15**

5.1 Belief space 15

5.2 Genetic algorithm 20

5.3 Performance measures 26

6 Results **29**

6.1 Parameter tuning and operator selection 29

6.2 Estimator comparison 31

7 Discussion **41**

7.1 Limitations and suggestions 42

Bibliography **44**

Appendices **49**

A Probability distribution of haemoglobin concentrations 49

B Number of feasible strategies in the belief space instances 52

C Graphs for parameter tuning and operator selection 52

1 Introduction

Currently, colorectal cancer (CRC) is the third most commonly diagnosed cancer worldwide (Schreuders et al., 2015) and is a major cause of death. It is estimated to have caused 881,000 worldwide deaths in 2018 (Bray et al., 2018).

To reduce CRC mortality, many countries implemented a CRC screening programme. Mortality is reduced in two ways. First, cancers are detected at an early stage so treatment is started earlier. Early treatment increases a patient's chances of survival. Second, screening prevents CRC occurrence. Colorectal cancers are preceded by irregularities in the colon called adenomas, which can be removed relatively easily, preventing them to develop into a harmful cancer (Schreuders et al., 2015; Lansdorp-Vogelaar et al., 2011). Several randomized controlled trials have shown that especially detection and removal of adenomas contributes to the effectiveness of CRC screening (Pignone et al., 2002).

However, such screening programmes have drawbacks as well. Screening tests may indicate false positive, leading to unnecessary treatment, or false negative, causing misplaced reassurance and a potentially harsher treatment later in life. Both cases reduce a patient's quality of life and should therefore be prevented. Additionally, screening programmes have financial costs to the society. Policy makers have to find a balance between improving quality of life while containing costs, also referred to as the cost-effectiveness.

To improve this balance, research has set its focus on personalised screening programmes. In such programmes, individuals with a high risk of developing CRC are screened more frequently compared to those with a low risk to minimise the harms and maximise the benefits. The Dutch FIT-based screening programme is well-suited for this. FITs (Faecal Immunochemical Tests) are harmless tests that measure the concentration of haemoglobin in patients' stool samples. Currently, these tests are used qualitatively: if the concentration is above a cut-off, patients are likely to have a adenoma or cancer and they are redirected to a hospital for further investigation. However, the concentrations do not only represent a patient's current status, Grobbee et al. (2017) have shown that these values also predict the risk of developing adenomas and cancers in the future. Therefore they advise to use the quantitative values measured in the past to establish suitable screening intervals for patients. This gives rise to a personalised screening strategy, as opposed to the current Dutch strategy in which the interval is fixed to two years for everyone.

In this thesis, a genetic algorithm is constructed to find cost-effective, personalised screening strategies. These strategies adapt to an individual's quantitative Hb concentrations obtained by FITs in the past. Such strategies use all the information provided by the screening tests and should therefore be more cost-effective than a strategy with a constant screening frequency. The

obtained screening strategies are evaluated using the MISCAN-Colon microsimulation model (Habbema et al., 1985).

This thesis starts with some background information on colorectal cancer, colorectal cancer screening and the MISCAN model in Section 2. A problem description is given in Section 3 and Section 4 describes the current literature on this topic. We elaborate on the properties of the algorithm in Section 5 and results are given in Section 6. Finally, we discuss the contents of this thesis and its strengths and weaknesses in Section 7.

2 Background Knowledge

In this section we discuss the natural history of colorectal cancer, the screening tests used in the Dutch screening programme and we introduce the MISCAN-Colon microsimulation model.

2.1 Colorectal cancer and colorectal cancer screening

Cancers in the colon develop in several stages. Most of them develop from adenomas, which come in three classes (Leslie et al., 2002). At first an adenoma's size is smaller than 5 mm. Once it has grown to 6 mm, it is classified as *non-advanced* and those larger than 10 mm are called *advanced adenomas*[1] (Brenner and Werner, 2017). If left untreated, it may progress into a cancer. At first, this cancer is undetected and it is named *preclinical*. Once detected, either by screening or by the patients themselves due to symptoms, it is classified as a *clinical cancer*. According to Holme et al. (2013) it usually takes about ten years for an adenoma to develop into a cancer. We also introduce the term *lesion*, which refers to both adenomas and cancers.

Adenomas are very common: about 40% of the Western population develops one or more adenomas during their lives. However, not all of them develop into a cancer as only 3% of the Western population actually develops a CRC (Leslie et al., 2002). Therefore we distinguish between *progressive* and *non-progressive adenomas*: the first will develop into a cancer while the latter will not. Note that this is a cause of overtreatment: removal of non-progressive adenomas brings harms and costs while they would not have developed into a cancer.

As mentioned before, CRC screening programmes are especially effective due to the existence of adenomas. They are detectable by screening in an early stage and can be removed without a surgical intervention (Pignone et al., 2002). Therefore, CRC can be prevented against relatively low costs and a small impact on the quality of life. Without screening, progressive adenomas develop into cancers and the patient needs to undergo surgical treatment to have the cancer removed. This intervention is much more expensive and has a much higher impact on the patient's life quality compared to adenoma removal.

It is known that patients which once had an adenoma removed, have a 30-35% of chance of growing new adenomas within 4 years (Leslie et al., 2002). Therefore these patients are brought under surveillance: they are screened more intensively such that their next adenomas are detected at an early stage.

The Dutch CRC screening programme consists of two types of tests. The first is an endoscopic tests, named *colonoscopy*, which is the backbone of the programme. With a colonoscopy, the

[1] Multiple defining properties of advanced adenomas exist and are described by Brenner and Werner (2017). In our model we characterise advanced adenoma by its size only.

colon and rectum are visualised and present lesions are observed by a surgeon directly. Found adenomas are removed immediately. Since the colon and rectum are visualised, it is a nearly perfect CRC screening tests. However, an important disadvantage of colonoscopies is that they are very expensive and invasive. A patient needs to visit the hospital, needs anaesthesia and a clinician is needed to perform and observe the colon.

Therefore the screening programme has been complemented with a less invasive and cheaper stool test, the Faecal Immunochemical Test (FIT). Robertson et al. (2017) and Day et al. (2013) write that this test has proven itself to be one of the least invasive and least expensive CRC screening tests.

FITs work as follows. Advancing lesions cause an increased concentration of haemoglobin (Hb) in a patient's stool sample. FITs measure this concentration, indicating the risk of having an adenoma or cancer. Currently, the test is scheduled biennially and used qualitatively in The Netherlands: if the Hb concentration is above the cut-off of 47μg haemoglobin (Hb) per gram faeces, the result is positive and the patient is redirected to a hospital for a colonoscopy. If the colon turns out to be clean, the FIT was false positive and the patient is not invited for a FIT for ten years. If an adenoma is found, the patient is brought under surveillance and a preclinical cancer leads to a treatment trajectory. In contrast, if the Hb concentration is below the cut-off (negative), a new FIT is scheduled in two years. This interval is pre-determined for every individual.

Note that the measured Hb concentration is only used qualitatively. However, Grobbee et al. (2017) show that the quantitative Hb value is predictive for the risk of developing a cancer in the future and should be used to establish a suitable interval. This relation is illustrated in Figure 2.1. The two axes show the concentrations found in two consecutive FIT rounds preceding a colonoscopy. The colour indicates the corresponding risk of finding a lesion with the colonoscopy. The map shows that the risk quickly increases for increasing Hb concentrations. Therefore it is is reasonable to adapt screening intervals to Hb values measured in the past.

2.2 MISCAN-Colon microsimulation model

Cancer screening has harms and benefits. While saving lives, it also brings financial costs, the impact of undergoing a colonoscopy and the harms of wrong test results. Day et al. (2013) enumerate more disadvantages of screening programmes. Therefore some screening strategies are better than others. However, comparing different programmes requires large scale trials in which large groups of patients follow different screening strategies. For such a procedure, a lot of people, resources and time are needed.

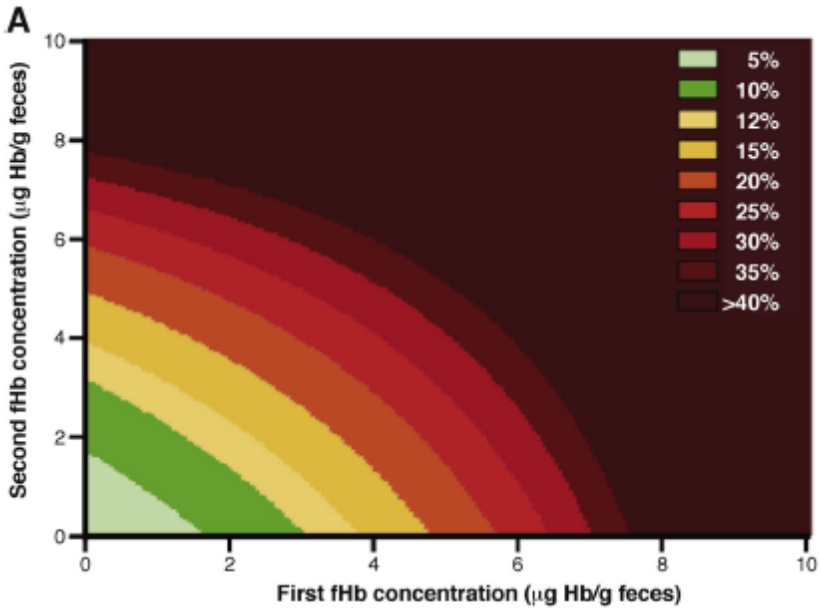

Figure 2.1: Heat map of the risk of finding a lesion with a colonoscopy, given the Hb concentrations obtained by two preceding (negative) FITs. The FIT result is negative if the concentration is less than 10μg/g. This plot illustrates that the risk increases with increasing Hb concentrations. Adapted from Grobbee et al. (2017)

To overcome these hurdles, the department of Public Health of the Erasmus MC Rotterdam set up the MISCAN model (Microsimulation Screening Analysis) in 1985, first described by Habbema et al. (1985). MISCAN simulates two parallel universes with the same population at an individual level. Part of the population develops a disease and needs to be treated. Screening is applied in only one of the universes, the other does not incorporate screening. Therefore some cases of the disease are prevented in the first universe. Since the simulation uses the same population in both universes, the costs and life years gained can be compared between the universes. This is how different screening strategies can be compared and the cost-effective one can be chosen without expensive clinical trials.

The model needs three layers of input. The first layer of parameters concerns the simulated population such as birth and life tables. The second layer concerns disease-related properties such as the risk of developing the disease, the progression speed of the cancer and how quickly someone gets symptoms. Finally it needs the properties of the screening programme as an input, such as screening frequency and sensitivities and specificities of the used tests.

MISCAN-Colon is the version that is especially adapted for colorectal cancer (Loeve et al., 1999). Each individual may develop one or multiple lesions. Each of them is a separate, continuous-time Semi-Markov process of which the discrete state space is shown in fig. 2.2. The state space nearly corresponds with the disease progression as described in section 2.1.

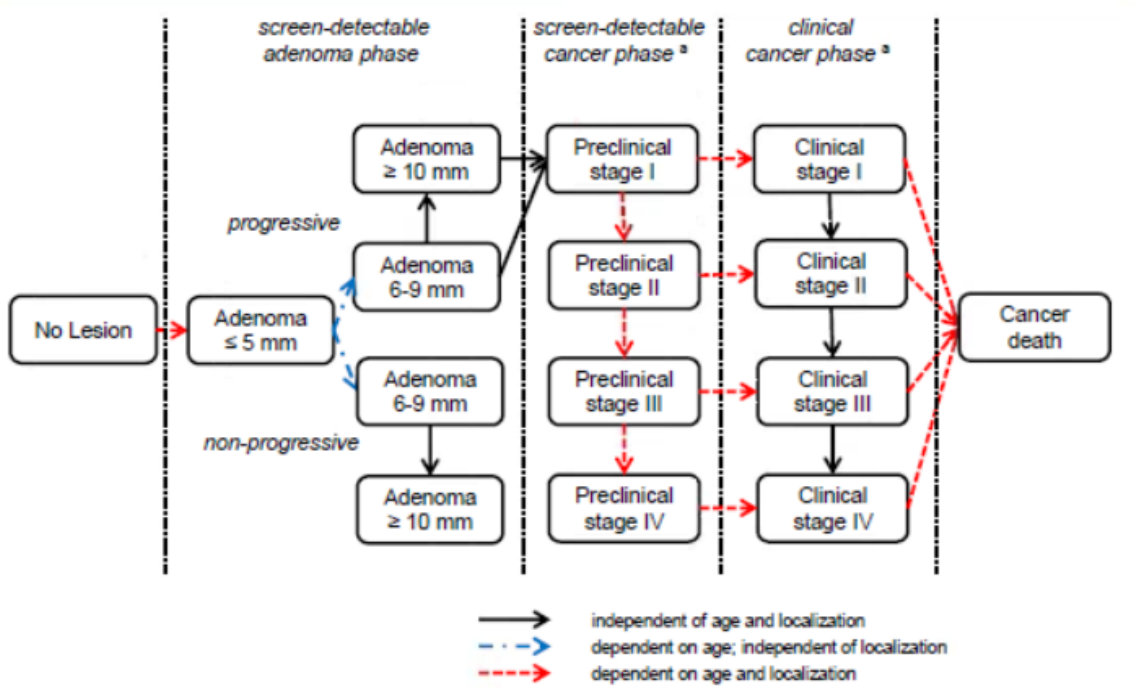

Figure 2.2: A detailed overview of the state space used in MISCAN-Colon to model the progression of one lesion. As shown in the figure, some transitions are dependent on age and possibly on localization. Most transitions are correlated. Adapted from Gini (2020)

Several features of the state space are noteworthy with respect to this thesis. First, lesions smaller than 10 mm are ignored in this thesis as they are undetected by the FIT. Second, some transitions are age-dependent and/or localisation-dependent: older patients and certain locations in the colon are more susceptible to developing a lesion. The localisation of lesion is also ignored in our algorithm since it cannot be derived from measured Hb values. Finally, some transition rates are highly correlated. For example, the transitions between the adenoma states: a small adenoma that progresses to non-advanced quickly will also quickly become advanced.

The number of lesions in an individual is determined by an individual risk factor, which is a random variable drawn at birth. Since each lesion is modelled as a separate stochastic process, a patient's state is characterised by multiple processes. In fact, the disease state of a patient is defined as the state of its most advanced lesion: the process of the patient is the 'maximum' of several Semi-Markov processes.

3 Problem Description

In this section we define the decision problem discussed in this thesis. First the assumed stochastic process for the cancer development is described. Next we consider the decisions that are to be optimised and we discuss the corresponding cost structure. Finally we define the objective function and our measure of optimality.

3.1 Colorectal cancer as an unobserved stochastic process

In our problem, we use a simplified state space $\mathcal{X}$ shown in Figure 3.1. It corresponds with Figure 2.2 as follows: the state Healthy (abbreviated by H) contains 'No lesion' and adenomas smaller than 9 mm. The Adenoma state (A) comprises advanced adenomas. The (Pre-)clinical Cancer (PC and CC resp.) states represent the four (pre-)clinical cancer stages in Figure 2.2. $\mathcal{X}$ also includes the states Cancer Death (CD) and Other-cause Death (OD). A CD occurs if treatment of a CC is unsuccessful, otherwise the death of a patient is labelled as an OD. The bold arrows show natural transitions, whereas the blue, dashed arrows indicate transitions induced by a colonoscopy. The disease of a patient is modelled by a continuous-time stochastic process $\{X_t \in \mathcal{X}, t \in [0, 100]\}$ with t the age of the patient.

$\mathcal{X}$ has two types of states: observed and unobserved states, distinguished by the shade in Figure 3.1. We assume that we can always observe whether a patient is in one of the unshaded states OD, CD and CC. In contrast, we need screening tests to get an indication of a patient's state if it is in one of the shaded states as the lesions of a patient are not directly observed. Where a colonoscopy indicates the patient's state with (near) certainty, a FIT only gives an indication. $\mathcal{X}^u \subset \mathcal{X}$ is defined as the set containing the unobserved states.

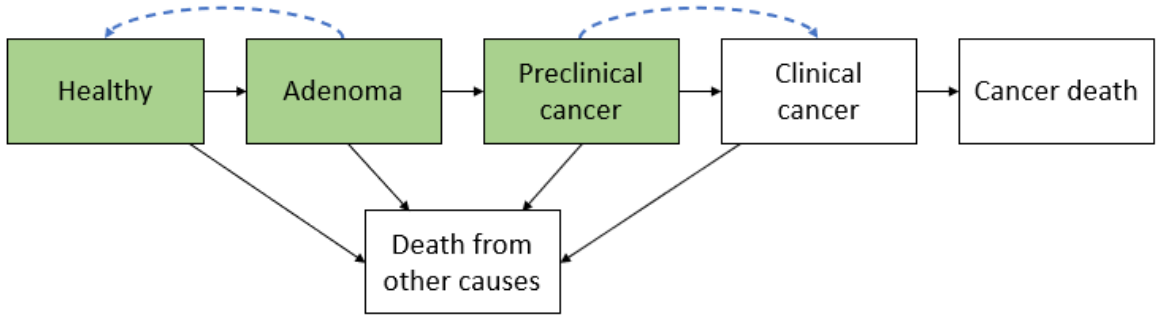

Figure 3.1: A simplified overview of the health states and possible transitions of CRC. The shaded states are unobserved and make up the set $\mathcal{X}^u$. The solid arrows indicate natural transitions while the dashed arrows indicate transitions induced by colonoscopies. Adapted from Zauber et al. (2009)

3.2 Screening strategies

FITs and colonoscopies are conducted according to a screening policy. This section defines the concept of a screening policy and introduces the corresponding notation.

Not everyone is eligible for CRC screening. The set $\mathcal{T} := \{40, 41, \ldots, 85\}$ contains the ages that the population is screen-eligible. Individuals do not undergo a test at every age in $\mathcal{T}$. $\tau_n \in \mathcal{T}$ denotes the age at which a patient undergoes his n^{th} FIT. Note that $\tau_{n-1} < \tau_n$ for all n. This test results in a quantitative Hb concentration S_n. S_n is a random variable and its distribution may depend on $S_1, \ldots, S_{n-1}$ and the disease state at the n^{th} measurement $X_n := X_{\tau_n}$. The currently used (unvalidated) model for the concentrations is described in Appendix A.

