Optimising Colorectal Cancer Screening with Decision Programming

Lotta Mäkinen

School of Science

Bachelor's thesis
Espoo 26.01.2024

Supervisor

Prof. Fabricio Oliveira

Advisor

MSc Olli Herrala

Aalto University
School of Science

Aalto University
School of Science

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi
Abstract of the bachelor's thesis

Author Lotta Mäkinen

Title Optimising Colorectal Cancer Screening with Decision Programming

Degree programme Bachelor's Programme in Science and Technology

Major Mathematics and Systems Sciences **Code of major** SCI3029

Teacher in charge Prof. Fabricio Oliveira

Advisor MSc Olli Herrala

Date 26.01.2024 **Number of pages** 26 **Language** English

Abstract

Colorectal cancer (CRC) is the second most prevalent cancer in Finland. Its growing incidence prompts a critical need for early detection of the cancer and its precursors. Finnish CRC screening programme, initiated as a pilot in 2019 and expanded nationwide in 2022, utilises the faecal immunochemical test (FIT) to detect the hemoglobin levels in stool samples. The patients with positive tests are further invited to colonoscopy. The FIT hemoglobin thresholds (µg/g) indicating test positivity are selected by the guidance of cost-effectiveness analyses.

This study employs the Decision Programming framework to analyze the cost-effectiveness of the Finnish CRC screening programme and to optimise age- and sex-specific FIT thresholds. An influence diagram describing the problem is constructed, featuring one decision node representing the optimal FIT threshold. The primary objective is to maximise the net monetary benefit derived from the screening.

An optimal strategy was successfully found in all different age groups for both sexes. Following a sensitivity analysis, the optimal FIT thresholds for men were determined as 25 µg/g for 55-64 year-olds and 10 µg/g for 65-74 year-olds, and for women, the thresholds were 25 µg/g for 55-59 year-olds and 10 µg/g for 60-74 year-olds. The results align with previous cost-effectiveness analyses, highlighting the reliability and applicability of Decision Programming in the healthcare context and its ability to solve more complex problems.

Keywords Decision programming, decision analysis, colorectal cancer, cancer screening, mixed integer linear programming, stochastic programming

Aalto University
School of Science

Aalto-yliopisto, PL 11000, 00076 AALTO
www.aalto.fi
Tekniikan kandidaatintyön tiivistelmä

Tekijä Lotta Mäkinen

Työn nimi Optimising Colorectal Cancer Screening with Decision Programming

Koulutusohjelma Teknillistieteellinen kandidaattiohjelma

Pääaine Matematiikka ja systeemitieteet **Pääaineen koodi** SCI3029

Vastuuopettaja Prof. Fabricio Oliveira

Työn ohjaaja DI Olli Herrala

Päivämäärä 26.01.2024 **Sivumäärä** 26 **Kieli** Englanti

Tiivistelmä

Suolistosyöpä on toiseksi yleisin syöpä Suomessa. Sen kasvaneen esiintyvyyden vuoksi syövän ja sen esiasteiden varhainen toteaminen on elintärkeää. Suomen suolistosyövän seulontaohjelma, joka alkoi pilottihankkeena vuonna 2019 ja laajentui koko maahan vuonna 2022, käyttää ulosteen immunokemikaalista testiä (eng. feacal immunochemical test, FIT) ulostenäytteiden hemoglobiinitasojen mittaamiseen. Positiivisen testituloksen saaneet potilaat kutsutaan jatkotutkimuksena järjestettävään suoliston tähystykseen. Seulonnan positiivista testitulosta merkitsevät FIT-raja-arvot (µg/g) on valittu kustannus-vaikuttavuusanalyysien tulosten perusteella.

Tässä työssä analysoitiin Suomen suolistosyövän seulontaohjelman kustannus-vaikuttavuutta ja optimoitiin ikä- ja sukupuolisidonnaiset FIT-raja-arvot käyttämällä päätösanalyysin Decision Programming-viitekehystä. Tutkimuksessa muodostettiin seulontaohjelman kulkua kuvaava vaikutuskaavio, jossa yksi päätössolmu edusti optimaalista FIT-raja-arvoa. Optimointiongelman tavoitteeksi asetettiin seulonnan kokonaistaloudellisen hyödyn maksimoiminen.

Optimaalinen strategia löydettiin onnistuneesti kaikille ikäryhmille ja kummallekin sukupuolelle. Herkkyysanalyysin jälkeen optimaalisiksi FIT-raja-arvoiksi päätettiin 25 µg/g 55-64-vuotiaille ja 10 µg/g 65-74-vuotiaille miehille, sekä 25 µg/g 55-59-vuotiaille ja 10 µg/g 60-74-vuotiaille naisille. Tulokset ovat yhteneviä aiempien kustannus-vaikuttavuusanalyysien kanssa. Tämä korostaa Decision Programming-viitekehyksen luotettavuutta ja käytettävyyttä terveydenhuollon kontekstissa ja monimutkaisempien ongelmien ratkaisussa.

Avainsanat suolistosyöpä, päätösanalyysi, syöpäseulonta, stokastinen ohjelmointi

Contents

1 Introduction

Colorectal cancer (CRC) has exhibited a steady increase in its incidence over recent decades. Comprising malignancies of both the colon and rectum, CRC was the second most prevalent cancer in Finland in 2021, with 3,825 newly diagnosed cases and 1,378 disease-related deaths (Finnish Cancer Registry, 2021). This increase in CRC incidence has been attributed to a range of environmental factors, including obesity, red/processed meat consumption, tobacco use, and alcohol consumption (Murphy et al., 2019). The majority of CRC cases are diagnosed in individuals aged 50 and above, with a higher prevalence among men than women (Thélin et al., 2015).

Most CRCs are adenocarcinomas that originate from the epithelial cells lining the gastrointestinal tract. The progression of malignant tumors follows the adenoma-carcinoma sequence, in which a minority of benign adenomas accumulate genetic mutations over an average span of 10-15 years, eventually transforming into colorectal cancer (Eide, 1986). With screening the asymptomatic cancers and their precursors can be found and removed, before they develop into malignant tumours. Thus, the European Union has been recommending CRC screening since 2003 to its member states (Off J Eur Union, 2003).