After each measurement, we choose a suitable action, based on the sequence of Hb concentrations $S_1, \ldots, S_n$. The first option is to send the patient to a hospital for a colonoscopy such that X_n is observed with near-certainty. This action corresponds with a positive test result and should be chosen for patients with high-risk Hb values. The remaining actions correspond with a negative test result: the time interval $I_n = \tau_{n+1} - \tau_n$ until the next FIT is chosen. The set of possible intervals $\mathcal{I}$ is a discrete, finite set. These actions are denoted by $\text{FIT}_I, I \in \mathcal{I}$. We define the complete action set by $\mathcal{A} := \{\text{FIT}_I, I \in \mathcal{I}\} \cup \{\text{COL}\}$.

A policy specifies what action to take, given the past observations and a patient's age. Therefore we can define a *policy* as a function $\pi : (S_1, \ldots, S_n, \tau_n) \mapsto a_n$ that maps the past observations and the current age to an action $a_n \in \mathcal{A}$. π is also referred to as a *screening strategy*.

Strategies can also be assigned to a part of the population. As a consequence, a strategy can be formed by assigning strategy π to a fraction w of the population and a strategy σ to the other part. In general we state that a convex combination of two screening strategies form a valid new strategy.

3.3 Costs and effects

Costs and effects are associated with the states of a patient and with taken actions. The costs are expressed in US dollars. No costs are made if a patient is healthy, has an undetected adenoma or a preclinical cancer (the states in $\mathcal{X}^u$). Similarly, no more costs are made if a patient has died due to either cause. Only once a patient has a clinical cancer, costs for treatment are charged. This is a fixed yearly amount which depends on the cancer stage the patient is in. Additionally, each FIT and colonoscopy comes with a fixed cost.

The effects are measured in terms of Quality Adjusted Life Years (QALYs). One QALY represents the value of a perfect life year without any complications. Therefore, patients are

rewarded one QALY for every year they are in one of the unobserved states $\mathcal{X}^u$. This yearly reward reduces to a value between zero and one for years with cancer treatment. The burdens of FITs and colonoscopies are represented by a fixed reduction in QALYs per FIT and colonoscopy.

The costs and effects of a policy on a patient are calculated by accumulating all the costs charged and QALYs obtained throughout one's life. The total costs and effects of a policy π are calculated by averaging the costs and effects of all individuals in the population. We denote these quantities by $K(\pi)$ and $E(\pi)$ and they are measured in \$/individual and QALY/individual respectively. It is impossible to calculate the costs and effects of a policy exactly due to the complex nature of the cancer and screening processes: they have to be simulated with MISCAN.

3.4 Objectives and Pareto dominance

Having defined all elements of the stochastic process, we formulate the optimisation problem. In our problem we aim to find strategies π that maximise the effects and minimise the costs. Therefore we formulate the problem as a bi-objective optimisation problem:

$$\max_{\pi \in \Pi} \begin{bmatrix} E(\pi) \\ -K(\pi) \end{bmatrix} \tag{3.1}$$

in which $K(\pi), E(\pi)$ are the costs and effects of a strategy π. Π represents the set containing all possible screening strategies. Often, in the setting of multi-objective optimisation, the concept of *Pareto dominance*, or *strong dominance*, is used:

Definition 3.1. *Let $\pi, \sigma \in \Pi$ be two screening strategies. π strongly dominates σ, denoted by $\pi \succ \sigma$, if and only if the two following properties hold:*

1. $E(\pi) \geq E(\sigma) \wedge K(\pi) \leq K(\sigma)$
2. $E(\pi) > E(\sigma) \vee K(\pi) < K(\sigma)$

Intuitively, strategy π dominates σ if (1) the values of the costs and effects of π are at least as good as those of σ and (2) at least one of the objectives is better. In that case strategy π is a better choice than strategy σ. The concept of strong dominance is illustrated in fig. 3.2 in which the objectives of a finite number of screening strategies are plotted in the cost-effectiveness space. Strategy E is strongly dominated by G as the effects of G are higher and the costs are lower. This is denoted by G $\succ$ E. In contrast, strategies B and E do not dominate each other: B scores better in terms of costs while E does better in terms of effects. Therefore we write B $\not\succ$ E and E $\not\succ$ B.

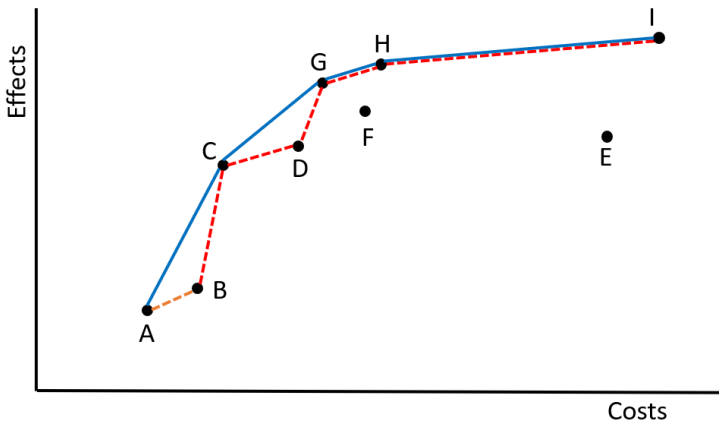

Figure 3.2: The costs and effects of screening policies A through I are plotted in the cost-effectiveness space. The red, dashed line represents the Pareto optimal frontier, the blue line represents the cost-effective frontier.

Since not all strategies are dominated by another, the solution of a multi-objective optimisation problem is represented by a set of solutions. In this set, none of the solutions are dominated by another solution in the set. As a consequence, the effects increase for increasing costs. To formalise this, Zitzler et al. (2003) define the notion of a *(strong) approximation set.*

Definition 3.2. *The strong approximation set* $\psi(S)$ *of a set* $S \subseteq \Pi$ *is characterised by*

$$\psi(S) := \{\pi \in S : \nexists \sigma \in S : \sigma \succ \pi.\}$$

Given a reference set S, its approximation set $\psi(S)$ contains all strategies in S that are not dominated by any other strategy in S: it contains the best solutions of the reference set. This leads to the formalisation of the *Pareto optimal frontier* (POF) $\psi^* := \psi(\Pi)$ as the set that contains all non-dominated solutions of the full solution space Π. This set is the optimal solution to the optimisation problem in Equation (3.1) with respect to Pareto dominance. The red line in Figure 3.2 represents the Pareto optimal frontier of the set containing strategies A through I.

3.5 The cost-effective frontier

Recall that convex combinations of policies are also valid policies. Pareto/strong dominance does not account for this. Figure 3.2 shows how the Pareto optimal frontier in red does includes strategies B and D as they are not dominated by another strategy. However they are dominated by a convex combination of A and C and C and G respectively. Therefore we introduce the concept of *weak dominance* as formalised by Dunnewind (2020):

Definition 3.3. *Let $\pi \in S \subseteq \Pi$ be a screening strategies. π is **not** weakly dominated with respect to S if and only if $\forall \sigma, \rho \in S$ and $\forall w \in [0,1]$ the two following properties hold:*

1. $E(\pi) \geq wE(\sigma) + (1-w)E(\rho) \wedge K(\pi) \leq wK(\sigma) + (1-w)K(\rho)$
2. $E(\pi) > wE(\sigma) + (1-w)E(\rho) \vee K(\pi) < wK(\sigma) + (1-w)K(\rho)$

Intuitively, a policy is weakly dominated if it is strongly dominated by a convex combination of two other strategies. Therefore strong dominance implies weak dominance. Similar to strong approximation sets, we also define *weak approximation sets*.

Definition 3.4. *The weak approximation set $\xi(S)$ of a set $S \subseteq \Pi$ contains all policies $\pi \in S$ that are not weakly dominated with respect to S.*

Note that $\xi(S) \subseteq \psi(S)$. Now we can define the *cost-effectiveness frontier* (CEF) ξ^* by the set $\xi(\Pi)$. This set is the optimal solution to the optimisation problem 3.1 with respect to weak dominance. The blue line in Figure 3.2 represents the cost-effectiveness frontier of the set containing strategies A through I.

Our aim is to approximate the cost-effective frontier of Π as closely as possible, such that policy makers can choose their preferred cost-effective screening strategy.

4 Literature Review

Several mathematical models for evaluating and generating cancer screening strategies are discussed in the literature due to its societal relevance and the numerous different types of cancers. The literature discusses three types of screening strategies: (i) periodic, population-based strategies; (ii) aperiodic, population-based strategies and (iii) individual strategies.

The first category assumes that screening takes place at a regular time interval, applied to the full population. The interval is independent of the individuals' previous test results or their risk factors. Examples of such methods are described by Frazier et al. (2000); Preston and Smith (2001); Clemen and Lacke (2001). While such population-based strategies may work well for the average population, they are not necessarily optimal to individuals. Furthermore, such strategies are often found using expert opinions: experts come up with strategies that could potentially work well and only these strategies are evaluated by the model. Out of all evaluated strategies, the best is decided to be optimal.

Second, aperiodic screening strategies appear in the literature: such strategies incorporate age-dependent screening intervals, accounting for the age-dependent disease progression. Dunnewind (2020) gives an example, which uses a genetic algorithm to optimise screening intervals and cut-offs. However, the resulting strategies are still population-based.

The final category concerns personalised screening strategies. The first step towards optimising such strategies was set by Parmigiani (1997). He formulates screening strategies by *screening intensity functions*, specifying the screening intensity for a given age. Everyone is assigned a personalised function, based on a (constant) individual risk factor which is assumed to be given. However, since this factor is unknown in our case, we prefer to base our strategy on past measurements instead.

The literature describes several models to find optimal screening strategies of this type. A popular choice is the Partially Observable (Semi) Markov Decision Process (POMDP). Examples are described by Ayer et al. (2012) and Maillart et al. (2008) for breast cancer, Otten et al. (2017) for prostate cancer and Erenay et al. (2014) for colorectal cancer. In a POMDP the cancer development is approximated by an unobserved Markov process. Each decision epoch, an observation is made by which the current state of the unobserved process can be estimated. Based on this estimate, an optimal action is decided upon. The POMDP is an interesting approach as the quantitative haemoglobin concentrations can be modelled by the observations. Additionally, the decision is based on the complete screening history. Since solution techniques are usually based on dynamic programming, screening strategies account for the complete future of the patient's life. A significant drawback is the fact that the disease process is approximated

by a Markov Process. The MISCAN model contains many correlations and therefore a Markov model is not suitable.

Ahuja et al. (2017) recognise the problem of the Markov approximation and therefore relax this assumption. It constructs a solution method very similar to the POMDP approach, however it does not require any assumption on the disease progression. Instead it needs a discrete set containing all trajectories by which the disease can develop and a probability for each of them. By sampling smartly from this set and using several theorems from POMDP theory, they approximate the optimal screening strategy.

Tomer et al. (2019) use a longitudinal joint models to optimize the cancer screening strategy for prostate cancer. Given the past screening measurements, the model generates a probability distribution for the survival of a patient for the next x years. Using this probability density, one can optimise the screening interval. The advantage of this method is that correlations within the disease process and the observations are incorporated relatively easily. However, the generated probability distribution does not incorporate the patient's complete screening future: it only calculates the survival probability density until the next screening test. This is a drawback compared to a POMDP.

Another suitable family of problems is the family of maintenance problems, aimed at constructing optimal maintenance strategies for expensive machines. Especially Condition-Based Maintenance (CBM) shows many comparisons with our problem. CBM considers a deteriorating machine and aims to *diagnose* its current status and to *prognose* the remaining time until breakdown. The level of deterioration is (stochastically) related to a performance measure which is measured continuously or at discrete times. Based on the value of the performance measure, one can decide to perform preventive maintenance or, in case of discrete measurement, to schedule a new performance measurement. CBM aims at minimising the costs of maintenance, charged during the machine's downtime, maintenance and performance measurements. The diagnostics part of CBM is similar to our problem as we apply maintenance to a person, based on the performance measure obtained by FITs. In terms of the CBM literature, our problem can be described as a CBM problem with discrete state deterioration which is partially observable and with discrete and non-periodic inspection.

Literature on CBM has been growing rapidly in the past decades (Alaswad and Xiang, 2017; Jardine et al., 2006). Solution methods for CBM can be split into two categories: statistical and data-driven methods (Peng et al., 2010). Statistical methods include model-based methods which assume a model for the deterioration process. Such models usually are (Hidden) Markov

Models for discrete deterioration as described by Kurt and Kharoufeh (2010); Byon and Ding (2010); Zhang et al. (2014) and Gaussian processes such as a Wiener or Gamma processes for continuous deterioration developed by Goode et al. (2000); van Noortwijk (2009). Since the Markov assumption is a key factor in these models, they are less applicable to our problem.

Examples of non-Markovian CBM models exist. Zhao et al. (2010) do not impose restrictions on the deterioration process but it assumes that the deterioration speed is related to the values of several covariates. These covariates change over time, according to a Markov process. However, due to the complexity of their model, the authors do not give an algorithm to find optimal strategies. Instead they simulate many different screening strategies chosen by experts to find the best out of those.

Wang (2012) does create an algorithm for a model with a general structure. The authors built a Bayesian network which incorporates all dependencies between the transitions and observations. If a measurement is obtained, they derive the probability that the machine is in the broken state and if this value is above a threshold, they apply maintenance. Using renewal theory, an optimal replacement policy is constructed.

Opposed to model-based methods, literature proposes data-driven methods using AI techniques. For such techniques the deterioration process is learnt by an AI model such as a neural network. This network estimates the time until breakdown, based on the observations. An overview of AI models used in CBM is given by Peng et al. (2010).

A hybrid form of the two categories is outlined by Marseguerra et al. (2002). In this paper a complex, multi-component machine is considered for which a complex and accurate simulation model of the degradation is known. The authors propose not to approximate this complex model, but to use it in the decision process for a closer adherence to reality: they combine it with a genetic algorithm. The policy space is explored using a Genetic Algorithm. Whenever a policy is selected by the algorithm, the corresponding costs are evaluated using the original simulation model.

5 Methodology

As discussed in the previous section, literature on Condition Based Maintenance often assumes a simplified model of the deterioration of a machine. However, in our analysis, we want to use the complex MISCAN model to simulate development of disease ("patient defects") and evaluate interventions. Marseguerra et al. (2002) acknowledge this problem and present a genetic algorithm that includes a complex degradation model to evaluate a policy. However, this model assumes that the machine is monitored continuously and therefore does not specify a screening interval. To adapt the method to our problem, we extend it with the main feature of POMDPs: the belief space. This section describes the abstract framework in which the two are combined to find cost-effective strategies. Concrete values will be given in Section 6.

5.1 Belief space

As long as patients are not redirected to a hospital, their health state is not directly observed. Therefore, we should estimate the risk of developing a preclinical cancer and choose an action based on the estimate. The space of all estimates is called the *belief space*.

As described in Section 2, the risk of developing CRC can be estimated by the Hb concentrations measured by the FIT. We define the risk estimator R^k. It maps the average of the past k Hb values to a risk value in the range $[0, 1]$, 0 indicating no risk and 1 a high risk. After measurement S_n it is evaluated by

$$R^k := \min\left\{\frac{1}{k}\sum_{i=0}^{k-1}\frac{S_{n-i}}{R_{max}}, 1\right\}. \tag{5.1}$$

In this equation, k represents the number of past Hb concentration included in the estimate and S_{n-i} is the concentration measured at measurement $(n-i)$. R_{max} is a normalisation constant, usually equal to 100μg/g, specifying the maximum feasible cut-off. Any average Hb value greater than R_{max} is equivalent to a risk of 1. This ensures that the algorithm focuses on finding suitable actions for Hb concentrations below 100 as for values above 100, there is no doubt about performing a colonoscopy. We denote the range of outcomes of R^k, $[0, 1]$, by $\mathcal{R}$ in the remainder of this thesis.

After estimating a patient's risk, we should choose an action from $\mathcal{A}$ based on the estimate. The optimal action is expected to be age dependent since the risk of developing a cancer increases with age. Therefore recall the set $\mathcal{T}$, defined as the set containing all eligible screening ages.

Now we can define the belief space $\mathbb{B}$. The belief space $\mathbb{B} := \mathcal{R} \times \mathcal{T}$ is the space of all combinations of outcomes of the risk estimator and patient ages. A vector in the belief space is denoted by $(r, \tau) \in \mathbb{B}$.We define a policy as a function $\pi : \mathbb{B} \to \mathcal{A}$ which maps every vector in the belief space to an action. In fact π makes partitions of $\mathbb{B}$ and links each part to one of the actions in $\mathcal{A}$. An example is shown in Figure 5.1a. Our aim is to find policies to approximate the cost-effective frontier.

5.1.1 Impact ordering assumption

In the following we give an alternative characterization of a policy, applicable to the genetic algorithm. For that we need the concept of *ordering by impact* and a related assumption.

Some actions in the action space have a higher impact on the patient's quality of life than others: a colonoscopy has a higher impact than a FIT as the patient has to visit a hospital instead of the general practitioner. The action FIT_{10} has a lower impact than FIT_5 as a patient has to see his doctor less often. A formal definition is given in definition 5.1.

Definition 5.1 (Impact ordering)**.** *Given two actions $a_1, a_2 \in \mathcal{A}$. The impact of a_1 is higher than (or equal to) the impact of a_2 if and only if they adhere to one of the following:*

1. *$a_1 = FIT_i$, $a_2 = FIT_j$ and $i \leq j$;*
2. *$a_1 = COL$.*

For this, we introduce the following notation $a_1 \succeq_{imp} a_2$. Additionally, the two actions are called consecutive if no action $a_3 \in \mathcal{A}$ exists with $a_1 \neq a_2 \neq a_3$ such that $a_1 \succeq_{imp} a_3 \succeq_{imp} a_2$.

Note that a colonoscopy has a higher impact than any other action in $\mathcal{A}$. Now we introduce the impact ordering assumption which specifies a property that every feasible policy should adhere to.

Assumption 5.2 (Impact ordering assumption)**.** *Given two belief vectors $(r_1, \tau), (r_2, \tau) \in \mathbb{B}$ with equal ages τ and $r_1 \neq r_2$. The corresponding actions as specified by policy π are a_1, a_2 respectively. If π is feasible, the following must hold:*

$$r_1 > r_2 \iff a_1 \succeq_{imp} a_2.$$

Intuitively, it says that a higher risk estimate must correspond with an action with a bigger impact. This is a natural assumption. Individuals with a higher risk are more likely to develop a cancer and need a intenser screening strategy, i.e. they need an action with a higher impact.

An example of a feasible and infeasible policy is given in Figure 5.1.

The impact ordering assumption has an important consequence as it restricts the way a policy π makes partitions of $\mathbb{B}$. For a fixed age τ, the $\mathcal{R}$-dimension is split into $|\mathcal{A}|$ (possibly empty) intervals, each corresponding with one of the actions. The order of the intervals is fixed due to the assumption. The intervals may differ by age. Therefore we can see the upper bound of the interval for action a as a function $\beta_a : \mathcal{T} \to \mathcal{R}$. Note that $\beta_{COL}(\tau) = 1$ for all τ since the intervals must form a partitioning of $\mathcal{R}$ (see Figure 5.1a). In fact π can be characterised by the set of functions $\{\beta_a(\tau)\}_{a \in \mathcal{A}}$. In the remainder of this thesis, the functions $\beta_a(\tau)$ will be referred to as *policy bounds*.

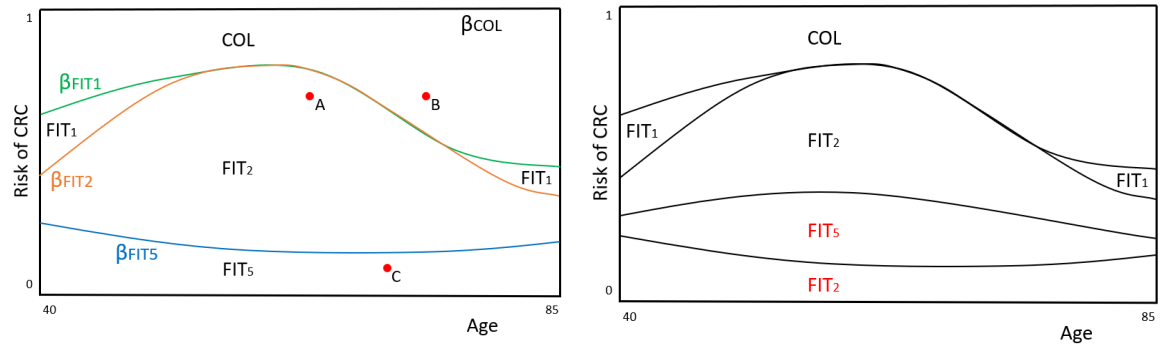

Figure 5.1: (a) A feasible policy. The belief space is partitioned and for each age a (possibly empty) interval is specified for each action. The red dots represent three patients: patients A and C are prescribed a FIT in two years, patient B is redirected to a hospital for a colonoscopy. The bold lines represent the policy borders, denoted by $\beta_a(\tau)$. Note that $\beta_{COL} = 1$. (b) An infeasible policy. Since the actions FIT_2 and FIT_5 are swapped, people with a higher risk may get a longer screening interval than people with a lower risk.

5.1.2 Discretisation of belief space

We derived that policies can be characterised by a finite set of functions in the belief space. In our algorithm we assume that they are piecewise constant. It allows us to discretise the belief space, which reduces the amount of feasible policies. Therefore the algorithm is likely to converge faster.

We discretise the spaces $\mathcal{R}$ and $\mathcal{T}$ into finite spaces, denoted by $\hat{\mathcal{R}}, \hat{\mathcal{T}}$ respectively. Usually $\hat{\mathcal{R}} = \{0, 0.1, \ldots, 1\}$, allowing the cut-offs to vary in steps of 10μg/g and $\hat{\mathcal{T}} = \{40, 45, \ldots, 85\}$, implying that policies may very their intervals every five years. Consequently we define $\hat{\mathbb{B}} = \hat{\mathcal{R}} \times \hat{\mathcal{T}}$ which is a two-dimensional grid that approximates the belief space $\mathbb{B}$.

Due to the discretisation, policies get an even simpler characterisation. The partition borders become functions $\hat{\beta}_a : \hat{\mathcal{T}} \to \hat{\mathcal{R}}$, defined on a discrete, finite domain and they can be represented

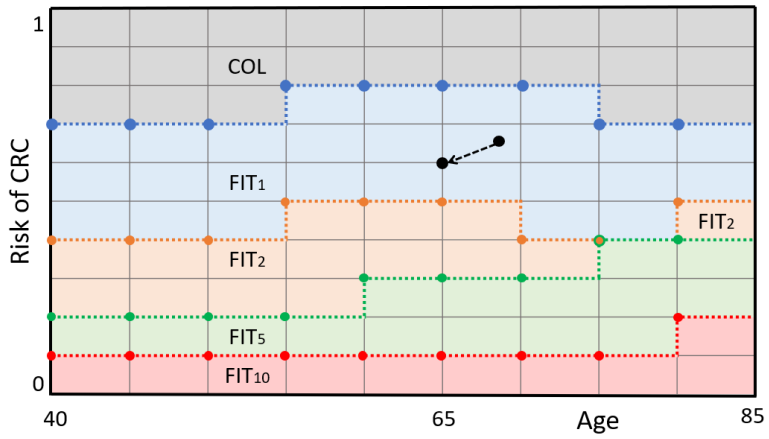

Figure 5.2: A policy in the discrete belief space. The policy edges (indicated by dotted lines) can only occupy parts of the grid. The coloured dots indicate the values included in each of the vectors $\boldsymbol{\beta}_a$. The black dot shows how the action is determined for an observation. The observation (69, 0.67), marked by the black dot, is treated as (65, 0.6) as indicated by the arrow.

by a vector $\hat{\boldsymbol{\beta}}_a \in \hat{\mathcal{R}}^{|\hat{\mathcal{T}}|}$. This characterisation is exploited in the genetic algorithm and from this point, a policy is represented by a set of vectors $\{\hat{\boldsymbol{\beta}}_a\}_a$. An example of a discrete policy is given in Figure 5.2. Since the policy borders are to be defined on the grid $\hat{\mathbb{B}}$, the policy borders are the same for all people between, for example, 60 and 64 years old.

5.1.3 Three-dimensional belief space

Up to now we have considered a two-dimensional belief space. We can also extend it into a third dimension. This allows us to incorporate a second risk estimator. We propose the *velocity estimator* V^k, based on the net change in the past $(k+1)$ Hb concentration measurements. This value is relevant as a large increase in concentration also hints at an increased risk of developing CRC. After measurement S_n, V^k is calculated by

$$V^k := S_n - S_{n-k}. \tag{5.2}$$

We denote the space of possible outcomes of V^k by $\mathcal{V}$.

Let us adapt some of the definitions and properties of the belief space to the three-dimensional case. The three-dimensional belief space is defined as $\mathbb{B} := \mathcal{R} \times \mathcal{V} \times \mathcal{T}$. A vector in the space is denoted by $(r, \boldsymbol{\tau}) \in \mathbb{B}$ with $\boldsymbol{\tau} := (v, \tau)$. Similar to the two-dimensional case, a policy π is a mapping between the space $\mathbb{B}$ and the set $\mathcal{A}$ and therefore partitions $\mathbb{B}$. The impact ordering assumption is adapted as follows:

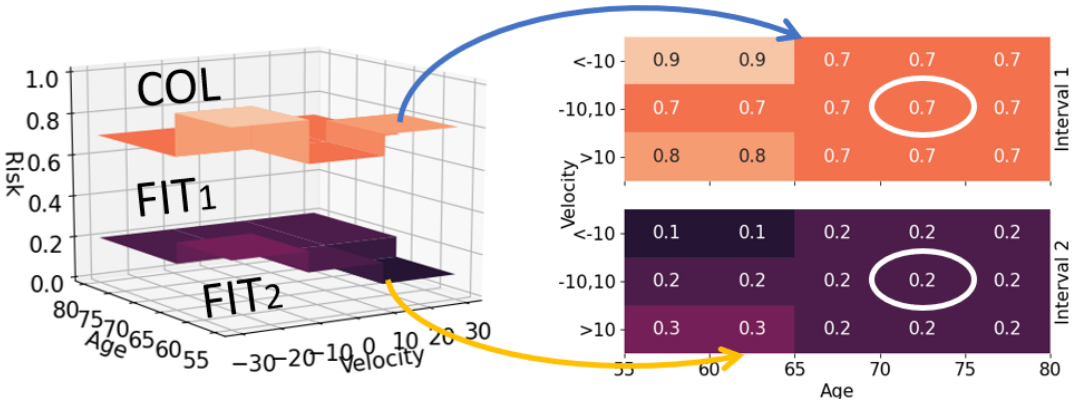

Figure 5.3: A 3-D policy is shown on the left. The dimensions correspond with the age, velocity and risk estimates. The belief space is partitioned and the policy borders are surface functions, represented by the matrices on the right (as shown by the arrows). In this case, a patient aged 70 with a velocity estimate between -10 and 10 undergoes a colonoscopy if $r \geq 0.7$, gets an interval of one year if $0.2 \leq r < 0.7$ and of two years otherwise.

Assumption 5.3 (Impact ordering assumption (3D))**.** *Given two belief vectors* $(r_1, \boldsymbol{\tau}), (r_2, \boldsymbol{\tau})$ *with equal ages and velocity estimates* $\boldsymbol{\tau}$ *and different risk estimates* r_i*. The corresponding actions as specified by policy* π *are* a_1, a_2 *respectively. If* π *is feasible, then*

$$r_1 > r_2 \iff a_1 \succeq_{imp} a_2.$$

Now the actions must be of increasing impact for an increasing r if both τ and v remain constant. Therefore, the partitioning of $\mathcal{R}$ may differ for each vector $\boldsymbol{\tau} \in \mathcal{V} \times \mathcal{T}$. We define the functions $\beta_a(\boldsymbol{\tau}) : \mathcal{V} \times \mathcal{T} \to \mathcal{R}$ which specify the upper bound of the interval for action a for each value of $\boldsymbol{\tau}$. A policy is now characterised by the set of functions $\{\beta_a(\boldsymbol{\tau})\}_{a \in \mathcal{A}}$.

In the three-dimensional setting, we impose a three-dimensional piecewise constant functional form to the policy bounds. Therefore we can also discretise the space $\mathcal{V}$ into $\hat{\mathcal{V}}$ and the partition borders become functions $\hat{\beta}_a : \hat{\mathcal{V}} \times \hat{\mathcal{T}} \to \hat{\mathcal{R}}$. Since these space are all finite and discrete, we can represent these functions by matrices $\hat{\boldsymbol{\beta}}_a \in \hat{\mathcal{R}}^{|\hat{\mathcal{V}} \times \hat{\mathcal{T}}|}$. Again, this representation is exploited in the genetic algorithm. From this point, a policy for the three-dimensional belief space is represented by a set of matrices $\{\hat{\boldsymbol{\beta}}_a\}_a$. An example of a policy is given in Figure 5.3.

In the remainder of this thesis we use the discretised spaces only. Therefore we use the notations $\mathcal{R}, \mathcal{T}, \mathcal{V}, \boldsymbol{\beta}_a$ to refer to the discretised elements $\hat{\mathcal{R}}, \hat{\mathcal{T}}, \hat{\mathcal{V}}, \hat{\boldsymbol{\beta}}_a$.

5.2 Genetic algorithm

Having defined the concept of belief space and policies, we introduce the genetic algorithm which optimises the policies. A genetic algorithm (GA) is an evolutionary algorithm. Its central idea is the principle of survival of the fittest from biology. The algorithm maintains a population of solutions and evaluates their fitness. Then it chooses the best solutions from the population and uses their features to generate new solutions, assuming that better solutions may arise from combining features of good solutions. New solutions are subject to random mutations to ensure a diverse population and to escape local minima.

Genetic algorithms have several features which make them appropriate for this problem. First, in each generation, the costs and effects of the policies in a population can be evaluated in parallel. Since running the MISCAN model consumes most of the running time of the algorithm, this speeds up the process compared to a sequential optimisation algorithm. Second, genetic algorithms are a popular choice for multi-objective optimisation problems. Since they keep track of a population of solutions, they can find complete approximation sets. Finally, genetic algorithms are known to escape local optima well due to their randomisation. This increases the chance of finding a (nearly) global optimum to the problem.

This section describes the key elements of a genetic algorithm and their relation with the MISCAN model and the belief space. If not specifically discussed, the characteristics are the same for the two- and three-dimensional belief spaces.

5.2.1 Chromosomes and initialisation

A genetic algorithm maintains a population of policies. As mentioned in Section 5.1.2, a policy is represented by a set of vectors (or matrices) $\{\boldsymbol{\beta}_a\}_{a\in\mathcal{A}}$. Before the algorithm is started, an initial population of policies is generated as follows. For each policy, each border is assigned a constant value: $|\mathcal{A}|$ values are drawn from $\mathcal{R}$, they are sorted and each of them is assigned to one of the policy borders, adhering to the impact ordering property. Now a policy border is a constant function as it has the same value for each age.

Then the policy borders are adapted to obtain piecewise constant functions. We apply the mutation operator as described in Section 5.2.5. The steps of the procedure are sketched in Figure 5.4.

We repeat this procedure N_{pop} times such that we generate an initial population containing N_{pop} policies.

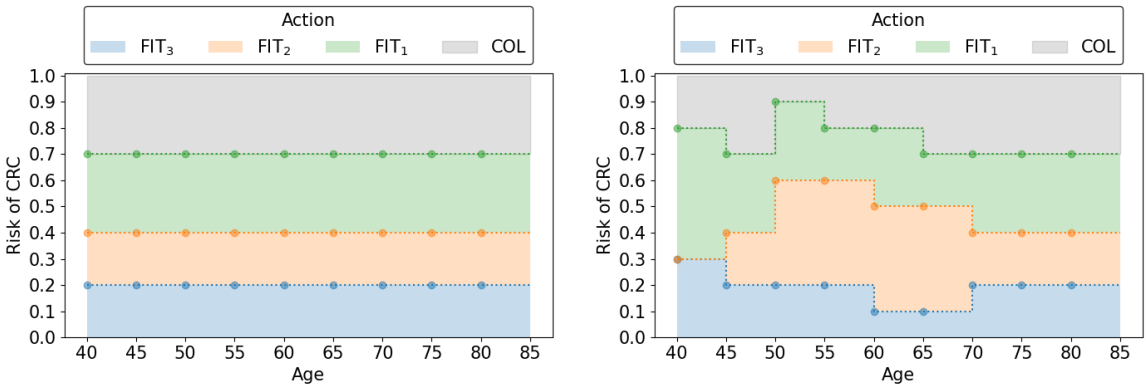

Figure 5.4: The initialisation consists of two parts. First policy edges are assigned a constant value, adhering to the impact ordering assumption (left figure). After that, random mutations are applied and the policy borders become piecewise constant (right figure).

5.2.2 Fitness function

Each of the policies in our population needs a fitness measure, to be used in the selection procedure (Section 5.2.3). It is based on a policy's costs and effects which are obtained from MISCAN. Since the optimisation problem is bi-objective, the obtained values are not directly comparable to determine the fittest policy and we need a Multi-Objective Genetic Algorithm (MOGA) to select the fittest policies. Dunnewind (2020) concluded that the Non-dominated Sorting Genetic Algorithm-II (NSGA-II) from Deb et al. (2000) worked best in his experiments. Therefore we use this algorithm to calculate the fitness of a policy. This paragraph explains the algorithm.

The fitness of a solution consists of two elements: its rank and its crowding distance. The rank is a measure for how non-dominated a solution is. Given a population of solutions P, the algorithm searches for those in its Pareto optimal frontier $\psi(P)$ and assigns them rank 1. Then these solutions are discarded and the algorithm determines the Pareto optimal solutions out of the remainder. These are assigned rank 2. This procedure is repeated until all solutions are ranked. A result of the algorithm is illustrated in Figure 5.5. The red, non-dominated solutions in this cost-effectiveness plot have rank 1. The yellow solutions are dominated by those with rank 1 only and obtain rank 2. The green solutions have rank 3 as they are dominated by those with ranks 1 and 2 only, etc. The lower the rank of a solution, the higher its fitness.

However, all solutions with the same rank seem equally fit while we may prefer certain solutions over others. For example, we prefer to maintain a widely-spread population in the cost-effectiveness space (expensive as well as cheaper policies) as this increases the chance of finding the full Pareto optimal frontier. To that end, NSGA-II assigns each solution a crowding distance, a measure for how isolated a solution is in the cost-effectiveness space. If a solution is

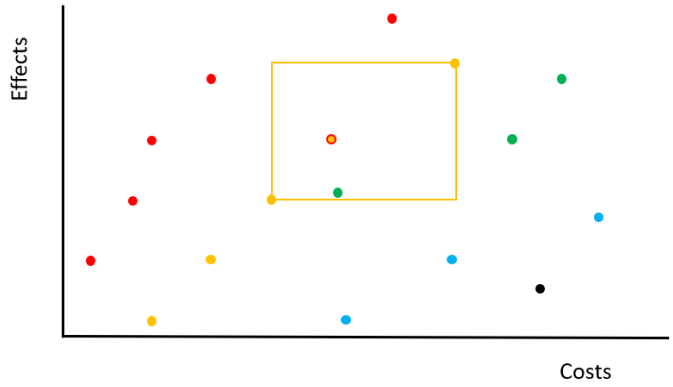

Figure 5.5: Illustration of the NSGA-II algorithm. The red, yellow, green, blue, black dots obtain rank 1 to 5 respectively. Additionally, the rectangle that defines the crowding distance is drawn for the yellow dot with a red edge.

in a crowded area, the probability that a similar solution is selected is larger because there are many of them. In contrast, there are no similar alternatives for an isolated solution. Therefore the NSGA-II algorithm has a slight preference towards isolated solutions.

The *crowding distance* is calculated as follows. Consider a frontier of solutions $\mathcal{F}$ containing n solutions. These solutions have the same rank because they are in the same frontier. They are sorted by increasing costs and decreasing effects to obtain a sorted sequence of policies $(\pi[1], \ldots, \pi[n])$. For each of these solutions $\pi[i]$ we draw a rectangle that touches its pre- and successor $\pi[i-1]$ and $\pi[i+1]$ (see Figure 5.5). The crowding distance of solution $\pi[i]$ is defined as the circumference of this rectangle. Note that this calculation implies that the crowding distances of $\pi[1]$ and $\pi[n]$ equal ∞.

The combination of the rank and crowding distance of solutions gives an ordering in the population. Those with a lower rank are better and in case of equal ranks, we prefer solutions with a high crowding distance. Based on these values, we can start the selection procedure.