Colonoscopy is the state-of-the-art method for CRC diagnosis, which has a high sensitivity and specificity in the detection of malignant lesions (Issa et al., 2017). This procedure affords complete visualization of the colorectum, including the distal part of the small intestine, with the option for histological evaluation through biopsies of abnormal findings. However, the colonoscopy is an invasive procedure carrying inherent risks for serious gastrointestinal complications. Furthermore, it is an expensive and a resource intensive method, making it unsuitable for primary screening purposes.

The CRC screening is widely implemented using the faecal immunochemical test (FIT), which detects the presence of hemoglobin in stool samples. The FIT allows for flexible cut-off levels, accommodating sex- and age-specific screening approaches. Individuals with positive FIT results are subsequently selected for diagnostic colonoscopy. In Finland, the FIT-based CRC screening pilot commenced in 2019 and transitioned into a nationwide programme in 2022. The current screening cohort consists of men and women aged 60-68, with plans to expand eligibility to individuals aged 56-74 by 2032. The decision to set the FIT cut-off level at 25 µg/g for both sexes was guided by a simulation study conducted by the Finnish Cancer Registry (FCR), which assessed the cost-effectiveness of various screening strategies (Heinävaara et al., 2022).

The primary objective of this thesis is to explore the feasibility of utilizing the Decision Programming framework to analyze the cost-effectiveness of CRC screening and to identify optimal sex- and age-specific FIT cut-off levels. The CRC screening will be modeled using Decision Programming framework, which was recently developed by Salo et al. (2022). The sensitivity of the model to different health benefit parameters will be assessed.

This thesis is structured as follows. CRC screening methods and an overview of the results of the Finnish CRC screening pilot study and prior cost-benefit analyses

are reviewed in Section 2. The methodology concerning Decision Programming and the development of the optimisation model are described in Section 3. The results of the optimisation model are presented in Section 4. Further discussion and future insights are concluded in Section 5.

2 Background

2.1 Colorectal Cancer Screening Methods

Colorectal cancer (CRC) screening relies on the detection of hemoglobin in fecal samples. Hemoglobin comprises two components: the heme and globin parts, which degrade into various products in the bowel through enzymatic processes. Consequently, fecal samples contain a mixture of hemoglobin and its degradation products, with variations in quantity and degree of degradation dependent on the tumor's location and bleeding pattern.

Two primary methods for detecting fecal blood are the guaiac fecal occult blood test (gFOBT) and the fecal immunochemical test (FIT) (Issa et al., 2017). The gFOBT employs a qualitative approach based on heme presence, where heme reacts with hydrogen peroxide, resulting in a color change to blue. This method requires a moderate heme quantity for a positive result, necessitating significant hemoglobin degradation. As a consequence, gFOBT exhibits lower sensitivity and specificity and may yield false positives due to dietary peroxidase reactions, impacting specificity (Hewitson et al., 2008). Additionally, gFOBT employs a fixed positivity threshold, which cannot be adjusted.

In contrast, FIT is a quantitative test that utilizes specific antibodies binding to the globin complex of human hemoglobin. Various immunoassay methods with high sensitivity to low hemoglobin concentrations measure these antibody-globin complexes. FIT offers flexibility in adjusting the positivity threshold to achieve desired sensitivity and specificity levels, making it the preferred method in CRC screening, gradually replacing gFOBT in Europe (Young et al., 2015).

However, research has indicated that FIT sensitivity and positive predictive value are lower in women compared to men, leading to more false-positive results in women (Arana-Arri et al., 2017). Men also appear to benefit more from gFOBT CRC screening than women (Shaukat et al., 2013). These disparities may stem from differences in tumor location and histology between genders. Women tend to have a higher proportion of tumors on the right side of the colon and more sessile serrated lesions, which bleed less and are harder to detect during colonoscopies (Koskenvuo et al., 2019) (Lash et al., 2010). Consequently, setting lower FIT thresholds for women than men could lead to better tumor detection and similar relative reductions in CRC mortality for both sexes.

Despite these gender-based variations, CRC screening programmes in Europe often employ uniform hemoglobin cutoff levels for both men and women. Nevertheless, different countries adjust their cutoff levels based on CRC incidence and desired FIT positivity rates. Denmark, for example, introduced its CRC screening programme in 2014 with a FIT cutoff level of 20 µg/g for both sexes, resulting in positivity rates of

9.4% in men and 6.0% in women (Njor et al., 2018). In the Netherlands, the screening programme initiated in 2014 with a FIT cutoff of 15 µg/g, but the threshold was adjusted to 47 µg/g due to a high initial positivity rate (10.6%) (Toes-Zoutendijk et al., 2017). Sweden implemented sex-specific cutoff levels in 2015, with 40 µg/g for women and 80 µg/g for men, resulting in positivity rates of 2.6% and 2.5%, respectively (Blom et al., 2019).

2.2 Finnish Colorectal Cancer Screening Pilot

The Finnish randomised gFOBT screening programme started in 2004. The target group included men and women aged 60 to 69, with 362,165 participants. Random allocation to screening or control groups was based on region, sex, and birth year. However, in 2015, the programme was suspended due to no observed difference in CRC mortality between the groups, and a non-significant increase in CRC mortality among women (Pitkäniemi et al., 2015).

In December 2016, the Ministry of Social Affairs and Health decided to relaunch CRC screening in Finland using a quantitative FIT method. The biennial Finnish FIT-based screening pilot commenced in April 2019 in nine municipalities, introducing sex-specific screening strategies with FIT cutoff levels of 25 µg/g for women and 70 µg/g for men (Sarkeala et al., 2021). These thresholds were determined after reviewing available information from existing European CRC screening programmes, with a focus on Nordic countries due to similar socioeconomic structures and CRC incidence rates. Target positivity rates were set at 3% for women and 5% for men, based on previous screening programme data and available colonoscopy resources. The target age group was expanded to include men and women aged 60-74, an extension from the previous gFOBT screening programme.

During the pilot's first year, participation was excellent, with 27,728 participants and a participation rate of 79.3%. FIT positivity rates were 2.8% for men and 2.4% for women. In total, 37 CRC cases and 116 advanced adenomas (AAs) were detected during the first year. In 2020, three additional municipalities joined the pilot, and FIT cutoff levels were lowered to 50 µg/g for men and 15 µg/g for women during the second year to increase positivity rates. Positivity rates rose to 3.6% for men and 3.7% for women in the second year (Kuoppa et al., 2022).