The calculation of the rank of each of the policies is the most complex, leading to an overall complexity of the NSGA-II algorithm of $O(2N_{pop}^2)$. The factor 2 corresponds with the dimensionality of the cost-effectiveness space (Deb et al., 2000).

5.2.3 Selection operators

Having assigned a fitness to each of the policies, we select which solutions are taken to the next generation and are used for cross-over. The set of selected solutions is called the mating pool and the *mating pool* in iteration g is denoted by M_g. During the selection phase, the mating pool is filled until it contains $N_{sel} := N_{pop}/2$ policies (this implies that N_{pop} must be even).

The literature provides different types of selection procedures, amongst which *elitist selection*

and *tournament selection*. In the first, the mating pool is filled with the solutions with the highest fitness, no randomisation is applied. The latter does apply randomisation: a small amount of solutions is randomly sampled from the population and only the fittest of this sample is added to the mating pool.

Our selection procedure is a combination of both types as it consists of two phases. In the *elitist phase* a number of solutions is selected in an elitist manner. Usually the mating pool is not completely filled yet. Then the *tournament selection phase* selects solutions with tournament selection until the mating pool is filled. In this thesis, we test three different procedures for the elitist phase. The tournament selection phase remains the same for these three cases.

Elitist phase Three procedures were tested for the elitist phase: strong, weak and no elitism. Given the current population P_g, *strong elitism* adds the strong approximation set $\psi(P_g)$ to M_g. This procedure ensures that all good solutions are included in the mating pool. *Weak elitism* only adds the weak approximation set $\xi(P_g)$ to M_g. It ensures that the best solutions are maintained, but also allows for more randomization as $\xi(P_g)$ is usually significantly smaller than $\psi(P_g)$. *No elitism* skips the elitist phase: the mating pool is only filled by tournament selection.

Tournament selection phase If the mating pool was not completely filled with N_{sel} solutions during the elitist phase, the tournament selection phase fills the remainder. Two policies are randomly sampled from the population. The policy that scores best in terms of rank or, if necessary, crowding distance is added to the mating pool. This is repeated until the mating pool is filled.

This procedure allows duplicates in the population: a policy can be selected in the elitist phase and multiple times in the tournament selection phase, especially those of higher rank. Since these solutions score well on fitness, they are likely to have good features. Therefore it could be advantageous that they are included in the mating pool more than once.

5.2.4 Cross-over operator

Having filled the mating pool M_g we use these solutions for cross-over to obtain offspring. In our method, 2-point cross-over is employed. The solutions in M_g are paired up randomly to obtain pairs of parents. Then, for each of the pairs, two ages τ_1, τ_2 are randomly selected and the policy bounds for the ages in the interval $[\tau_1, \tau_2]$ are exchanged. The resulting two sets of policies are added to the set of offspring obtained in generation g, O_g. After all of the policies in M_g have created offspring, O_g has a size of $N_{pop}/2$. The operator is illustrated in Figure 5.6.

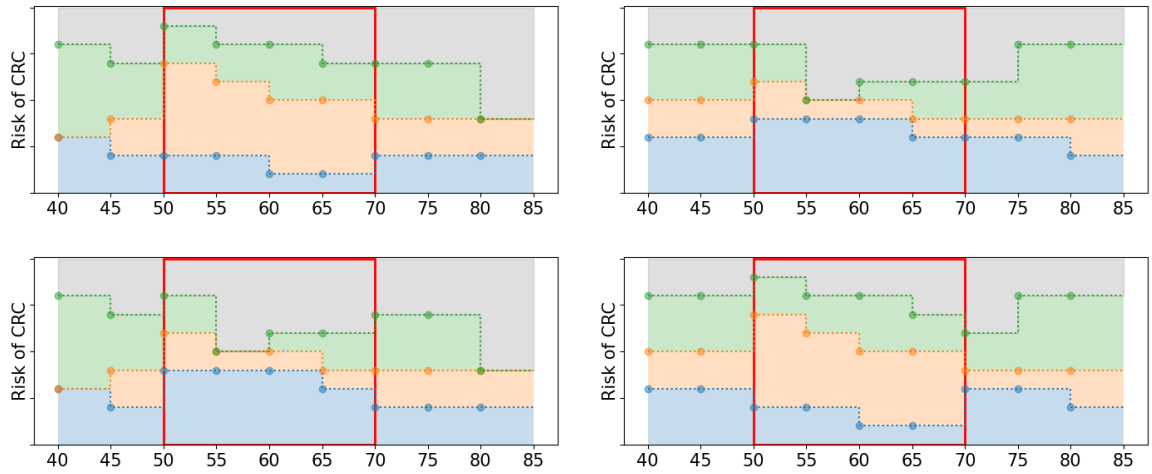

Figure 5.6: An illustration of the cross-over operator. The two upper figures represent the parent policies, the lower figures show the offspring. Two random ages are selected (50 and 65 years old) and the values in between those ages, including the selected ages, are exchanged (red box).

3-D belief space In case of the three-dimensional belief space, cross-over is slightly adapted. Two ages are chosen randomly in the space $\mathcal{T} \times \mathcal{V}$ and the area in the rectangle between these two points is exchanged between the two policies (Figure 5.7).

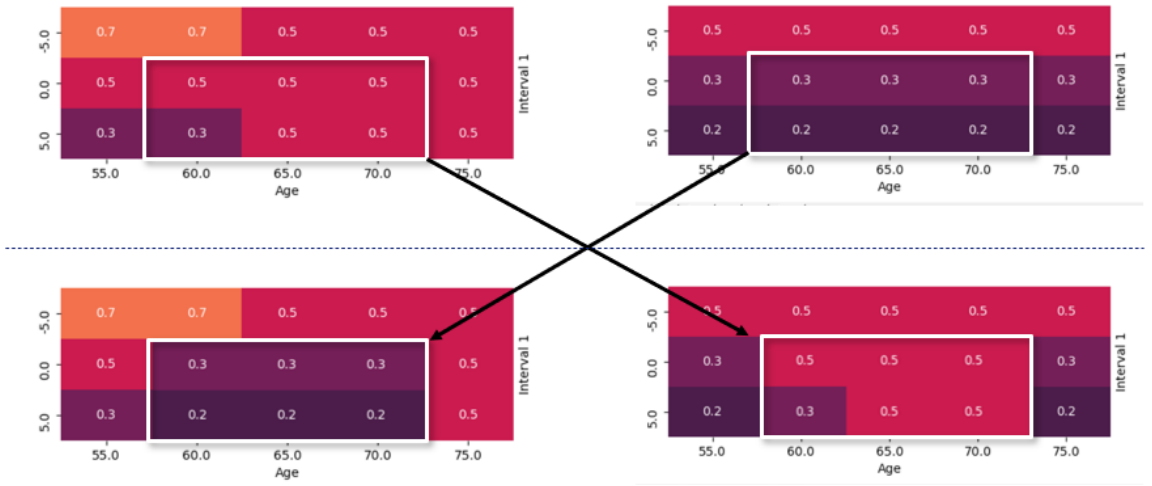

Figure 5.7: An illustration of the three-dimensional cross-over operator. The two upper figures represent one $\boldsymbol{\beta}_a$ of the parent policies, the lower figures show those of the offspring. Two $\boldsymbol{\tau}$ are selected ($(60, 5)$ and $(70, 0)$) and the values in between (in the white box) are exchanged, including the two selected $\boldsymbol{\tau}$.

5.2.5 Mutation operators

After cross-over, offspring is subjected to random mutations with probability p_M. Two different mutation operators were tested: *uniform* and *bounded mutation*.

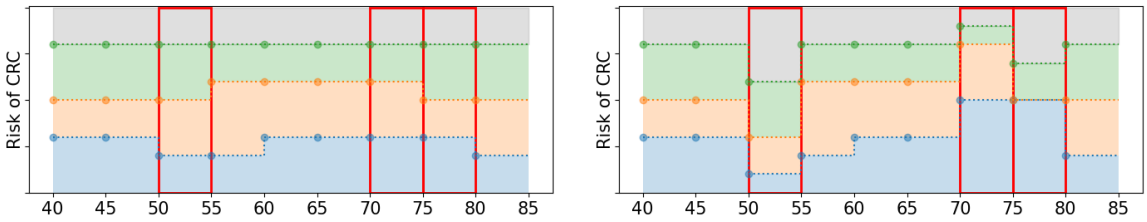

Figure 5.8: An illustration of the uniform mutation operator. The original policy is shown on the left, the mutant on the right. The selected element, (ages 50, 70 and 75) are marked by a red box. Each of the mutant's policy edges obtains a new random value for these ages. In this case $p_e = 3/9$ as three out of the nine age groups were selected for mutation.

Uniform mutation If we mutate a policy, we select a percentage p_e of the elements in its vectors $\boldsymbol{\beta}_a$. The values of all policy borders $\{\boldsymbol{\beta}_a\}_a$ are mutated for these elements. Uniform mutation does so by drawing $|\mathcal{A}|$ random values from $\mathcal{R}$, with replacement. It then sorts these values and assigns them to the vectors, adhering to the impact ordering assumption. See Figure 5.8 for an illustration.

Bounded mutation A drawback of uniform mutation is that it is unlikely to put different policy boundaries close together as the assigned values are sampled uniformly from $\mathcal{R}$ while might be a characteristic of a good policy. For example, a policy shown in Figure 5.9 is unlikely to be generated. To overcome that problem, a second mutation operator was introduced. In bounded mutation, the values are sampled from a subinterval of $\mathcal{R}$. This increases the chance of creating a policy in which the policy edges are close to each other.

The operator works as follows. For each of the to be mutated elements in $\boldsymbol{\beta}_a$, the operator first samples one value $\tilde{r} \in \mathcal{R}$. This value is an upper bound or a lower bound with 50% probability. If it is an upper/lower bound, $|\mathcal{A}|$ random values are selected from $\mathcal{R}$ below/above $\tilde{r}$ respectively. Then these values are assigned to the vectors $\{\boldsymbol{\beta}_a\}_a$ while adhering to Assumption 5.2. An illustration is given in Figure 5.10.

3-D belief space The two operators work the same way for the three-dimensional belief space as for the two-dimensional space. Here p_e% of the elements in the matrix $\boldsymbol{\beta}_a$ are mutated with the chosen operator.

The mutated policies replace their original policy in the set of offspring O_g. Together with the mating pool M_g, this set forms the new population used in the next generation: $P_{g+1} = M_g \cup O_g$.

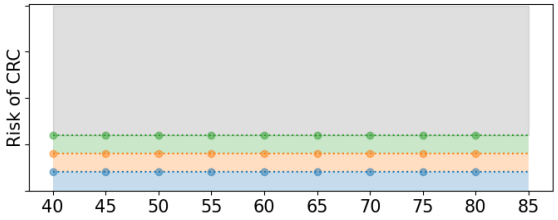

Figure 5.9: An example policy unlikely to be obtained with uniform mutations. Since uniform mutation samples the values uniformly from $\mathcal{R}$, it is unlikely to put the policy edges this close together. This policy is more likely to be obtained with bounded mutation.

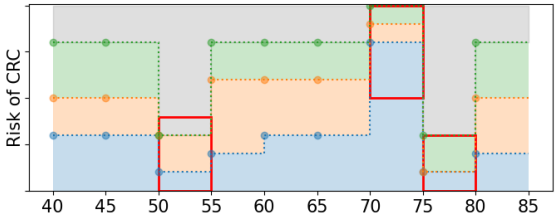

Figure 5.10: An illustration of the bounded mutation operator. The original policy is the same as in Figure 5.8. New policy borders are sampled from a subinterval of $\mathcal{R}$ (red box). At ages 60 and 75, $\tilde{r}$ is 0.4 and 0.3 resp. and is an upper bound. For age 70, $\tilde{r} = 0.5$ is a lower bound.

5.2.6 Memory and stopping criterion

Since some of our selection procedures are non-elitist, we need a memory to ensure that the best solutions obtained throughout the generations are not forgotten. We do so by memorising the strong approximation set of all solutions obtained up to iteration g, denoted by $\Psi_g := \psi(\bigcup_{i=1}^{g} P_g)$. In each generation, Ψ_g is updated with respect to Ψ_{g-1} and the number of added policies is memorised. The stopping criterium is satisfied if this number has equalled 0 for N_{stop} consecutive iterations. If meeting this criterion takes too long, the algorithm is stopped after a fixed number of iterations. The frontier Ψ_g of the final generation is the output of the algorithm as it contains the best found solutions.

5.3 Performance measures

In the previous section, we introduced several parameters and a small amount of selection and mutation operators. The performance of the algorithm depends on the chosen operators as well as the chosen parameter values. One way to compare the quality of two parameter sets is to plot the obtained approximation sets and compare them visually. However, it is hard and time consuming to do this for many sets at the same time.

Performance measures (PMs) solve this problem. They summarise the quality of an approximation set by one number, allowing for an easy comparison between the parameter sets. A requirement to a performance measure is that it is *dominance preserving* (Zitzler and Künzli, 2004). That means that if each solution in an approximation set B is dominated by at least one solution in approximation set A, the performance measure must assign a higher value to A than to B. Additionally, there are four properties of an approximation set that a performance measure can capture (Li and Yao, 2019).

1. *Convergence*: How distant is an approximation set from the POF, provided that the POF is known?

2. *Spread*: To what extent are the solutions in the approximation set spread throughout the cost-effectiveness space?

3. *Uniformity*: Are the solutions in the approximation set uniformly spread throughout the cost-effectiveness space?

4. *Cardinality*: How many solutions are included in the approximation set?

A performance measure that captures most of these properties is preferred.

A large variety of performance measures exists, but the one that is used most prominently in the literature is the hypervolume (Riquelme et al., 2015; Audet et al., 2020), introduced by Zitzler and Thiele (1999). The hypervolume (HV) of an approximation set represents the area covered by the rectangles between each of the points in the set and a reference point, as shown in Figure 5.11a. It captures three properties of an approximation set: the convergence, spread and cardinality.

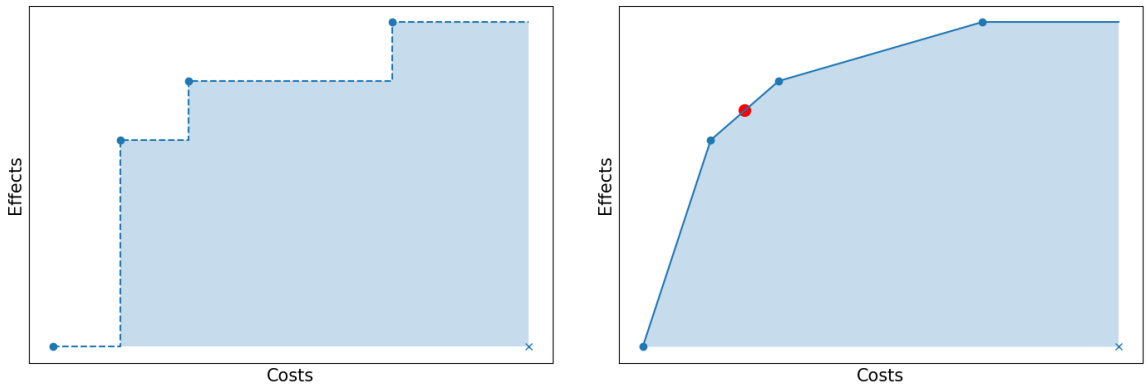

Figure 5.11: (a) The hypervolume equals the area between the step function connecting the approximation set and a reference point (marked by the cross).
(b) The extended hypervolume equals the area between the lines connecting a weak approximation set and a reference point. While the cardinality increases by adding the red policy to the blue set, the EHV remains the same.

However, the hypervolume is not dominance preserving for weak dominance. This is illustrated in Figure 5.12 where two hypothetical weak approximation sets are shown. The blue includes the points (0.1,0), (500,10) and (600,10.5), the orange includes (0,0), (510,10.5), (600,10.6). The orange approximation set weakly dominates the blue one as all (linear combinations of) blue strategies are weakly dominated by a linear combination of orange policies. Therefore we would expect the HV of the orange to be larger than that of the blue approxima-

tion set. However, the HV of the blue set equals 1000, while that of the orange set equals 945. The hypervolume is not dominance preserving with respect to weak dominance.

Therefore we introduce the *extended hypervolume* (EHV). It is defined as the area between the line connecting a weak approximation set and the reference point (see Figure 5.11b). This performance measure does assign a higher value to the orange approximation set of Figure 5.12 compared to the blue one.

A drawback of this extension is that the measure loses the reflection of cardinality, as shown in Figure 5.11b. The EHV of the blue approximation set remains unchanged if the red point is added, while the cardinality is increased. It can be discussed, however, to what extent this reduces the quality of the performance measure since the red point is, in fact, already covered by a linear combination of the policies in the blue approximation set.

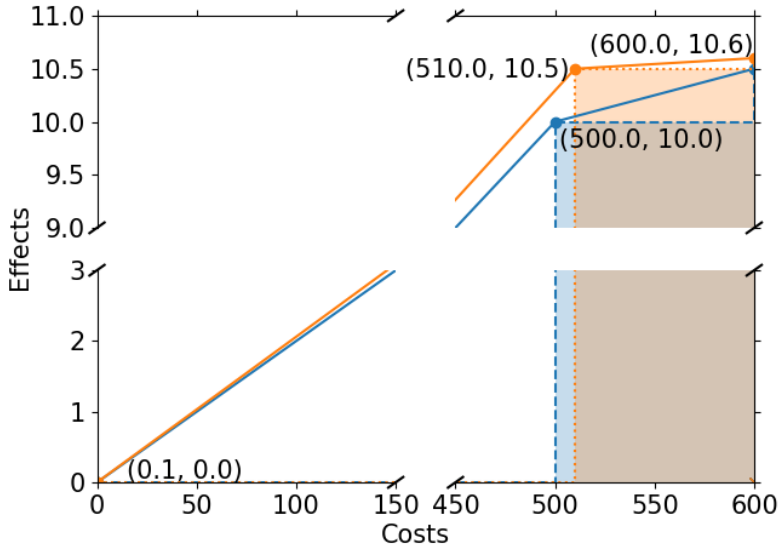

Figure 5.12: The HV does not represent the quality of a weak approximation set well. The orange set dominates the blue but the HV of the blue is higher than that of the orange. Note the distorted scale on the axes. Not all coordinates are shown: the coordinates of the leftmost orange point and rightmost blue points are (0,0) and (600, 10.5) resp.