The CRC screening protocol, established during the pilot study and currently employed in the nationwide screening programme, follows these steps:

1. Invitations are sent to potential participants, including a FIT test kit, screening information, testing instructions, a questionnaire, and a return envelope. Participants conduct the test at home and return the sample. Two reminders are sent at 4 and 8 weeks. If laboratory analysis fails, two additional test kits are sent.

2. Samples are analyzed in the screening laboratory.

3. Participants receive test results by mail. For positive results, the letter includes contact information for a municipality screening nurse.

4. Following a pre-colonoscopy interview with the screening nurse, eligible participants are invited for screening colonoscopy.

5. Screening colonoscopies are performed, with lesion removal (polypectomy) and pathological evaluation.

6. Pathologists analyze histological samples. If CRC is detected, patients are informed, and further treatments, usually surgery, are arranged.

2.3 Cost-effectiveness Analysis of CRC Screening

To assess the financial effectiveness of new medical interventions, such as treatments or diagnostic tests, cost-benefit analyses are employed. In these analyses, a monetary value is assigned to gained health benefits. The most frequently used parameter to represent the health outcomes of such interventions is Quality Adjusted Life Years (QALYs). A QALY is calculated by multiplying the number of life-years influenced by the intervention with the quality of one affected life-year. As medical interventions usually impact either the quantity or quality of a patient's life, both of these changes can be quantified in terms of QALYs. However, in the context of cancer screening, a simpler metric is often favored — life-years gained (LYGs) from the screening. The health outcomes are converted to monetary values in cost-benefit analysis using a willingness-to-pay threshold. This threshold represents the price that society is willing to pay per one additional QALY or LYG. The cost-benefit analyses compare different decision strategies with the aim of finding an optimal strategy. These strategies are typically presented as decision trees, and Markov models are commonly used to simulate them.

The Finnish Cancer Registry (FCR) conducted a comprehensive assessment of the cost-effectiveness of the FIT screening pilot programme with the goal of determining the most optimal screening strategy (Heinävaara et al., 2022). This study utilized data from the pilot programme's first year and evaluated 180 different FIT strategies with various cut-off levels, screening intervals, and target age groups.

The analysis was executed through a stochastic MISCAN-Colon microsimulation model. This model generated outcomes for each individual and reported the distribution of different health outcomes. The model simulated 10 million healthy individuals aged 50 until either their death or reaching age 100. Different screening strategies were compared to a no-screening strategy. The FIT positivity cut-off levels considered were 10, 25, 40, 55 and 70 µg/g, with sex-specific FIT parameters for sensitivity and specificity based on the first-year results of the Finnish pilot study. The health benefits derived from the study were quantified in terms of life-years gained (LYGs) from screening, as well as the number of prevented CRCs and CRC-related deaths. The overall referral rate to colonoscopy following FIT test was capped at a maximum of 5%, which was decided following the guidance of a clinical expert group at FCR.

Among the feasible strategies assessed, the optimal strategy was identified as annual screening for men aged 50-79 years with a cut-off of 25 µg/g and for women aged 55-69 years with a cut-off of 10 µg/g. If the same target age group and screening interval was maintained, the optimal strategy was found to be annual screening for

individuals aged 55-74 years, with a cut-off of 25 µg/g for men and 10 µg/g for women. Among the biennial strategies evaluated, the optimal strategy also featured the same cut-off level of 25 µg/g, for men and women aged 55-74 years.

While the results supported sex-specific strategies, it was ultimately concluded that justifying and implementing different target groups and cutoff levels for each gender presented challenges and could potentially hinder screening adherence. Therefore, an efficient strategy suitable for both sexes was deemed preferable, even though it would result in greater benefits for men than for women.

Based on the findings of the cost-effectiveness analysis, the Finnish Government issued a decree in August 2021 to establish a biennial CRC screening programme (Fin Gov, 2021). The target demographic was expanded to include men and women aged 56-74, and the FIT cutoff level was standardized at 25 µg/g for both sexes. Consequently, the Finnish CRC screening programme was rolled out nationwide in 2022, inviting all 60-68-year-old men and women to participate. This gradual implementation will continue until the target age group of 56-74 is reached in 2032.

3 Methodology

3.1 Decision Programming

Decision Programming is an optimisation framework that combines stochastic programming and decision analysis to solve multi-stage decision problems under uncertainty. By representing a problem as an influence diagram, it can be formulated into a mixed integer linear programming (MILP) problem within this framework. The formulated MILP problem can be efficiently solved using off-the-shelf commercial solvers. In Decision Programming, the objective is to find an optimal decision strategy Z, which maximises the expected utility function.

An influence diagram is a directed, acyclic graph $G = (N, A)$ with the arcs A and the nodes $N = C \cup D \cup V$, where the chance nodes C represent the realisations of uncertain events associated with random variables; the decision nodes D represent the decisions among discrete alternatives; and the value nodes V represent the consequences resulting from the realisation of random events and decisions made. Each chance and decision node $j \in C \cup D$ is associated with a finite set of possible states S_j, representing either possible random events or possible decisions, with each individual state denoted as s_j.

The arcs $A = \{(i, j) \mid i, j \in N\}$ represent informational dependencies between nodes i and j. For a given node $j \in N$, the *information set*

$$I(j) \subseteq \{i \in N \mid (i, j) \in A\} \tag{1}$$

is defined as a set of nodes from which there is an arc to node j. The *information state* $s_I(j) \in S_{I(j)}$ of a node j includes all the states of the information set $I(j)$, with $S_{I(j)} = \prod_{i \in I(j)} S_i$ being the set of all possible information states for node j.

A *path* is a sequence of states $s_i \in S_i$, where a specified state is defined for all

chance and decision nodes $i \in C \cup D$. The set of all possible paths is

$$S = \{(s_i)_{i=1,...,n} \mid s_i \in S_i, i = 1, ..., n\} \tag{2}$$

with $n = |C| + |D|$.

Next, we define a random variable $X_j \in S_j$ as the realised state of a chance node $j \in C$. A *local decision strategy* $Z_j : S_{I(j)} \mapsto S_j$ is a mapping between the realisation s_j of a decision node $j \in D$ and its information state $s_{I(j)}$. This mapping can be represented by an indicator function $\mathbb{I} : S_{I(j)} \times S_j \mapsto \{0, 1\}$. We have

$$\mathbb{I}(s_{I(j)}, s_j) = \begin{cases} 1, & \text{if } Z_j(s_{I(j)}) = s_j, \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

A *decision strategy* Z contains a local decision strategy for each decision node $Z = \{Z_j \mid j \in D\}$. A decision strategy Z is said to be compatible with a path $s \in S$, if $Z_j(s_{I(j)}) = s_j$ for all $j \in D$. The set of all decision strategies is denoted by $\mathbb{Z}$. Only the set of compatible paths $S(Z) \subseteq S$ are considered active paths. Binary *decision variables* $z(s_j \mid s_{I(j)})$ are defined, equal to one if $\mathbb{I}(s_{I(j)}, s_j) = 1$ and to zero otherwise.

For each chance node, there is a discrete probability distribution corresponding to one of the information states. Therefore, we can define a conditional probability of observing a given state s_j for node $j \in C$ as $\mathbb{P}(X_j = s_j \mid X_{I(j)} = s_{I(j)})$. The *path probability* $\mathbb{P}(s \mid Z)$ is a conditional probability of a path s being observed given a strategy Z, and it is defined as

$$\mathbb{P}(s \mid Z) = \left(\prod_{j \in C} \mathbb{P}(X_j = s_j \mid X_{I(j)} = s_{I(j)}) \right) \left(\prod_{j \in D} \mathbb{I}(s_{I(j)}, s_j) \right). \tag{4}$$

Let $x(s) \in [0, 1], s \in S$ be the *path compatibility variables* representing the right-hand term $\prod_{j \in D} \mathbb{I}(s_{I(j),s_j})$ in Equation (4). Using the notation of decision variables $z(s_j \mid s_{I(j)})$, we have $x(s) = \prod_{j \in D} z(s_j \mid s_{I(j)})$. We can see that $x(s)$ take the value one if the strategy Z is compatible with the path $s \in S$, and zero otherwise. Therefore, $x(s)$ serve as indicator variables for whether a path s is compatible with the decision strategy Z defined by the decision variables z. Moreover, the upper bound of the path probability $\mathbb{P}(s \mid Z)$ can be defined as

$$p(s) = \left(\prod_{j \in C} \mathbb{P}(X_j = s_j \mid X_{I(j)} = s_{I(j)}) \right), \tag{5}$$

and it follows that Equation (4) can be reformulated as $\mathbb{P}(s \mid Z) = p(s)x(s)$.

Each value node $v \in V$ has a utility function $U_v : S_{I(v)} \mapsto \mathbb{R}$ mapping its information state $s_{I(v)}$ to a utility value $U_v(s_{I(v)})$. Furthermore, the utility of a path s is the aggregated utilities of individual value nodes $U(s) = \sum_{v \in V} U_v(S_{I(v)})$. The default objective in choosing a best strategy $Z \in \mathbb{Z}$ is to maximise the expected utility, which can be defined as

$$\max_{Z \in \mathbb{Z}} \sum_{s \in S} U(s)p(s)x(s). \tag{6}$$

.

With these building blocks, the influence diagram can be converted to an optimisation problem. However, as the computational performance of Decision Programming formulation is highly dependent on the number of variables in the model, it is advisable to limit the number of paths included. First, let us define the notion of *locally compatible paths* $S_{s_j|s_{I(j)}}$, which is the collection of paths s compatible with local decision strategies Z_j when $z(s_{I(j)}, s_j) = 1$. Only paths that are compatible with the selected strategy have a probability different than zero, thus

$$\sum_{s \in S_{s_j|s_{I(j)}}} p(s)x(s) \leq z(s_j \mid s_{I(j)}), \forall j \in D, s_j \in S_j, s_{I(j)} \in S_{I(j)}. \tag{7}$$

As $x(s) \in [0, 1]$ and it must hold for $z(s_j \mid s_{I(j)}) = 1$ if $x(s) = 1, s \in S_{s_j|s_{I(j)}}$, we can derive the following inequality by considering only the locally compatible paths

$$\sum_{s \in S_{s_j|s_{I(j)}}} x(s) \leq |S_{s_j|s_{I(j)}}| z(s_j \mid s_{I(j)}), \forall j \in D, s_j \in S_j, s_{I(j)} \in S_{I(j)}, \tag{8}$$

where $|S_{s_j|s_{I(j)}}|$ is the number of locally compatible paths.

To further limit the number of paths in the model, we can consider only the *active* locally compatible paths $S_{s_j|s_{I(j)}} \cap S(z)$. At every other decision node $d \in D \setminus j$, only one alternative $s_d \in S_d$ will be selected. Thus, the number of active locally compatible paths can be estimated as

$$|S_{s_j|s_{I(j)}} \cap S(z)| = \frac{|S_{s_j|s_{s_I(j)}}|}{\prod_{d \in D \setminus \{j, I(j)\}} |S_d|}, \tag{9}$$

and, therefore, we can reformulate Equation (8) into the form

$$\sum_{s \in S_{s_j|s_{I(j)}}} x(s) \leq \frac{|S_{s_j|s_{s_I(j)}}|}{\prod_{d \in D \setminus \{j, I(j)\}} |S_d|} z(s_j \mid s_{I(j)}), \forall j \in D, s_j \in S_j, s_{I(j)} \in S_{I(j)}. \tag{10}$$

Depending on the problem at hand, some subpaths might be unrealisable and never be observed. These paths are called *ineffective* or *forbidden paths*. When these subsets of forbidden subpaths are removed from the set of all paths, we get *effective paths* $S^* \subset S$. Moreover, $S^*_{s_j|s_{I(j)}}$ is the set of effective locally compatible paths. If the model has forbidden paths, then $|S^*_{s_j|s_{I(j)}}| < |S_{s_j|s_{I(j)}}|$. Depending on the problem structure, either the number of effective paths or the number of active locally compatible paths is smaller, and it is safer to consider the minimum of these two bounds. Therefore, we can further reformulate Equation 10 into

$$\sum_{s \in S_{s_j|s_{I(j)}}} x(s) \leq \min(|S^*_{s_j|s_{I(j)}}|, \frac{|S_{s_j|s_{s_I(j)}}|}{\prod_{d \in D \setminus \{j, I(j)\}} |S_d|}) z(s_j \mid s_{I(j)}). \tag{11}$$

In Decision Programming (Salo et al., 2022) the influence diagram is converted into a mixed-integer linear programming (MILP) problem using the abovementioned

concepts. The objective is to maximise the expected utility similarly to Equation (6). This thesis uses the improved model formulation of the Decision Programming, which enhances the numerical performance of the model by considering only the necessary paths. The new formulation and the more detailed proof have been described in Hankimaa et al. (2023).