The hypervolume provides us information on how much an approximation set has converged towards the Pareto optimal frontier, the extended hypervolume does so for the cost-effective frontier. Therefore we consider both performance measures when comparing different parameter settings. If both measures are the highest for certain settings, it is likely that these settings work well in that specific case.

6 Results

The results section consists of two parts. First we discuss how the parameter values and operators were chosen with a small instance of the belief space. The chosen settings were used to run the algorithm on a large instance to compare different estimators which is discussed secondly.

Some of the settings remained the same in both instances. $\mathcal{T}$ was discretised in steps of 5 years. The action set was chosen as $\mathcal{A} = \{\text{COL}, \text{FIT}_1, \text{FIT}_2, \text{FIT}_3\}$ which includes three screening intervals. If patients turn out to be healthy after a colonoscopy, they are not screened for a fixed interval of 5 years. We assumed screening participation of 100%. The reference point for the (extended) hypervolume was set to a cost of 5000$/individual and an effect of 0.

The costs and effects of a policy were obtained by simulation with MISCAN. While a simulation size of at least 1e7 individuals is common to ensure robust results, a smaller size was chosen to limit the running time. Common seeds ensure that the simulation results of different policies are still comparable. The Hb concentrations measured by the FITs were simulated using the unvalidated zero-inflated negative binomial distribution, as described in Appendix A. The costs and effects of a policy were discounted by 3% annually from the age of 40 years and were calculated with respect to the current Dutch screening strategy: a biennial FIT between the ages 55 and 75.

6.1 Parameter tuning and operator selection

Section 5.2 introduced several parameters and operators. To choose the best settings for our algorithm, we tested several on a small instance of the belief space. The amount of feasible policies on this space is limited, which allows us to evaluate the costs and effects of all of them before starting the algorithm. This way we can derive the optimal POF and CEF by enumeration and the algorithm runs much faster as we do not need simulation time. Therefore we can quickly evaluate how closely the POF and CEF are approximated with certain parameter settings.

The small belief space is two-dimensional. People are screened between ages 55 and 75 so $\mathcal{T} = \{55, 60, 65, 70\}$. We applied risk estimator R^1 with $R_{max} = 50\mu$g/g. $\mathcal{R}$ is discretised into $\{0, 0.25, 0.5, 0.75, 1\}$. This combination of parameters gives 1,500,625 feasible policies (see Appendix B). All of them were evaluated using the Extreme-scale Model Exploration with Swift (EMEWS) framework (Ozik et al., 2016) on the Bebop supercomputer at the Argonne National Laboratory. These simulations were done with 2e6 individuals.

Each dot in Figure 6.1 represents a feasible policy. We can see that policies with negative costs and positive effects exist, strongly dominating the current Dutch strategy. The POF of this set of policies has a HV of 134.49 and the CEF has an EHV of 135.70.

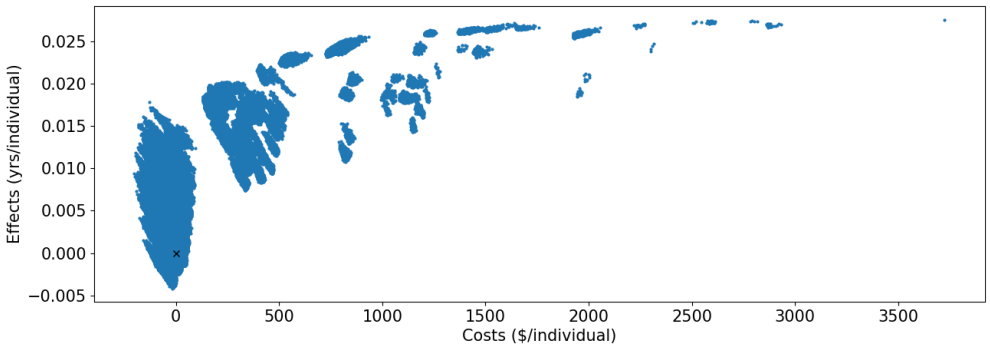

Figure 6.1: Results of the small belief space. Each dot represents a feasible policy. The origin of the plot, marked by a '+', corresponds with the current Dutch screening strategy.

6.1.1 Population size and selection operator

First we experimented with the population size N_{pop} and the selection operator. We tested the three different elitist selection phases from Section 5.2.3 (strong, weak and none). N_{pop} takes values between 100 and 600, increasing with steps of 100.

All combinations of population sizes and selection operators were tested as follows. The algorithm was run fifty times until convergence for each combination. The seed used for the initial population remained the same for each setting. In every iteration g, the HV (EHV) of the found approximation sets Ψ_g ($\xi(\Psi_g)$) were evaluated and averaged over the fifty runs. In these runs, the mutation settings were fixed at $p_M = 0.2$, $p_e = 0.6$ with bounded mutation. N_{stop} was set at 20. Their evolution is shown in Figure C.1 in the appendix.

It turns out that all settings perform very well for this small instance: the optimality gap is less than 0.7% for both PMs and all settings. We observe that weak elitism usually performs better than no elitism or strong elitism. While no elitism apparently selects too randomly, possibly leading to a local optimum, strong elitism is likely to be too elitist, including a lot of duplicate policies in the population, leading to a lack of randomisation. Therefore we choose to select by weak elitism in our algorithm. Finally, we see that performance usually increases with population size. For practical reasons, we choose a population size of 400 for our algorithm.

6.1.2 Mutation operator and parameters

Next we select the mutation operator and mutation rates for our algorithm. Recall the uniform and bounded mutation operators specified in Section 5.2.5 and the symbols p_M and p_e which represent the mutation rate and the percentage of elements of a vector $\boldsymbol{\beta}_a$ selected for mutation. p_M was chosen to vary between 0.1 and 0.5 with steps of 0.1 and p_e varies between 0.2 and 0.8 with steps of 0.2. The variation of p_M is limited as we expect worse performance if more than

half of the offspring is mutated, leading to too much randomisation. We do test high values for p_e as it may be beneficial to have a very rigorous mutation operator. The other parameters remained as described above.

There are 2*5*4=40 combinations of these settings. The algorithm was run fifty times for each combination. The average values of the PMs after convergence for each setting are shown in Figure C.2 in the appendix. The two top graphs show that the algorithm performs poorly for $p_e = 0.2$. Then the mutation is so marginal that the algorithm is unlikely to escape local minima.

The two bottom graphs show the same graphs with a different scale on the vertical axis to reveal the differences between the best parameter settings. There is no clear difference in performance between uniform and bounded mutation: both operators outperform the other in several cases for both PMs. We neither see a clear pattern between the different parameter values. The algorithm seems to perform better for increasing p_M but this is not always the case. Bounded mutation with mutation rates $p_M = 0.3$ and $p_e = 0.6$ is one of the better parameter sets. These values are used in the remainder of this thesis.

6.2 Estimator comparison

Having selected good parameters and operators for our algorithm, we test the different estimators R^k and V^k on a large instance of the belief space. In this space, individuals undergo screening between the ages 40 and 85 so $\mathcal{T} = \{40, 45, \ldots, 80\}$. We allow cut-offs up to $R_{max} = 100\mu$g/g, discretised in steps of 10 so $\mathcal{R} = \{0, 0.1, \ldots, 0.9, 1\}$. Policies are evaluated with a simulation size of 1e6 individuals. The stopping criterion is $N_{stop} = 30$. The remainder of the settings are the same as above. This combination of policies gives 12.8e21 feasible policies (see Appendix B). As it is impossible to evaluate all of them, we need our algorithm

First we describe the results of the algorithm in the baseline in which we use estimator R^1. Then we consider R^2 and R^3. Finally we test the combination of R^1 and V^1, which involves a three dimensional belief space.

6.2.1 Baseline (R^1)

In the baseline, R^1 is used to estimate the risk of CRC. Figure 6.2 shows the evolution of the algorithm throughout the generations. The upper two graphs show the evolution of the (E)HV, the lower shows the amount of policies that were added and removed from the memory Ψ_g per generation. MISCAN was the bottleneck of the algorithm in terms of running time: simulation of a generation took approximately five minutes on the Bebop computer, while the operators of

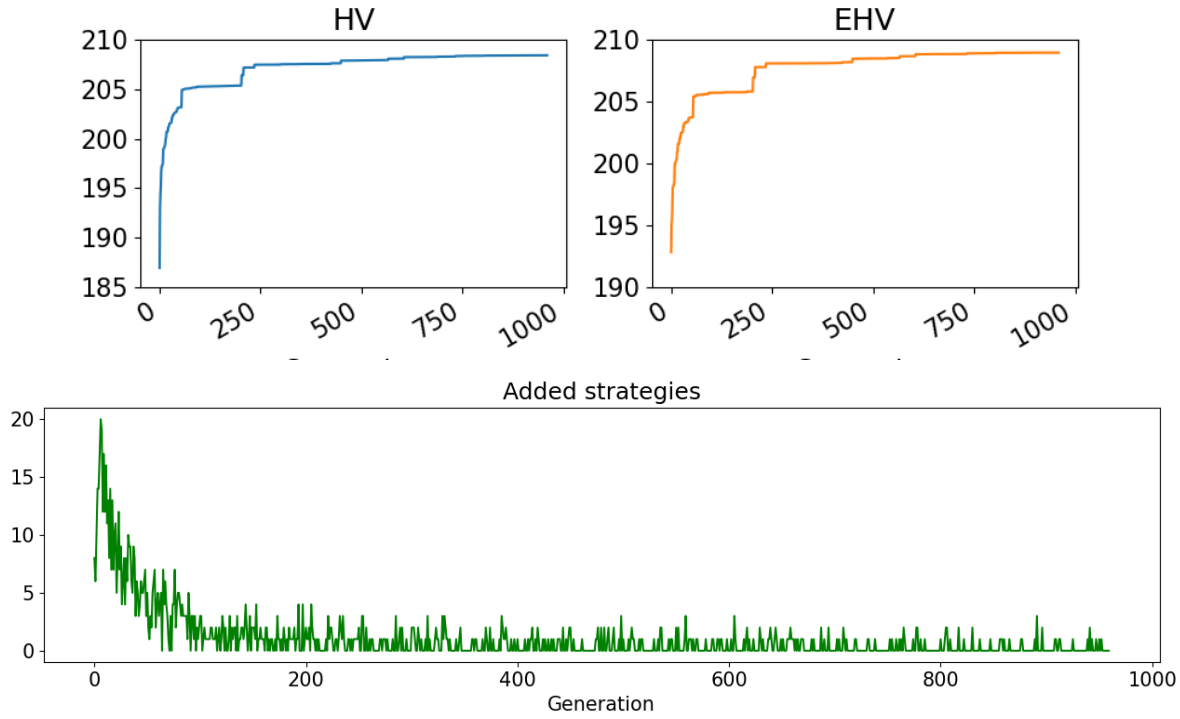

Figure 6.2: Evolution of the genetic algorithm. The upper plots show the evolution of (E)HV over the generations. The bottom plot shows the number of solutions added to the memory Ψ_g in every iteration.

the genetic algorithm took less than a second. The algorithm was stopped after 958 iterations, before it met its stopping criterion. We are unlikely to have found the optimal frontiers of the instance as solutions were still added to the POF in the final iterations and the performance measures were still slightly increasing. The final policies obtained by the algorithm with R^1, collected in the set $\Psi_{958}^{R^1}$, were evaluated once more with a simulation size of 2.5e6 and a different seed for robuster costs and effects.

Figure 6.3 shows the obtained cost-effective frontier $\xi(\Psi_{958}^{R^1})$. It has an EHV of 199.94. To provide a frame of reference for the found policies, the costs and effects of several other good policies are displayed. Two strategies are marked by a plus: the current Dutch screening strategy (the origin) and a situation without screening (costs $\approx -\$500$). Strategies marked by a diamond are the efficient strategies that were advised by Knudsen et al. (2020) to the United States Preventive Task Force (USPTF) in October 2020 and incorporate FITs, colonoscopies or both[2]. These strategies cannot be feasibly generated by our algorithm, for example because screening starts after age 40, ends before age 85 or the interval after a negative colonoscopy is longer than 5 years. Still they do provide us a frame of reference. Strategies marked by a square are trivial strategies that could have been generated by our algorithm. These either

[2]In this study, efficiency was measured in QALYs per colonoscopy, contrasting our efficiency measure QALYs/\$. That is why not all of their policies are efficient with respect to each other in our figure.

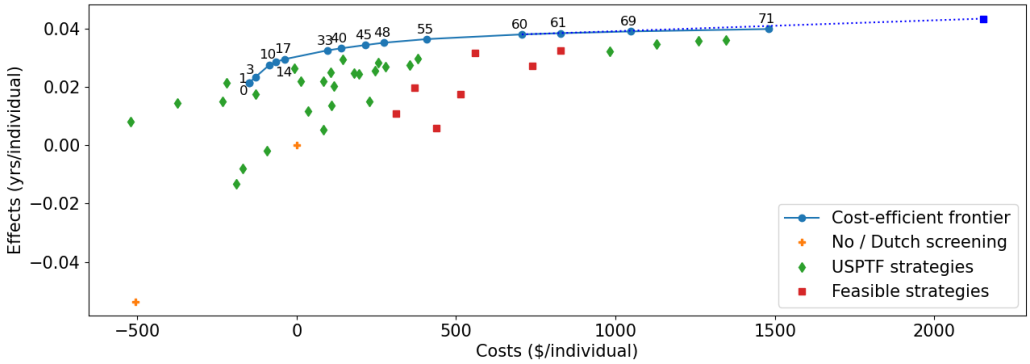

Figure 6.3: Several types of policies displayed in the cost-effectiveness space. The connected dots represent the CEF found by the algorithm. The diamonds indicate the strategies with FIT and colonoscopies recommended to the US Preventive Task Force. The pluses show the current Dutch screening policy (origin) and a situation without screening. The squares indicate strategies that can be feasibly generated by the algorithm. These strategies were annual/biennial/triennial FIT with cut-off either 20 (common in the US) or 47μg/g (common in NL), and a five-yearly colonoscopy, all between ages 40 and 85. Linear combinations of the colonoscopy policy and policy 60 weakly dominate 61, 69 and 71, as illustrated by the dotted line.

prescribe an annual/biennial/triennial FIT with a cut-off of 47μg/g (common in NL) or 20μg/g (common in US) or a five-yearly colonoscopy, all between the ages 40 and 85. We should note that the costs and effects of the all reference strategies were re-evaluated with our unvalidated Hb concentration model: to determine the test result of a FIT, an Hb value was generated by the model from Appendix A and compared with the cut-off of the FIT. As a consequence, the reference strategies may contain the same biases as the strategies found by our algorithm, while the latter are optimised to that bias.

Most of the reference strategies are dominated by our cost-efficient frontier, which shows that the found strategies are of better quality than strategies with non-personalised screening intervals. But not all reference strategies are dominated. The strategy on the right of the spectrum, marked by the blue square, prescribes a five-yearly colonoscopy between ages 40 and 85. Even though it is feasible in the belief space, it was not found by our algorithm. The dotted line in the figure shows that it weakly dominates 61, 69 and 71 in combination with policy 60. This confirms that we have not found the optimum yet after 958 generations. Moreover, there are six non-dominated USPTF policies on the left of the spectrum. These policies all start screening at the age of 50 or 55 and stop screening at 70 or 75. Since our start and stop age are fixed at 40 and 85, these strategies cannot be generated by our algorithm. This suggests that it could be beneficial to relax our start and stop ages.

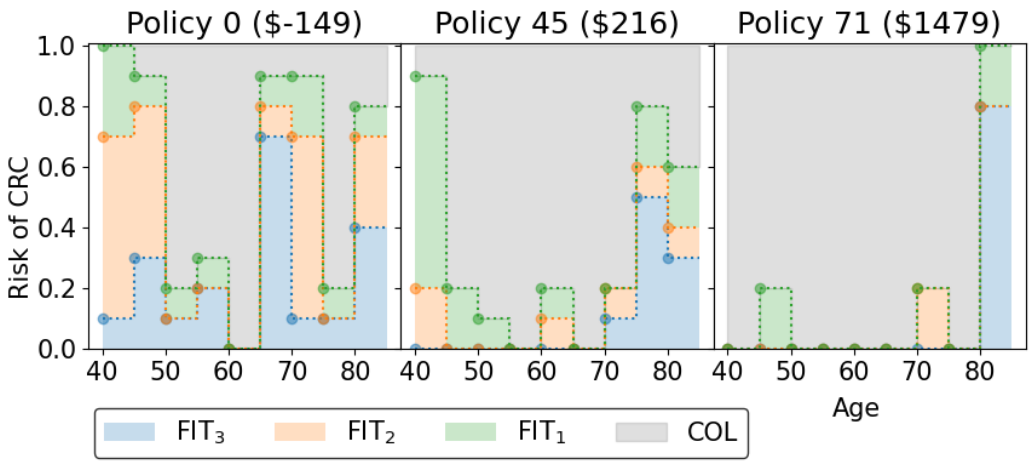

Figure 6.4: From left to right: a policy on the CEF with minimum, medium and maximum costs. The titles show the policy number, corresponding with the number in Figure 6.3, and the costs of the specific policy.

Obtained policies

To analyse the policies in further detail, we show three of them in Figure 6.4: the cheapest, the most expensive and one of medium costs. We observe several features of these policies.

First, they all prescribe at least one guaranteed colonoscopy as the cut-off equals zero for at least one of their age groups. At these ages, everyone gets a colonoscopy independent of their Hb concentration, for example at ages 60 to 64 in policy 0 or 55-59 and 65-69 in policy 45. This is not unique for the three highlighted policies, Figure 6.5 shows that every policy in Ψ_{958}^{R1} has at least one guaranteed colonoscopy. Each circle represents a policy. The horizontal axis shows their costs, the vertical axis represents the number of colonoscopies they guarantee. The value of this line is always greater than zero and increases for increasing costs. Therefore, all obtained policies guarantee at least one colonoscopy to every individual throughout their lives.

The fact that the number of guaranteed colonoscopies increases with costs is in line with previous findings of Dunnewind (2020). However, the fact that even the cheapest policies have at least one guaranteed colonoscopy contrasts his findings: they are only prescribed in policies of medium and high costs. This could be explained by the probability distribution that generates the Hb concentrations. This distribution is highly zero-inflated: an Hb concentration of zero is measured with more than 80% of the healthy patients, 70% of the patients with an adenoma and 6% with a preclinical cancer, while the occurrence of other values is about a factor 10 lower for each of the stages (see also Figure A.1). The algorithm tries to reduce the risk of missing a lesion due to these zeroes by guaranteeing at least one colonoscopy.