The optimal decision strategy can be obtained from the following optimisation model

$$\max_{Z \in \mathbb{Z}} \sum_{s \in S^*} U(s)p(s)x(s) \tag{12}$$

$$\text{s.t.} \sum_{s_j \in S_j} z(s_j \mid s_{I(j)}) = 1, \qquad \forall j \in D, s_{I(j)} \in S_{I(j)} \tag{13}$$

$$\sum_{s \in S_{s_j|s_{I(j)}}} x(s) \leq \Gamma(s_j \mid s_{I(j)}) z(s_j \mid s_{I(j)}), \qquad \forall j \in D, s_j \in S_j, s_{I(j)} \in S_{I(j)} \tag{14}$$

$$\sum_{s \in S^*} p(s)x(s) = 1, \tag{15}$$

$$0 \leq x(s) \leq 1, \qquad \forall s \in S^x \tag{16}$$

$$z(s_j \mid s_{I(j)}) \in \{0, 1\}, \qquad \forall j \in D, s_j \in S_j, s_{I(j)} \in S_{I(j)} \tag{17}$$

where $\Gamma(s_j \mid s_{I(j)}) = \min(|S^*_{s_j|s_{I(j)}}|, \frac{|S_{s_j|s_{s_I(j)}}|}{\prod_{d \in D \setminus \{j, I(j)\}} |S_d|})$.

In this optimisation model, constraint (13) ensures that each local decision strategy maps each information state to exactly one decision. Constraint (14) ensures that only the path probability variables $x(s)$ that are associated with locally compatible paths $s \in S_{s_j|s_{I(j)}}$ are equal to one. Constraint (15) limits the sum of all possible path probabilities to one and constraint (16) restricts the value of $x(s)$ between zero and one, the value being one if the chosen path is compatible with a given strategy and otherwise zero, as previously discussed. Constraint (17) declares the variables $z(s_j \mid s_{I(j)})$ to be binary variables.

3.2 Constructing the Optimisation Model

The constructed model determines an optimal decision strategy for selecting the cut-off level for feacal immunochemical testing in CRC screening. This model incorporates the prior risk of the patient having a disease based on the 2019 CRC incidence rates in Finland. The FIT threshold is implemented as a discrete variable with five different levels. The problem setting mirrors the course of the CRC screening programme in Finland. The patient chooses to participate in CRC screening and submits a screening sample to a testing laboratory. The FIT screening test is conducted and provides either a positive or negative result based on the chosen threshold level and the patient's prior risk of CRC. If the FIT result is positive, the patient is referred for a colonoscopy. Based on the patient's prior risk, the colonoscopy may reveal a normal bowel, an adenoma, an advanced adenoma or a colorectal cancer. If an abnormal lesion is found, a polypectomy (removal of the lesion for further analysis) is performed. The patient may suffer from adverse effects following the colonoscopy,

and the risk is dependent on whether a polypectomy is performed. This model was individually constructed for both sexes and different age groups, allowing for different thresholds for various target groups.

3.2.1 Influence Diagram

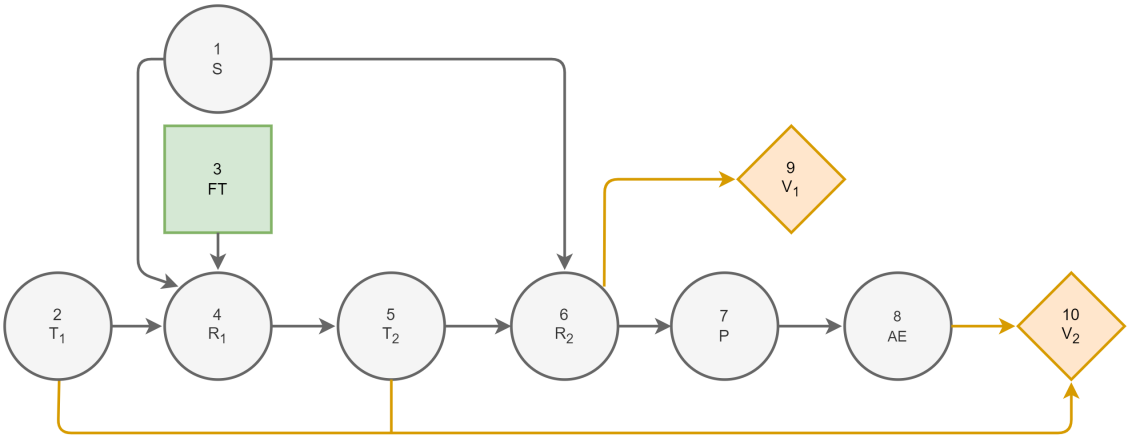

Figure 1: Influence diagram representation of the CRC screening optimisation problem.

The influence diagram representing the problem is shown in Figure 1. The green square node FT is a decision node, reflecting the chosen FIT threshold. The grey circular nodes $S, T_1, R_1, T_2, R_2, P, AE$ are chance nodes. Node S signifies the health state of the patient's bowel, i.e. the prior risk of CRC and other findings. Node T_1 denotes the chance that the FIT test is taken and returned by the patient. Node R_1 represents the FIT test result, which has both the threshold FT and the bowel state S in its information set affecting the test result.

Node T_2 corresponds to the chance of a patient undergoing further examinations, i.e., a colonoscopy. Node R_2 represents the outcome of the colonoscopy, which is dependent on the patient's prior risk of having an abnormal lesion in their bowel. Consequently, node R_2 includes node S in its information set in addition to node T_2. Node P represents the probability of a polypectomy (biopsy of a tumour) being performed. Node AE represents the probability of the patient suffering from serious gastrointestinal adverse effects, such as bleeding or perforation.

There are two value nodes in the diagram, represented by orange diamonds. Node V_2 represents the screening costs, mapping the performed tests and procedures, i.e. FIT (node T_1) and colonoscopy (node T_2), and the treatment of possible adverse effects (node AE) to their costs. Node V_1 represents the health benefits obtained from the screening. The information set of node V_1 includes node R_2 and it maps the colonoscopy findings to their achieved health benefits.

3.2.2 Information States and Probability Distributions

The states of each node are summarised in Table 1. The states of node S represent whether a patient has a normal bowel (N), adenoma (A), advanced adenoma (AA), or

Table 1: The states for each node j in the influence diagram representing the problem.

Node	States
S	{N, A, AA, CRC}
T_1	{yes, no}
FT	{10, 25, 40, 55, 70}
R_1	{+, -, NA}
T_2	{yes, no}
R_2	{N, A, AA, CRC, NA}
P	{yes, no}
AE	{yes, no}

malignant colorectal cancer (CRC). The classification criteria were devised similarly to the classification used in the CRC screening pilot study in Finland (Sarkeala et al., 2021). An advanced adenoma (AA) is defined as an adenoma with 25% or greater villous component, high-grade dysplasia, or a size of 10mm or larger. These adenomas are considered premalignant and have the potential to develop into a malignant state. Non-villous and smaller adenomas were considered benign adenomas (A), and their risk of developing into CRC was considered nearly insignificant. The probability of a patient having a malignant disease was based on the 2021 CRC incidence in Finland (Finnish Cancer Registry, 2021). The incidences of adenomas and advanced adenomas were not readily available, so they were estimated using the results of the pilot study of CRC screening (Sarkeala et al., 2021). The ratios of adenomas and advanced adenomas to malignant tumours reported in the study were applied to population level to determine the needed incidences.

The states of the decision node FT correspond to different thresholds for FIT. We chose to use the same thresholds 10, 25, 40, 55, 70 µg/g as in the study by Heinävaara et al. (2022), as sensitivities for FIT with these thresholds were readily available, allowing for easy comparisons with previous studies.

The states of T_1 represent whether a patient returns a viable screening sample to the laboratory. The probability was approximated to be 80% based on the results of the pilot study (Sarkeala et al., 2021). The same value was used for both sexes. The states of R_1 represent the results of the FIT. If the test is not returned, the result is NA; otherwise, the test is either positive or negative. If a FIT is conducted, its result corresponds to the sensitivity of the test with the given threshold. The sensitivity of a diagnostic test is defined as the probability of a positive test given that the patient has the disease. The sensitivities for different FIT thresholds were obtained from the study by Heinävaara et al. (2022) and they are summarised in Table 2.

The states of T_2 correspond to the probability that a colonoscopy is performed. For simplification purposes, it is assumed that if the patient has a positive test result from the FIT test, a colonoscopy is always performed. In reality, this decision depends on the patient's medical history and overall health. The states of R_2 correspond to the result of screening colonoscopy: normal bowel (N), adenoma (A), advanced adenoma (AA), or colorectal cancer (CRC). If a colonoscopy is not performed, the state is NA. The specificity of the colonoscopy was assumed to be 100%, and the

Table 2: Sensitivities of feacal immunochemical test for both sexes with different threshold levels for different bowel states.

Bowel state	Sex	10 µg/g	25 µg/g	40 µg/g	55 µg/g	70µg/g
Normal	Male	0.04	0.01	0.01	0.01	0.01
	Female	0.03	0.02	0.01	0.01	0.01
Adenoma	Male	0.27	0.26	0.24	0.18	0.14
	Female	0.09	0.09	0.08	0.06	0.03
Advanced adenoma	Male	0.6	0.55	0.5	0.35	0.21
	Female	0.4	0.35	0.25	0.2	0.15
Colorectal cancer	Male	0.9	0.8	0.75	0.7	0.65
	Female	0.57	0.5	0.35	0.25	0.19

sensitivities were obtained from the study by Heinävaara et al. (2022).

The node P had states representing the chance of a polypectomy being performed. It is assumed that if any tumour (A, AA or CRC) is found, a polypectomy always follows. Node AE represents the possibility of the patient suffering from adverse effects following the colonoscopy. The risk for adverse effects is dependent on whether a polypectomy is performed, the risk being 9.4/1000 for the patients undergoing a polypectomy and 2.8/1000 for others (Warren et al., 2009).

3.2.3 Utility Function

As the default objective in the Decision Programming framework is the expected utility, we selected the objective function to be the expected *net monetary benefit (NMB)*, which is defined as

$$NMB = health\ outcomes \cdot WTP - costs, \tag{18}$$

where WTP describes the societal willingness-to-pay threshold. NMB converts the expected health benefits (QALYs or LYGs) to monetary values, making them easier to compare with the monetary costs.

To apply the Decision Programming model effectively in this context, separate health outcome values, either as QALYs or LYGs, would be needed for each possible value of node V_1 corresponding to the states of node T_2. This would require evaluating the exact number of additional QALYs or LYGs the patient would gain if an adenoma, an advanced adenoma, or a colorectal cancer is found. However, these values were not readily available in the literature, and constructing a separate simulation model would be necessary to evaluate them. Furthermore, the WTP threshold for CRC screening is not defined in Finland. Heinävaara et al. (2022) used a WTP threshold of 10,000€/LYG in their study. It was decided that evaluating health outcome values was beyond the scope of this thesis.

However, it was hypothesised that the model was not very sensitive to the selection of these parameters. Therefore, we decided to use an approximated monetary parameter called health benefits (HB), representing the product of health outcomes and WTP. Thus, NMB is defined as

$$NMB = HB - costs, \tag{19}$$

in our optimisation model.

We approximated health benefit parameter values based on the primary treatment costs of CRC, estimated to be 22,200€ in Finland (Färkkilä et al., 2015). Thus 22,200€ was set to be the parameter value of finding a malignant tumour during colonoscopy (HB_C). As the malignant potential of an advanced adenoma is considered to be significant, the parameter value of finding an advanced adenoma (HB_{AA}) was estimated to be half of the CRC value (11,100€). Given the low risk of a benign adenoma developing into a malignant tumour compared to its incidence, the parameter value for finding an adenoma (HB_A) was set relatively low, just double the cost of colonoscopy (800€). The parameter value for a normal colonoscopy finding (HB_N) was set to zero. The health benefit values were used as the values for node V_1, and are presented in Table 3 as monetary values (€). The sensitivity of the decision model to the health benefit parameters is later assessed in this thesis.