Second, we observe that policies 0 and 45 are U-shaped in the sense that cut-offs are high and long screening intervals apply to the ages 40 to 45/50 and 70/80 to 85. We explain this

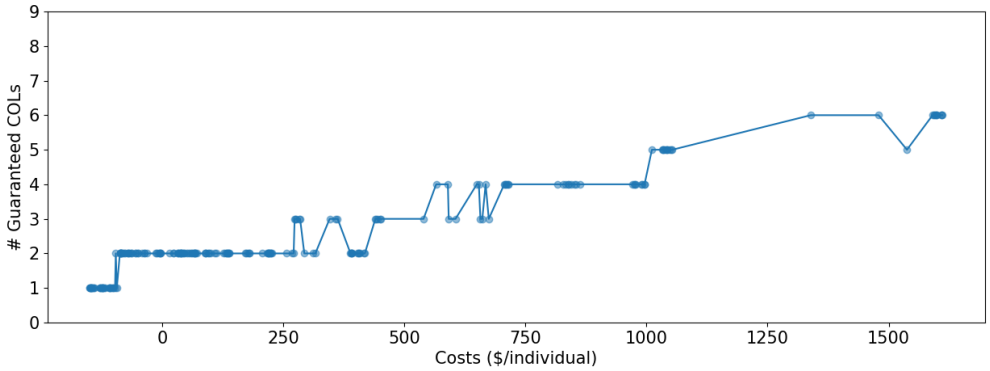

Figure 6.5: Each dot in this graph represents a policy in $\Psi_{958}^{R^1}$. Its costs are shown on the horizontal axis. The vertical axis represents the number of guaranteed colonoscopies in each policy. We observe that each strategy, even the cheapest, guarantee at least one colonoscopy.

epidemiologically. Young patients are unlikely to develop an adenoma or cancer so screening often leads to false positive results. Screening with older patients is not cost-effective as such patients do not have many QALYs to gain. Therefore screening is most effective for the middle ages, causing lower cut-offs and shorter intervals around the age of 60.

However, we also observe that the cut-offs are either above 80μg/g, doubling the cut-off in the current Dutch screening programme, or below 20 for nearly all ages. These extremes are likely to be caused by the skewness of the distribution of the Hb concentrations for patients with a preclinical cancer. Table 1 shows that healthy people and those with an adenoma have a 0.5% and 3% chance respectively of measuring an Hb concentration greater than 100μg/g at the age of 55. This probability is 84% for people with a preclinical cancer. And these probabilities are comparable for 80-years-olds. Furthermore, Figure 6.6 shows the probability of suffering from a certain cancer stage, given the measured Hb concentration. We observe that people that measure a concentration up to 100μg/g, have at least a 70% probability of being healthy. This contrasts Figure 2.1 in which a concentration of 10μg/g corresponds with a 40% probability of having a lesion. This explains why the algorithm uses very low or very high cut-offs. With a low cut-off, almost everyone gets a colonoscopy to detect the adenomas. With high cut-offs, only people that almost certainly have a preclinical cancer are screened.

Table 1: Each entry represents the probability that an estimator obtains a value above 100μg/g, given the age and the disease state of a male patient. H=Healthy, A=Adenoma, PC=Preclinical Cancer.

Estimator	Age	H	A	PC
R^1	55	0.47%	3.0%	84.4%
	80	1.1%	5.0%	85.8%
R^2	55	0.13%	1.8%	96.9%
	80	0.54%	4.3%	97.4%
R^3	55	0.01%	0.3%	80.3%
	80	0.08%	1.0%	82.1%

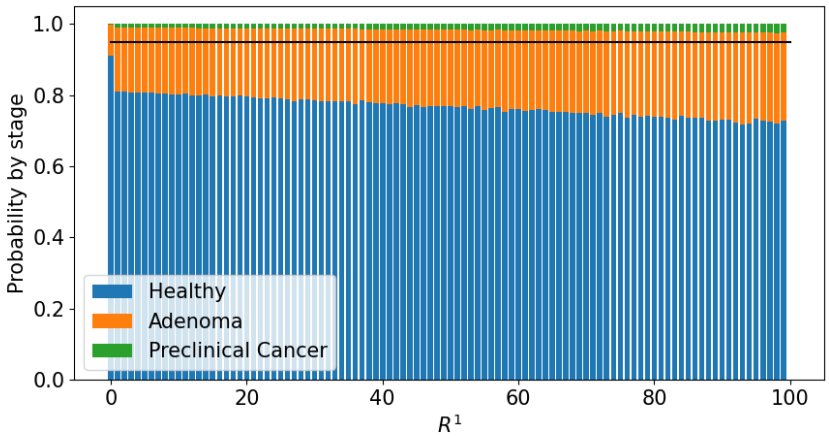

Figure 6.6: This plot shows the probability of being in a certain disease state, given the measured value of the estimator R^1. We observe that, even for values up to 100μg/g, patients have a 70% chance of being healthy. This plot was generated by simulating 1e7 Hb concentrations for 80 years old male for each disease state and applying Bayes' theorem.

6.2.2 Risk estimators R^2 and R^3

Next we tested the effect of using R^2 or R^3 instead of R^1. The algorithm was initialised differently compared to the baseline. We initialised the population of the algorithm with the final population obtained in the baseline (i.e. $P_0^{R^i} = P_{958}^{R^1}$) and its memory with the baseline's final memory ($\Psi_0^{R^i} = \Psi_{958}^{R^1}$ for $i = 2, 3$) to speed up convergence. Since the costs and effects were expected to change with the new estimator, all policies were re-evaluated before the algorithms were started. They were stopped before meeting the stopping criterion after approximately 460 generations. Even though the number of generations is smaller, we can still compare them with the baseline because of the smarter initialisation. After the 460 generations, the final policies obtained were re-evaluated with a population of $2.5e6$ individuals and with a different seed, to increase the robustness of the results.

Figure 6.7 shows the CEFs obtained with the baseline, R^2 and R^3 and Table 2 shows the EHVs of each of them. We observe that the differences between them are minimal. In fact,

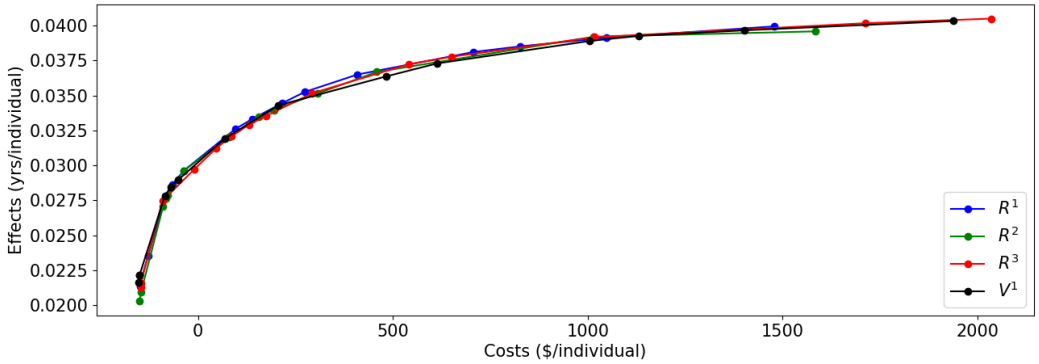

Figure 6.7: The four CEFs obtained by the algorithm with the four different estimators: R^1 (the baseline), R^2, R^3 and V^1 (in combination with R^1).

Table 2: The Extended Hypervolumes of the four CEFs displayed in Figure 6.7.

	R^1	R^2	R^3	V^1
EHV	199.943	198.383	201.462	200.834

none of the CEFs is dominated by another. Apparently, incorporating a larger history of Hb concentrations in the decision process does not improve the quality of the policies.

Obtained policies

For both estimators, three policies from the obtained CEFs are shown in Figure 6.8, chosen such that their costs are comparable to the policies in Figure 6.4. The same trends are observed as in the baseline policies: in general the policies are U-shaped and there is always at least one guaranteed colonoscopy. The cut-offs are either very high or very low due to the fixed start and stop age of the policies and due to the Hb concentration distribution. However, we observe that the highest cut-offs have slightly decreased for policies with R^3, for example for the age group 80-84 in all three policies and the age groups 45-49 and 70-74 in policy 0. Since R^3 incorporates more Hb values, it is robuster to outliers, leading to decreased cut-offs. This is illustrated by Figure 6.9. If $R^3 = 100$, the probability of being healthy is 45% and of having a preclinical cancer is 10%, so the risk of having a lesion is larger if $R^3 = 100$ than if R^2 or R^1 equal 100. This legitimates the decreased cut-offs for 80-years-olds with R^3 compared to R^2.

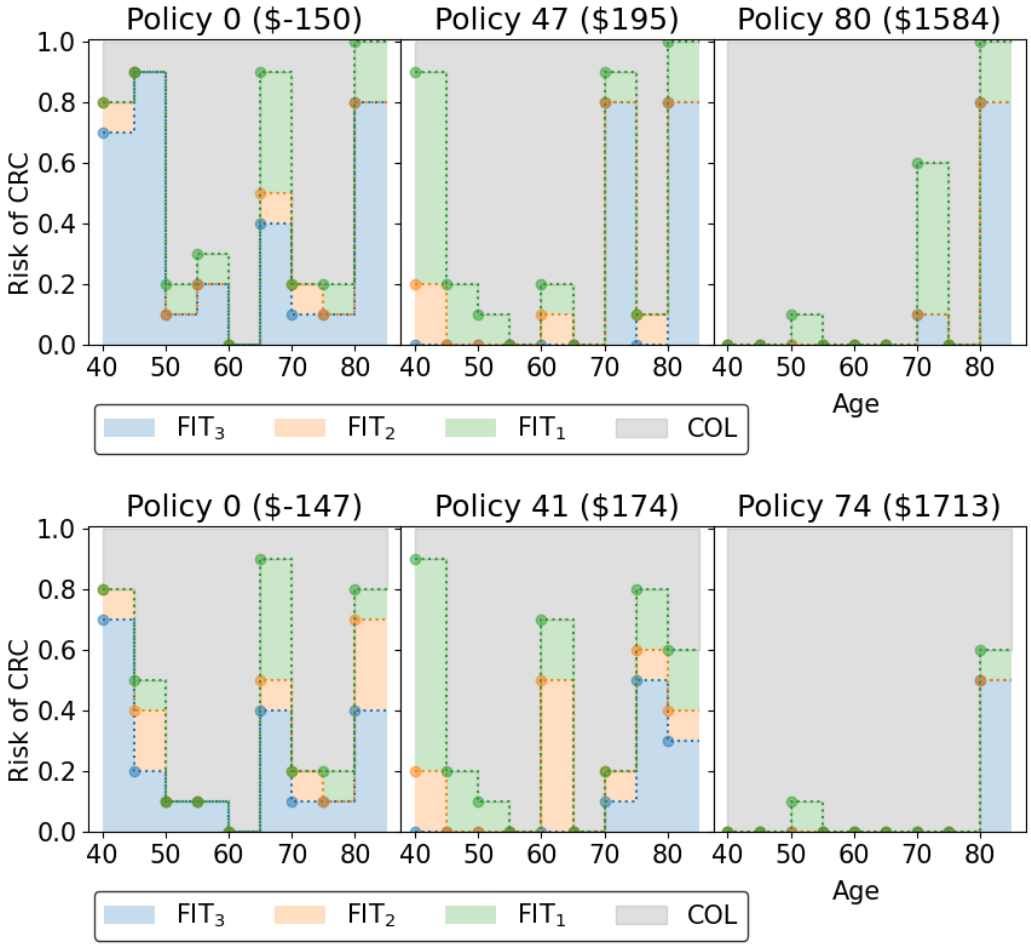

Figure 6.8: Three policies on the CEF that estimate the risk of CRC with R^2 (top) or R^3 (bottom). They are chosen such that their costs are comparable to those of the policies in Figure 6.4.

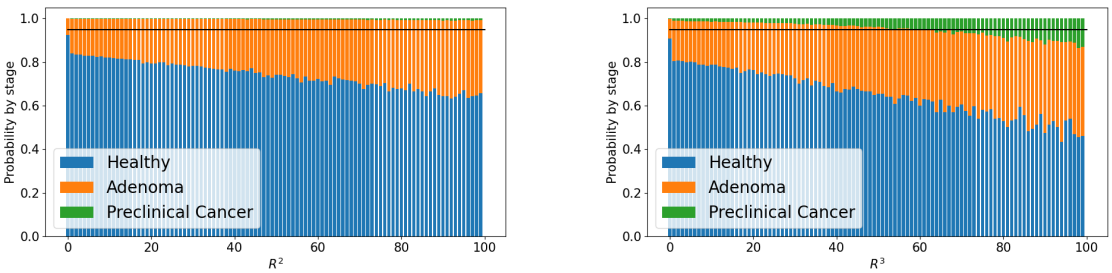

Figure 6.9: The probability of being in a certain disease state, given the value of the estimator R^2 (left) or R^3 (right). These plots were generated for 80 years old male.

6.2.3 Velocity estimator V^1

Finally we test the effect of the three-dimensional belief space in our algorithm by combining velocity estimator V^1 with risk estimator R^1. The $\mathcal{V}$-dimension was discretised in three categories: the change in Hb concentration is smaller than -10μg/g, larger than 10μg/g or in between. This way we can distinguish between large increases, large decreases or negligible changes. This three-dimensional space has $2e66$ feasible policies (see Appendix B).

The algorithm was initialised as described in Section 5.2.1. Since the number of feasible policies is much larger compared to the baseline, it was stopped after 1825 generations instead of 958, without meeting the stopping criterion. The obtained CEF and its EHV is shown in Figure 6.7 and Table 2. Again, the difference with the other CEFs is negligible and this CEF is not dominated by nor dominates any of the other CEFs.

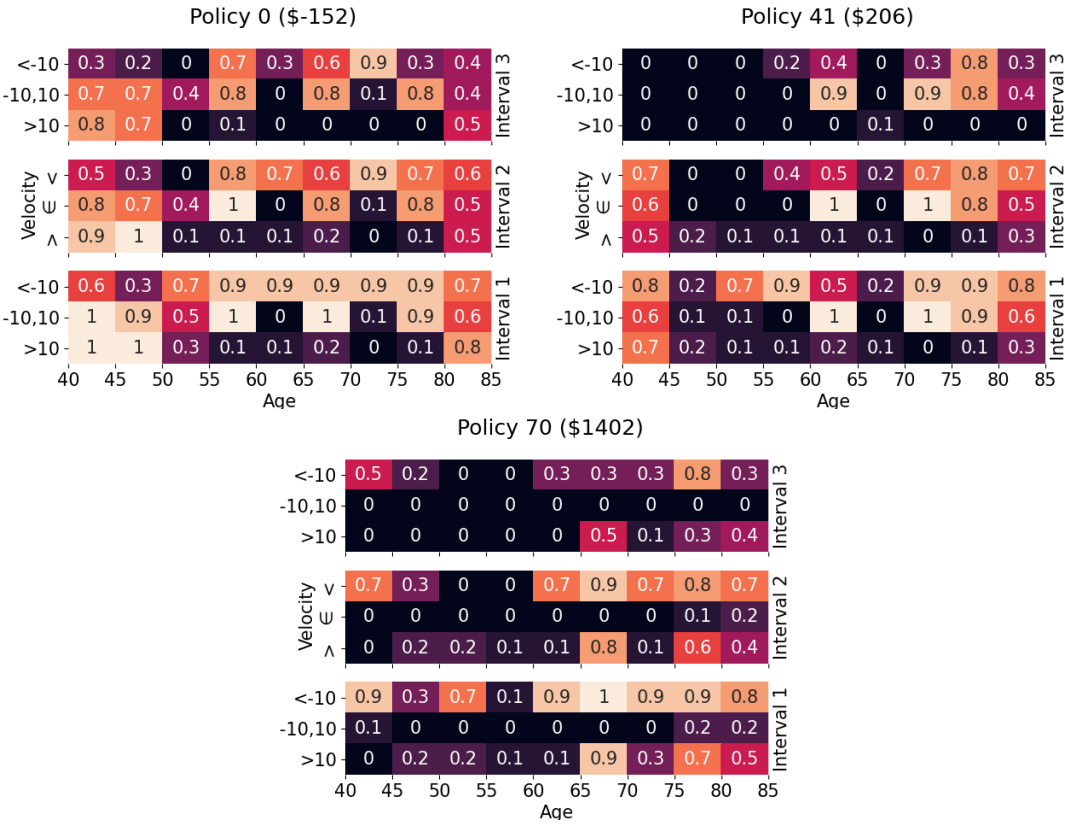

Figure 6.10: Three policies on the CEF that estimate the risk of CRC with V^1, chosen such that their costs are comparable to those of the policies in Figure 6.4. The $\mathcal{V}$-dimension is split into three categories: $V^1 < -10$, $V^1 \in [-10, 10]$ and $V^1 > 10\mu$g/g.

Obtained policies

Figure 6.10 shows three policies from the CEF with costs comparable to the policies in Figure 6.4. First we consider the middle row of each of the intervals, corresponding with $V^1 \in [-10, 10]$.

We observe some of the aforementioned characteristics of the two-dimensional policies: policies 0 and 29 assign high cut-offs and longer intervals to the youngest and oldest age groups and all policies, even the cheapest, have at least one guaranteed colonoscopy. The expensive policy has a lot of age groups with a cut-off of zero. We observe a lot of cut-offs above 70 and below 30μg/g.

We note that the third dimension is exploited by the policy. Policy 0, which guarantees a colonoscopy at the ages 60-64. 5 years later, the cut-off is very high (100μg/g) if the Hb concentrations value remains constant, which is similar to the two dimensional policies. However, if the concentration has increased significantly, five years after the colonoscopy, the cut-off is 20: almost all people with significant increases are screened once again. This way, the algorithm increases the screening intensity for high-risk people.

On the other hand, there are cases that the policy reduces the screening intensity with the third dimension. In policy 0, the cut-off is very high for significant decreases at the ages 60-64. Therefore, if a patient gets an Hb concentration of 90μg/g at the age of 59 and 80μg/g at the age of 60, the "guaranteed colonoscopy" is skipped. This also occurs in policy 70 for people aged 60-74: a significant decrease in concentration results in skipping a guaranteed colonoscopy. This way the algorithm deals with outliers in the measured Hb values.

We can conclude that the policies generated with R^3 and V^1 show some signs that they exploit the properties of the estimator, such as velocity dimension. However, this is not reflected in the quality of the obtained CEFs: none of them dominates another. This may be caused by the Hb concentration distribution. Appendix A describes that in this distribution, every patient is assigned a fixed individual risk factor γ, which influences the mean of the distribution and therefore ensures that Hb values within a patient are correlated. This distribution is zero-inflated, which means that a large proportion of the values is assigned a value 0. However, this proportion only depends on patients' disease stage, not on their individual risk factor. The zero-inflated part is responsible for approximately 4 out of 5 or 2 out of 3 of the simulated Hb concentrations for patients in the healthy or adenoma state respectively. The used estimators take actions based on at most three past concentrations and therefore they cannot capture the effects caused by γ. Consequently, the algorithm is unable to derive if patients have a systematically high Hb value, even though they are healthy, or whether the high value caused by a cancer or adenoma.

7 Discussion

This thesis proposes an algorithm to personalise stool-based screening strategies. Based on haemoglobin (Hb) concentrations in patients' stool, measured by the Faecal Immunochemical Tests (FITs), strategies prescribe whether patients require further screening in a hospital with a (near-)perfect colonoscopy or within what time interval they should return for another FIT. High-quality screening strategies are characterised by low costs and many Quality Adjusted Life Years gained. To generate such strategies, we combine the concept of the belief space, adapted from Partially Observed Markov Decision Processes, with a genetic algorithm.

Using two different performance measures, we tested what combination of parameter values and operators worked out best on a small instance of the belief space. The algorithm achieved optimality gaps smaller than 0.7% for most of these settings. We selected one of the best performing settings for the remainder of our thesis.

To convert the measured Hb concentrations to a risk value for developing CRC, two types of estimators were tested. R^k maps the mean of the past k concentrations to a probability and V^k represents the net change in Hb concentration over the past $k+1$ measurements. With our algorithm, we searched for cost-efficient strategies for the estimators R^1, R^2, R^3 and a combination of V^1 and R^1.

The generated policies show some interesting characteristics. First, the policies using R^1 (weakly) dominate most of the screening strategies recommended to the United States Preventive Task Force (USPTF) by Knudsen et al. (2020) and (variations of) the current Dutch screening strategy. It seems that personalised screening strategies improve the quality of a screening programme. However, we should be careful with this conclusion. The costs and effects of the reference policies were also evaluated with our Hb concentrations model and may therefore have some bias. In that case the strategies from the algorithm may perform better because they are optimised to that bias.

Second, variation in the estimators does not seem to improve the quality of the policies. The cost-efficient frontiers obtained with the four different estimators merely differ, while we expected that including more observations would give robuster policies. We expect that this is caused by the model that simulates the measured Hb concentrations. Even though the model includes a factor responsible for correlation between the values measured with the same person over time, this factor is used in a minority of the drawn values. Therefore a good estimator would need to incorporate more than three past Hb concentrations to base its decision on.

Third, we observe that in general, the policies either prescribe very low or very high cut-offs, usually between 0 and 20μg/g or between 60 and 100μg/g. High cut-offs especially happen at

older and younger ages which suggests that we should relax our fixed start and stop ages of 40 and 85 for our screening policies. This is also in line with the USPTF strategies: the six of them that were not dominated by the strategies obtained with our algorithm prescribe screening for people aged around 50 to 70 only. However, the fact that these cut-offs are either very low or very high and that this also occurs at other ages, shows that it is also caused by the used Hb concentration model. We have shown that if the estimators equal 100μg/g, a value twice as high as the current cut-off in the Dutch screening programme, patients still have a probability of at least 45% to be healthy. This contrasts the findings by Grobbee et al. (2017).

This is also confirmed by the fact that every policy, even the cheapest, prescribe at least one colonoscopy to the full population, independent of the measured Hb concentration. This contrasts the finding of Dunnewind (2020) in which such guaranteed colonoscopies are only prescribed in medium and high cost policies. However, due to the Hb concentration model, more than 70% of the people with an adenoma obtain an Hb value equal to zero in our model. Therefore screening every individual with a colonoscopy at least once during their lives ensures that the adenomas that would remain undetected by FITs are removed.

7.1 Limitations and suggestions

This study has several limitations. First, the quality of our output strategies hinges on the quality of the simulation model for the Hb concentrations. A lot of the characteristics of our policies are explained with this model. Therefore the quality of our conclusions is highly influenced by the quality of this model. Since our algorithm uses a preliminary, unvalidated model to simulate these concentrations and we have seen that some of its results conflict the observations of Grobbee et al. (2017), it is unlikely that our policies are optimal in practice. Therefore we emphasize the need for further research on and further calibration of the Hb concentration model to make strong conclusions.

Second, the risk estimators used in the method plays an important role. This thesis explored simple estimators and has shown that these find policies of equal quality. However, more sophisticated prediction models, such as a Cox model, which incorporate all measured Hb concentrations and the intervals between these measurements, may significantly increase the quality of policies. Especially if the Hb concentration model is improved. Such models also waive the need for a third dimension in the belief space, which reduces the amount of feasible policies by drastically, leading to faster convergence. Therefore, our second suggestion for future research concerns finding a suitable risk prediction model. A more sophisticated model, however, may lead to implementation problems in, for example, the United States where the absence

of a centralised screening system requires doctors to track their patient's past Hb concentrations themselves. A sophisticated model leads to a sophisticated decision process by the doctor, possibly leading to erroneous advices.

Third, we are aware of some limitations of the algorithm developed in this thesis. Parameter values and operators were chosen by their performance on a small instance which was solved to near-optimality by almost all parameter settings. This suggests that the instance may have been too small for a fair comparison. Also, the performance was measured with two very similar performance measures, while at least three different types of performance measures are preferred (Zitzler et al., 2003). Therefore we may not have chosen one of the better parameter sets. We are also aware that the algorithm was stopped before meeting the convergence criterium. Therefore we are unlikely to have found the most cost-efficient strategies. Furthermore, the NSGA-II algorithm prefers strategies with a large crowding distance to maintain a widespread POF. However, this prevents the algorithm from focussing on a part of the cost-effectiveness space in detail and improving the strategies in that part, leading to a local optimum. Especially because in practice, we are not necessarily interested in the most expensive strategies future researchers may choose to assign a better fitness to cheaper strategies such that the left part of the cost-effectiveness space is explored well.

We also suggest some improvements to the algorithm. Currently we do not explicitly model that an individual has undergone a colonoscopy with a negative result. This is valuable information to the algorithm as these patients are likely to have increased Hb concentrations without having a lesion. By incorporating this, for example, as a third dimension in the belief space, we could improve the quality of the obtained strategies.

Next, we also suggest to relax the fixed start and stop ages of the screening strategies. In our algorithm, they were fixed at 40 and 85 respectively. However, most of our strategies prescribe high cut-offs for 80 to 85-years-olds, and the non-dominated USPTF strategies were those that started screening at the age of 45 or 50 and ended at 70 or 75. The quality might improve if people are not screened at all at these ages.

Finally, MISCAN can provide more information than the costs and effects of a strategy. For example, the age at which patients die from cancer or the number of preclinical cancers in a population at a certain point in time. Therefore we suggest to enhance the operators of the genetic algorithm with this extra information, especially since their running times are negligible to that of MISCAN. For example, if many 60-years-olds have a preclinical cancer, the mutation operator should mutate the strategy at the younger ages such that more cancers are detected. Such enhanced operators could result in better screening policies.

References

Ahuja, K., Zame, W., and van der Schaar, M. (2017). Dpscreen: Dynamic personalized screening. In *Advances in Neural Information Processing Systems*, pages 1321–1332.

Alaswad, S. and Xiang, Y. (2017). A review on condition-based maintenance optimization models for stochastically deteriorating system. *Reliability Engineering & System Safety*, 157:54–63.

Audet, C., Bigeon, J., Cartier, D., Le Digabel, S., and Salomon, L. (2020). Performance indicators in multiobjective optimization. *European journal of operational research*.

Ayer, T., Alagoz, O., and Stout, N. K. (2012). Or forum—a pomdp approach to personalize mammography screening decisions. *Operations Research*, 60(5):1019–1034.

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 68(6):394–424.

Brenner, H. and Werner, S. (2017). Selecting a cut-off for colorectal cancer screening with a fecal immunochemical test. *Clinical and translational gastroenterology*, 8(8):e111.

Byon, E. and Ding, Y. (2010). Season-dependent condition-based maintenance for a wind turbine using a partially observed markov decision process. *IEEE Transactions on Power Systems*, 25(4):1823–1834.

Clemen, R. T. and Lacke, C. J. (2001). Analysis of colorectal cancer screening regimens. *Health care management science*, 4(4):257–267.

Day, L. W., Bhuket, T., and Allison, J. (2013). Fit testing: an overview. *Current gastroenterology reports*, 15(11):357.

Deb, K., Agrawal, S., Pratap, A., and Meyarivan, T. (2000). A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: Nsga-ii. In *International conference on parallel problem solving from nature*, pages 849–858. Springer.

Dunnewind, N. (2020). Finding cost-effective colorectal cancer screening strategies using multi-objective evolutionary algorithms and the miscan-colon microsimulation model. Master's thesis.

Erenay, F. S., Alagoz, O., and Said, A. (2014). Optimizing colonoscopy screening for colorectal cancer prevention and surveillance. *Manufacturing & Service Operations Management*, 16(3):381–400.

Frazier, A. L., Colditz, G. A., Fuchs, C. S., and Kuntz, K. M. (2000). Cost-effectiveness of screening for colorectal cancer in the general population. *Jama*, 284(15):1954–1961.

Gini, A. (2020). *Microsimulation Models to Inform Colorectal Cancer Screening Decisions: From validated tools to tailoring recommendations*. PhD thesis.

Goode, K., Roylance, B., and Moore, J. (2000). Development of model to predict condition monitoring interval times. *Ironmaking & Steelmaking*, 27(1):63–68.

Grobbee, E. J., Schreuders, E. H., Hansen, B. E., Bruno, M. J., Lansdorp-Vogelaar, I., Spaander, M. C., and Kuipers, E. J. (2017). Association between concentrations of hemoglobin determined by fecal immunochemical tests and long-term development of advanced colorectal neoplasia. *Gastroenterology*, 153(5):1251–1259.

Habbema, J., Van Oortmarssen, G., Lubbe, J. T. N., and Van der Maas, P. (1985). The miscan simulation program for the evaluation of screening for disease. *Computer methods and programs in biomedicine*, 20(1):79–93.

Holme, O., Bretthauer, M., Fretheim, A., Odgaard-Jensen, J., and Hoff, G. (2013). Flexible sigmoidoscopy versus faecal occult blood testing for colorectal cancer screening in asymptomatic individuals. *Cochrane Database of Systematic Reviews*, (9).

Jardine, A. K., Lin, D., and Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical systems and signal processing*, 20(7):1483–1510.

Knudsen, A. B., Rutter, C. M., Peterse, E. F. P., Lietz, A. P., Seguin, C. L., Meester, R. G. S., Perdue, L. A., Lin, J. S., Siegel, R. L., Zauber, A. G., Kuntz, K. M., and Lansdorp-Vogelaar, I. (2020). Colorectal cancer screening: A decision analysis for the u.s. preventive services task force.

Kurt, M. and Kharoufeh, J. P. (2010). Monotone optimal replacement policies for a markovian deteriorating system in a controllable environment. *Operations Research Letters*, 38(4):273–279.

Lansdorp-Vogelaar, I., Knudsen, A. B., and Brenner, H. (2011). Cost-effectiveness of colorectal cancer screening. *Epidemiologic reviews*, 33(1):88–100.

Leslie, A., Carey, F., Pratt, N., and Steele, R. (2002). The colorectal adenoma–carcinoma sequence. *British Journal of Surgery*, 89(7):845–860.

Li, M. and Yao, X. (2019). Quality evaluation of solution sets in multiobjective optimisation: A survey. *ACM Computing Surveys (CSUR)*, 52(2):1–38.

Loeve, F., Boer, R., van Oortmarssen, G. J., van Ballegooijen, M., and Habbema, J. D. F. (1999). The miscan-colon simulation model for the evaluation of colorectal cancer screening. *Computers and Biomedical Research*, 32(1):13–33.

Maillart, L. M., Ivy, J. S., Ransom, S., and Diehl, K. (2008). Assessing dynamic breast cancer screening policies. *Operations Research*, 56(6):1411–1427.

Marseguerra, M., Zio, E., and Podofillini, L. (2002). Condition-based maintenance optimization by means of genetic algorithms and monte carlo simulation. *Reliability Engineering & System Safety*, 77(2):151–165.

Otten, J., Witteveen, A., Vliegen, I., Siesling, S., Timmer, J. B., and IJzerman, M. J. (2017). Stratified breast cancer follow-up using a partially observable mdp. In *Markov decision processes in practice*, pages 223–244. Springer.

Ozik, J., Collier, N. T., Wozniak, J. M., and Spagnuolo, C. (2016). From desktop to large-scale model exploration with swift/t. In *2016 Winter Simulation Conference (WSC)*, pages 206–220. IEEE.

Parmigiani, G. (1997). Timing medical examinations via intensity functions. *Biometrika*, 84(4):803–816.

Peng, Y., Dong, M., and Zuo, M. J. (2010). Current status of machine prognostics in condition-based maintenance: a review. *The International Journal of Advanced Manufacturing Technology*, 50(1-4):297–313.

Pignone, M., Saha, S., Hoerger, T., and Mandelblatt, J. (2002). Cost-effectiveness analyses of colorectal cancer screening: a systematic review for the us preventive services task force. *Annals of internal medicine*, 137(2):96–104.

Preston, A. J. and Smith, W. (2001). Disease screening designs: sensitivity and screening frequency. In *Proceedings of the annual meeting of the American statistical association*, pages 5–9.

Riquelme, N., Von Lücken, C., and Baran, B. (2015). Performance metrics in multi-objective optimization. In *2015 Latin American Computing Conference (CLEI)*, pages 1–11. IEEE.

Robertson, D. J., Lee, J. K., Boland, C. R., Dominitz, J. A., Giardiello, F. M., Johnson, D. A., Kaltenbach, T., Lieberman, D., Levin, T. R., and Rex, D. K. (2017). Recommendations on fecal immunochemical testing to screen for colorectal neoplasia: a consensus statement by the us multi-society task force on colorectal cancer. *Gastroenterology*, 152(5):1217–1237.

Schreuders, E. H., Ruco, A., Rabeneck, L., Schoen, R. E., Sung, J. J., Young, G. P., and Kuipers, E. J. (2015). Colorectal cancer screening: a global overview of existing programmes. *Gut*, 64(10):1637–1649.

Tomer, A., Nieboer, D., Roobol, M. J., Steyerberg, E. W., and Rizopoulos, D. (2019). Personalized schedules for surveillance of low-risk prostate cancer patients. *Biometrics*, 75(1):153–162.

Van Gils, P., Schoenmaker, C., and Polder, J. (2013). Hoeveel mag een gewonnen levensjaar kosten? onderzoek naar de waardering van de qaly. *Nederlands Tijdschrift voor Geneeskunde*, 157.

van Noortwijk, J. M. (2009). A survey of the application of gamma processes in maintenance. *Reliability Engineering & System Safety*, 94(1):2–21.

Wang, W. (2012). A simulation-based multivariate bayesian control chart for real time condition-based maintenance of complex systems. *European Journal of Operational Research*, 218(3):726–734.

Zauber, A. G., Knudsen, A. B., Rutter, C. M., Lansdorp-Vogelaar, I., Savarino, J. E., Kuntz, K., et al. (2009). Cost-effectiveness of ct colonography to screen for colorectal cancer.

Zhang, M., Ye, Z., and Xie, M. (2014). A condition-based maintenance strategy for heterogeneous populations. *Computers & Industrial Engineering*, 77:103–114.

Zhao, X., Fouladirad, M., Bérenguer, C., and Bordes, L. (2010). Condition-based inspection/replacement policies for non-monotone deteriorating systems with environmental covariates. *Reliability Engineering & System Safety*, 95(8):921–934.

Zitzler, E. and Künzli, S. (2004). Indicator-based selection in multiobjective search. In *International conference on parallel problem solving from nature*, pages 832–842. Springer.

Zitzler, E. and Thiele, L. (1999). Multiobjective evolutionary algorithms: a comparative case study and the strength pareto approach. *IEEE transactions on Evolutionary Computation*, 3(4):257–271.

Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C. M., and Da Fonseca, V. G. (2003). Performance assessment of multiobjective optimizers: An analysis and review. *IEEE Transactions on evolutionary computation*, 7(2):117–132.

Appendices

A Probability distribution of haemoglobin concentrations

As described in Section 2, the FIT test returns the haemoglobin concentration in a patient's stool. To incorporate this in our simulation, we developed a probability distribution for the Hb concentrations.

The Hb concentration depends on the a patient's age, gender, and individual risk factor and the cancer stage in $\mathcal{X}^u$ the patient is in. The Hb concentrations were modelled by a mixed-effects zero-inflated negative binomial model (ZINB), expressed by:

$$\mathbb{P}(S_n|\boldsymbol{x}_n) \sim \begin{cases} NB(S_n|\boldsymbol{x}_n), & \text{if } u \geq \varphi(x\_STAGE) \\ 0, & \text{otherwise.} \end{cases} \tag{A.1}$$

Here S_n is the Hb value obtained at the n^{th} test and $\boldsymbol{x}_n = [x\_\text{AGE}, x\_\text{GENDER}, x\_\text{STAGE}]$ is the vector that contains the age, gender and cancer stage of the patient respectively at the moment of the test. The model is called *zero-inflated* because it contains a lot of zeroes compared to a regular negative binomial distribution. The number of zeroes is determined by the value $\varphi(x) \in [0, 1]$ which depends on the patient's current state $x \in \mathcal{X}^u$. A standard uniform random variable u is drawn at every FIT to determine whether the new Hb value is a zero or is to be drawn from the NB distribution.