The values of node V_2 representing the costs of screening and treatment of possible adverse effects were inserted into the model as monetary values (€). The costs were obtained from the study by Heinävaara et al. (2022) and they are presented in Table 4 as the parameter values used in the model.

Table 3: Approximated monetary health benefit parameter values for different colonoscopy findings.

Parameter	Value (€)
Normal (HB_N)	0
Adenoma (HB_A)	800
Advanced adenoma (HB_{AA})	11,100
Colorectal cancer (HB_C)	22,200

Table 4: The parameter values used in the model for indicating the testing costs of FIT and colonoscopy, and treating costs of colonoscopy adverse effects.

Parameter	Value (€)
FIT test (C_F)	12.4
Colonoscopy (C_C)	400
Adverse effects (C_{AE})	3,280

The constructed model was implemented in the Julia language using the DecisionProgramming.jl package (Oliveira et al., 2021). The Gurobi Optimization solver was used for the optimisation (Gurobi Optimization, 2023).

4 Results

4.1 Optimal Decision Strategy

The Decision Programming model was executed separately for different age groups and both sexes to determine the optimal strategy for selecting a FIT threshold. An optimal strategy was found in each case. The analysis covered age groups of 55-59

years, 60-64 years, 65-69 years, and 70-74 years for both sexes, aligning with the planned target age group for CRC screening in Finland. The summarised results are presented in Table 5. The obtained optimal thresholds are consistent with the findings from the study by Heinävaara et al. (2022). According to our model, the optimal FIT threshold for men is 25 µg/µl in the age group of 55-69 year-olds and 10 µg/µl in the age group of 70-74 year-olds. For women, the optimal threshold is 25 µg/µl for ages 55-64 years and 10 µg/µl for ages 65-74.

The prevalence of different events with this screening strategy, according to the model, is summarised in Table 6. Given our model's assumption that all patients with a positive FIT result undergo colonoscopy, the number of colonoscopies is equivalent to the positivity rate of FIT. Thus, the positivity rate of FIT is 3.0% for men and 3.3% for women, resulting in 16,497 and 19,135 colonoscopies for men and women, respectively. A total of 4,371 AAs would be found (2,978 in men and 1,393 in women), along with 994 cases of CRC (670 in men and 324 in women). The prevalence of serious gastrointestinal adverse effects is estimated at 0.01%, meaning 142 patients would suffer from adverse effects (75 men and 67 women). It is important to note that screening in Finland is conducted biennially, meaning these events, i.e., colonoscopies, found AAs and CRCs, and adverse effects, would occur over a two-year time frame.

Table 5: The optimal hemoglobin threshold levels (µg/g) in feacal immunochemical testing for different target age groups and both sexes with initial parameters.

Sex	55 - 59 y	60 - 64 y	65 - 69 y	70 - 74 y
Men	25	25	25	10
Women	25	25	10	10

Table 6: The outcomes of optimal screening strategies for different target groups with initial parameters. The prevalence of different events is shown as absolute numbers and as percentages from the number of FIT tests performed: COLs = number of performed, colonoscopies, AAs = found advanced adenomas, CRCs = found colorectal cancers, AEs = adverse effects from colonoscopy.

Sex	Age	FITs	COLs (%)	AAs (%)	CRCs (%)	AEs (%)
Men	55 - 59	145,602	2,588 (1.8)	341 (0.2)	81 (0.06)	10 (0.007)
	60 - 64	139,589	2,900 (2.1)	562 (0.4)	122 (0.09)	13 (0.009)
	65 - 69	134,472	3,314 (2.5)	827 (0.6)	180 (0.1)	17 (0.01)
	70 - 74	129,081	7,695 (6.0)	1,249 (1.0)	287 (0.2)	35 (0.03)
Total	55 - 74	548,744	16,497 (3.0)	2,978 (0.5)	670 (0.1)	75 (0.01)
Women	55 - 59	145,809	3,751 (2.6)	167 (0.1)	26 (0.02)	11 (0.008)
	60 - 64	144,372	3,886 (2.7)	250 (0.2)	57 (0.04)	13 (0.009)
	65 - 69	145,667	5,550 (3.8)	380 (0.3)	90 (0.06)	19 (0.01)
	70 - 74	146,097	5,948 (4.1)	595 (0.4)	140 (0.1)	23 (0.02)
Total	55 - 74	581,945	19,135 (3.3)	1,393 (0.2)	324 (0.06)	67 (0.01)

Table 7: The results of the sensitivity analysis for men. The optimal FIT thresholds after a single parameter change by either -50%, 50%, or 100% are shown. The original FIT thresholds (µg/g) and the updated thresholds based on the sensitivity analysis are displayed in the bottom of the table.

Parameter	Change (%)	55 - 59 y	60 - 64 y	65 - 69 y	70 - 74 y
HB_N	-50	25	25	25	10
	+50	25	25	25	10
	+100	25	25	25	10
HB_A	-50	25	25	25	10
	+50	25	25	25	10
	+100	25	25	25	10
HB_{AA}	-50	25	25	25	25
	+50	25	25	10	10
	+100	25	25	10	10
HB_C	-50	25	25	25	25
	+50	25	25	25	10
	+100	25	25	10	10
Original results		25	25	25	10
Updated results		**25**	**25**	**10**	**10**

4.2 Sensitivity Analysis

Given that the parameters representing health benefits were approximated in this study, a sensitivity analysis was conducted to assess the model's sensitivity to these parameters. The analysis involved running the model multiple times for each case with different parameter values. Each time, only one parameter was changed while keeping others constant. The chosen parameter was either decreased by 50%, or increased by 50% or by 100%. If the model's results remained unchanged after these adjustments, it was concluded that the model was not sensitive to that specific parameter in that case. The results of the sensitivity analyses for men and women are presented in Tables 7 and 8, respectively.

It is important to emphasize that when selecting parameter values for health benefits, priority was given to ensuring that these values were conservatively estimated on the higher side. This was done to prevent setting FIT thresholds too high, which could result in reduced cancer detection rates, leading to higher CRC mortality and an ineffective screening programme. Overestimating these values, on the other hand, would primarily lead to more colonoscopies and increased costs. Therefore, if the sensitivity analysis indicated that increasing the parameter values led to lower optimal FIT thresholds, it would suggest that the initial parameter estimates might have been too conservative. In such cases, we opted to adopt the lower FIT threshold. If the results differed when decreasing parameter values, we decided to retain the original FIT thresholds.