The pdf of this NB distribution is specified as follows:

$$f(S_n|\boldsymbol{x}_n) = \frac{\Gamma(S_n + \theta)}{S_n!\Gamma(\theta)} * \left(\frac{\theta}{\theta + \mu_n}\right)^{\theta} * \left(\frac{\mu_n}{\theta + \mu_n}\right)^{S_n} \tag{A.2}$$

with θ the dispersion factor, constant throughout the algorithm, and $\Gamma(.)$ the gamma function. The mean μ_n is defined by

$$\begin{aligned} \mu_n = \mu(\boldsymbol{x}_n) = \exp\{&b_{\text{INTERCEPT}} + \gamma + b_{\text{AGE}} * x\_\text{AGE} + b_{\text{MALE}} * I(x\_\text{GENDER} = \text{MALE}) + \\ &b_{\text{A}} * I(x\_\text{STAGE} = \text{A}) + b_{\text{PC}} * I(x\_\text{STAGE} = \text{PC})\}. \end{aligned} \tag{A.3}$$

Here b_i stands for the weights of each of the properties in $\boldsymbol{x}_n$ and the intercept, I(.) represents the indicator function which equals one if the statement between brackets holds true and zero otherwise. A and PC stand for the adenoma and preclinical stages respectively. γ is an individual risk factor, a random value that is drawn for all patients individually and remains constant during their lives. It is responsible for the correlation between the Hb concentrations obtained in the

Table 3: Overview of the variables in the model for Hb concentrations and their values.

Symbol	Description	Value
$\varphi(H)$	Probability of inflated zero for healthy stage	0.840
$\varphi(A)$	Probability of inflated zero for adenoma stage	0.644
$\varphi(PC)$	Probability of inflated zero for preclinical stage	0.032
$b_{\text{INTERCEPT}}$	Intercept	1.481
b_{AGE}	Weight of age factor	0.0181
b_{MALE}	Weight of gender factor	0.2832
b_{A}	Weight of adenoma stage	0.609
b_{PC}	Weight of preclinical stage	8.74
θ	Dispersion factor	0.262
σ	Standard deviation of individual risk factor distribution	0.1120

same patients. Its distribution is $\gamma \sim \mathcal{N}(0, \sigma)$.

Table 3 gives an overview of the calibrated values. Figure A.1 displays histograms of the theoretical Hb concentration distribution for each of the three stages for a fixed age and gender. Note that the distribution is highly zero-inflated, especially for the healthy and adenoma stages. Therefore, the histograms are also shown for larger Hb values. In turn, the distribution for the preclinical stage is highly skewed: the three theoretical distributions generate maximum Hb values of 632, 1406 and 4968590 respectively.

Figure A.2 shows the Hb value distribution for a given cancer stage as observed in the Dutch screening programme. We observe that the theoretical distribution needs further calibration, especially the preclinical cancer stage, since it is not similar to the observed distribution. This is also confirmed by the maxima of the distributions mentioned above as the observed Hb values merely exceed 300μg/g.

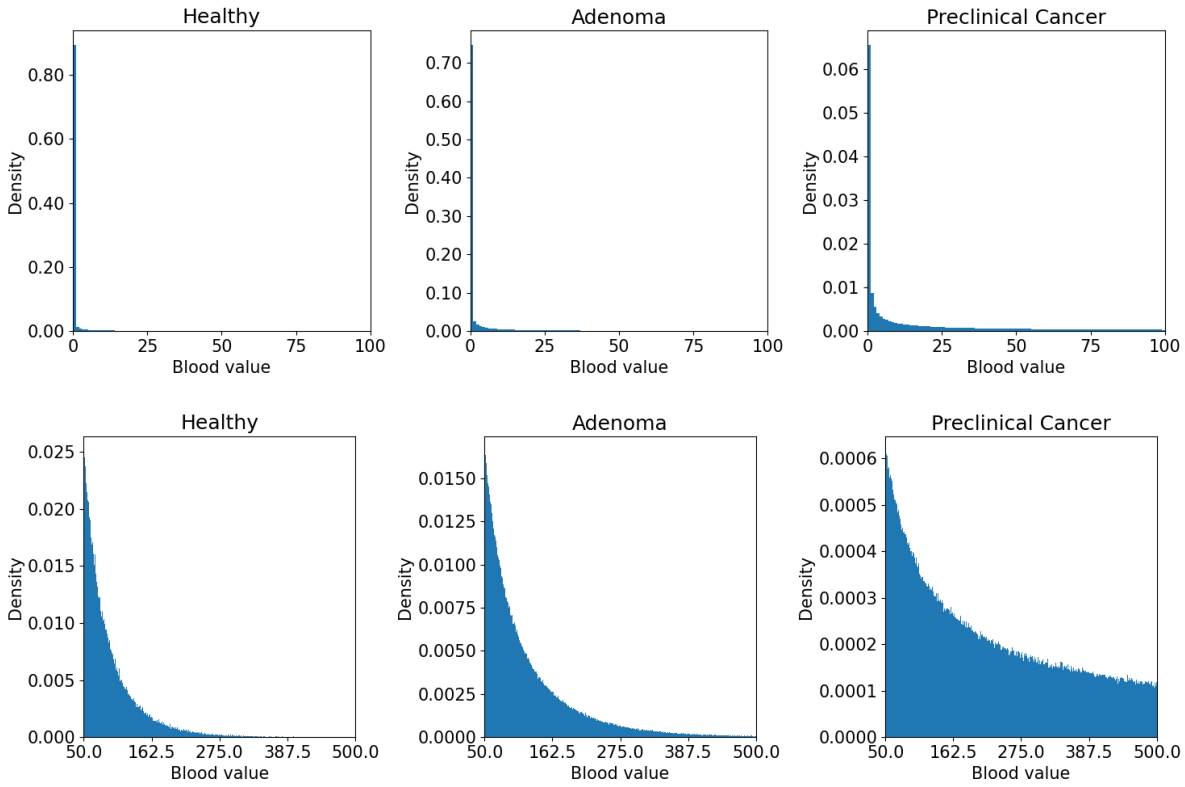

Figure A.1: Histograms of the simulated ZINB-distributions for each of the three cancer stages with the given parameters. The upper plot displays the Hb concentrations between 0 and 100, the bottom plot shows them between 50 and 500. The age is set constant at 55 and gender is fixed at male. Each of the histograms was generated with 1e7 individuals.

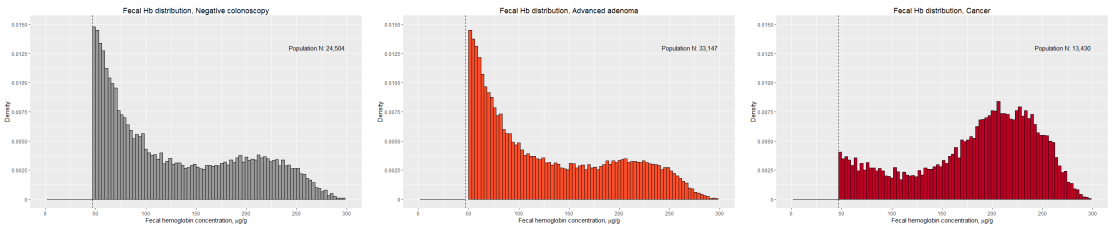

Figure A.2: Histograms of the observed Hb concentration given the result of a consequent colonoscopy in the Dutch screening programme. From left to right, the states healthy, adenoma and preclinical cancer are displayed. The horizontal axis represents the observed Hb value, ranging from 50 to 300μg/g. The vertical axes show the density, ranging from 0 to 0.0150. Only Hb values above 50μg/g are shown: we only observe a patient's cancer stage with a colonoscopy after the patient tested positive, i.e. obtained a concentration above 47μg/g.

B Number of feasible strategies in the belief space instances

Two instances of the belief spaces are used in this thesis, a small, a large and a three-dimensional one. This appendix shows how the number of feasible policies in these spaces is calculated.

First we consider one age τ only in the small instance. If there were one screening interval only, i.e. $\mathcal{A} = \{\text{FIT}_1,\text{COL}\}$, there are 5 unique screening policies: since $\mathcal{R} = \{0, 0.25, 0.5, 0.75, 1\}$, the single policy border $\boldsymbol{\beta}_{FIT_1}$ can take five values. When adding an extra interval, we need to account for the impact ordering assumption. If $\boldsymbol{\beta}_{FIT_2} = 0$, $\boldsymbol{\beta}_{FIT_1}$ can take any of the five values. If $\boldsymbol{\beta}_{FIT_2} = 0.25$, $\boldsymbol{\beta}_{FIT_1}$ can take the four values equal or greater than 0.25. If $\boldsymbol{\beta}_{FIT_2} = 0.5$, $\boldsymbol{\beta}_{FIT_1}$ can take the three values equal or greater than 0.5. This leads to $5+4+3+2+1 = 15$ options.

When adding a third interval, the above calculation applies to the case that $\boldsymbol{\beta}_{FIT_3} = 0$: there are 15 feasible policies in that case. If $\boldsymbol{\beta}_{FIT_3} = 0.25$, there are $4+3+2+1 = 10$ policies. In the end, this leads to $15 + 10 + 6 + 3 + 1 = 35$ options for one τ with three intervals. Since there are four independent ages in the small belief space, there are $35^4 = 1,500,625$ unique policies.

Generalizing the above calculation, we can calculate the number of feasible policies in the two-dimensional belief space instances by

$$\left(\sum_{k=1}^{|\mathcal{R}|}\sum_{j=1}^{k} j\right)^{|\mathcal{T}|}$$

and that of the three-dimensional instance by

$$\left(\sum_{k=1}^{|\mathcal{R}|}\sum_{j=1}^{k} j\right)^{|\mathcal{T}\times\mathcal{V}|},$$

assuming the action space consists of three screening intervals.

In the large belief space instance, $|\mathcal{R}| = 11$ and $|\mathcal{T}| = 9$. Therefore it has $286^9 = 12.8e21$ feasible policies. In the three-dimensional instance, $|\mathcal{V}| = 3$ which leads to $286^{27} = 2.1e66$ feasible policies.

C Graphs for parameter tuning and operator selection

In Section 6.1 we showed how the parameter values and operators were selected. The following figures show the performance of the algorithm for several combinations of parameters and operators.

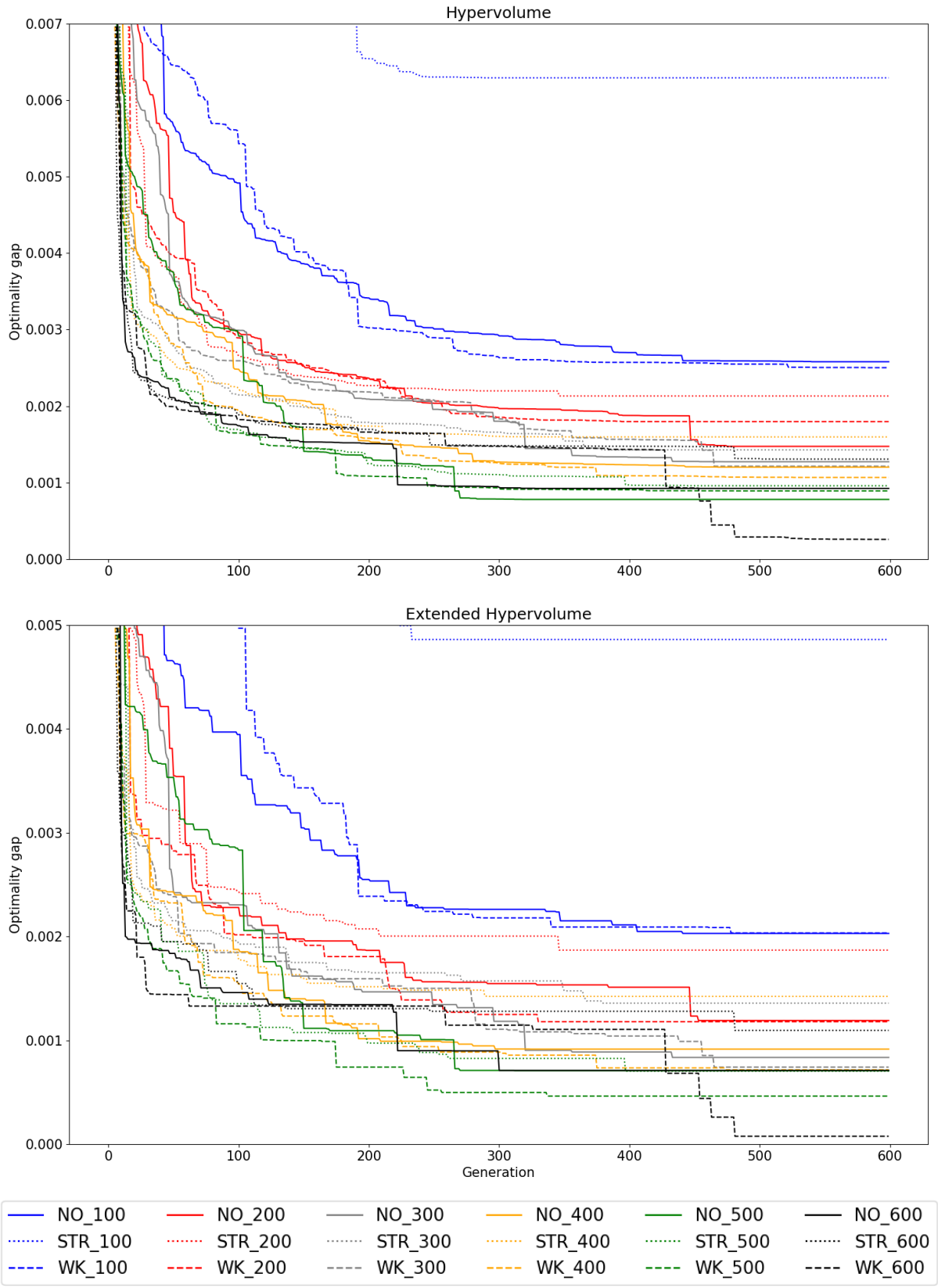

Figure C.1: Evolution of the average of the performance measures per selection operator and population size. The upper and lower plot display the HV and EHV resp. The vertical axis displays the optimality gap from the known POF/CEF, relative to the known optimum. The three tested selection operators are no elitism (NO), strong elitism (STR) or weak elitism (WK). The legend indicates the selection operator and N_{pop} separated by a "_".

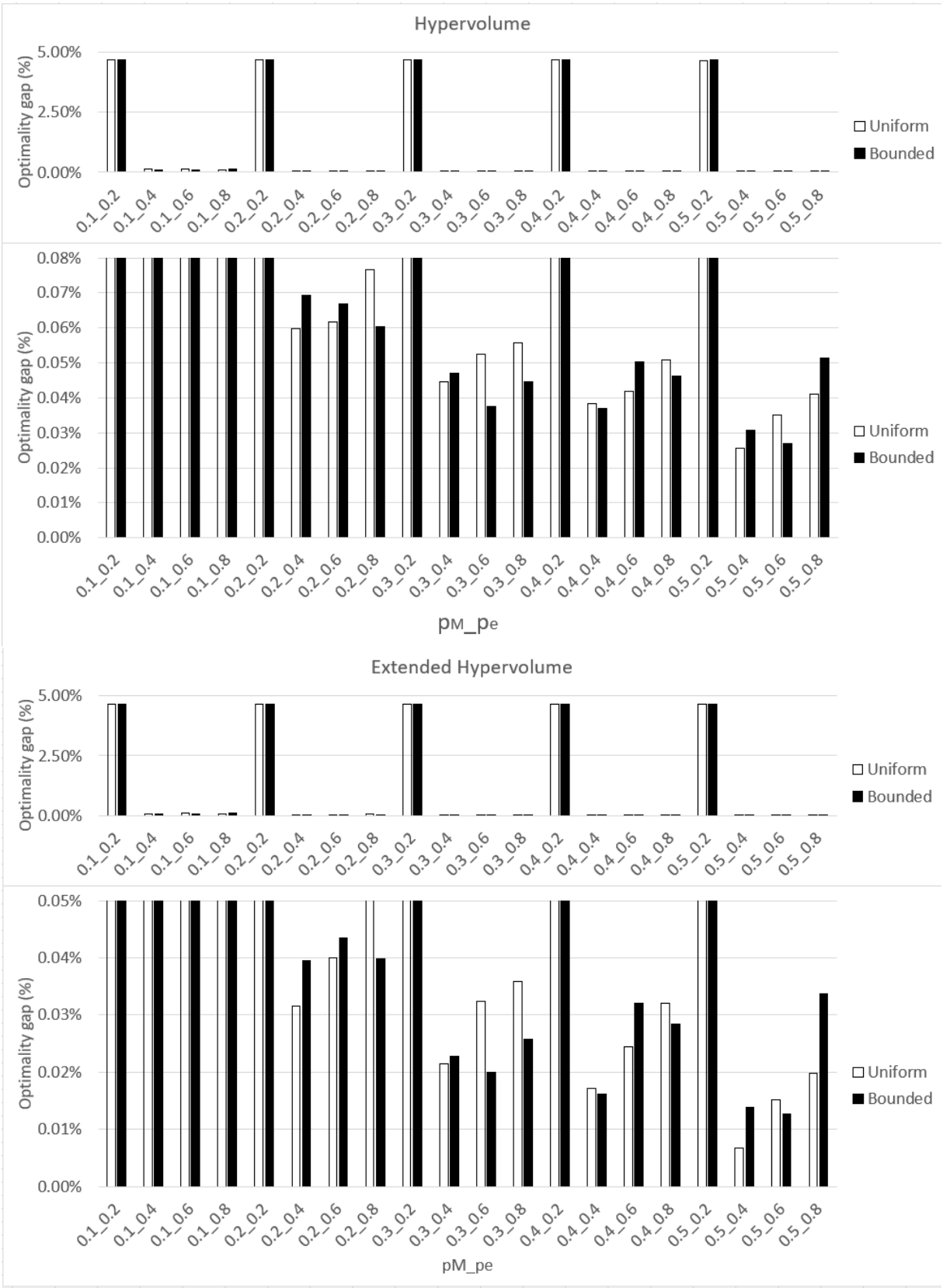

Figure C.2: The average of the performance measures per mutation setting. The algorithm was run fifty times for each setting. The average of the PMs after convergence are shown. The two upper and two bottom plot display the HV and EHV resp. The vertical axis displays the optimality gap from the (E)HV of the known POF/CEF as a percentage of the optimum. The first and second plot are the same but have a different axis scaling to clarify the differences between well-performing settings. The legend indicates the mutation operator, the horizontal axis labels show the values p_M and p_e, separated by a _.