Looking at the results for men (Table 7), the FIT threshold remained unchanged after the parameter changes in age groups 55-59 and 60-64 years. This suggests that the model is not sensitive to the selection of parameters in these age groups, and

Table 8: The results of the sensitivity analysis for women. The optimal FIT thresholds after single parameter change by either -50%, 50% or 100% are shown. The original FIT thresholds (µg/g) and the updated thresholds based on the sensitivity analysis are displayed in the bottom of the table.

Parameter	Change (%)	55 - 59 y	60 - 64 y	65 - 69 y	70 - 74 y
HB_N	-50	25	25	10	10
	+50	25	25	10	10
	+100	25	25	10	10
HB_A	-50	25	25	25	10
	+50	25	25	10	10
	+100	25	25	10	10
HB_{AA}	-50	25	25	10	10
	+50	25	10	10	10
	+100	25	10	10	10
HB_C	-50	25	25	10	10
	+50	25	10	10	10
	+100	10	10	10	10
Original results		25	25	10	10
Updated results		**25**	**10**	**10**	**10**

the original results are accurate. However, in the age group 65-69 years, increasing parameters HB_{AA} by 50% and HB_C by 100% resulted in the optimal threshold of 10 µg/g. Consequently, the lower threshold of 10 µg/g was selected for this age group. In the 70-74 age group, reducing parameters HB_{AA} and HB_C by 50% resulted in higher optimal thresholds of 25 µg/g; otherwise, the results remained unchanged, and the threshold was not adjusted in this age group.

For women, the model's results remained unchanged after parameter changes in the youngest age group (55-59 years) and oldest age group (70-74 years). In these target groups, the model was considered insensitive to parameter variations. In the age group 60-64 years, increasing the parameters HB_{AA} and HB_C by 50% resulted in a lower threshold value of 10 µg/g. Therefore, the lower threshold was chosen for this age group. In the 65-69 age group, decreasing HB_{AA} by 50% led to a higher threshold, but otherwise, the original threshold was preferred, as the results remained consistent.

The updated thresholds resulted in an increased number of colonoscopies, as well as more detected CRCs and AAs. However, it also led to a higher number of adverse effects. The prevalences of various events are shown in Table 9. The FIT positivity rates increased to 3.7% in men and 3.5% in women. A total of 4,481 AAs and 1,025 CRCs were found, representing an additional 110 AAs and 31 CRCs compared to the previous thresholds. Adverse effects following colonoscopy were estimated to affect 147 patients.

Table 9: The outcomes of optimal screening strategies for different target groups with the new FIT thresholds after the sensitivity analysis. The new values after sensitivity analysis are in bold. The prevalence of different events is shown as absolute numbers and as percentages from the number of FIT tests performed: COLs = number of performed, colonoscopies, AAs = found advanced adenomas, CRCs = found colorectal cancers, AEs = adverse effects from colonoscopy.

Sex	Age	FITs	COLs (%)	AAs (%)	CRCs (%)	AEs (%)
Men	55 - 59	145,602	2,588 (1.8)	341 (0.2)	81 (0.06)	10 (0.007)
	60 - 64	139,589	2,900 (2.1)	562 (0.4)	122 (0.09)	13 (0.009)
	65 - 69	134,472	**7,340 (5.5)**	**902 (0.7)**	**203 (0.2)**	**30 (0.02)**
	70 - 74	129,081	7,695 (6.0)	1,249 (1.0)	287 (0.2)	35 (0.03)
Total	55 - 74	548,744	**20,523 (3.7)**	**3,053 (0.6)**	**693 (0.1)**	**75 (0.02)**
Women	55 - 59	145,809	3,751 (2.6)	167 (0.1)	26 (0.02)	11 (0.008)
	60 - 64	144,372	**4,965 (3.7)**	**286 (0.2)**	**65 (0.05)**	**17 (0.01)**
	65 - 69	145,667	5,550 (3.8)	380 (0.3)	90 (0.06)	19 (0.01)
	70 - 74	146,097	5,948 (4.1)	595 (0.4)	140 (0.1)	23 (0.02)
Total	55 - 74	581,945	**20,214 (3.5)**	**1,428 (0.2)**	**332 (0.06)**	**72 (0.01)**

5 Discussion and Conclusions

In this thesis, we successfully applied the Decision Programming framework to optimise the threshold of the faecal immunochemical test (FIT) used in colorectal cancer (CRC) screening. Despite the need for approximations in certain critical parameters representing the health benefits of the screening, our results closely aligned with those of previous studies. Our analysis illustrates the potential effectiveness of the Decision Programming framework within the context of healthcare. With more precise parameter selection, the framework can potentially be expanded to address even more complex healthcare problems.

Our influence diagram, a key element of the Decision Programming model, was relatively straightforward and consisted of only one decision node. Consequently, Decision Programming is not deemed necessary for solving the problem in question, as alternative methods such as dynamic programming and simulation could potentially be employed. However, the results highlight that the problem was extremely easy to construct and solve with Decision Programming. Thus, the simplicity of our model paves the way for the development of more intricate models with additional parameters and decision nodes, which would be computationally less complex and demanding to solve with Decision Programming than with alternative methods.

To the best of our knowledge, this is the first study to optimise FIT thresholds for different age groups, in addition to distinguishing between sexes. Our findings suggest that implementing sex-specific and age-specific thresholds can be beneficial. From simply a cost-benefit perspective, sex- and age-specific thresholds may be preferred in CRC screening. However, it is crucial to consider the potential impact on screening adherence. Lowering the threshold can increase an individual's perceived benefit from screening, as it increases the likelihood of detecting bowel abnormalities. Conversely, higher thresholds may lead to reduced participation rates among individuals who do

not perceive sufficient benefit from the screening.

The target positivity rates in the Finnish CRC screening pilot study were 3% for women and 5% for men. Our model indicates positivity rates of 3.5% for women and 3.7% for men, resulting in approximately 40,737 individuals undergoing colonoscopy over two years. In the study by Heinävaara et al. (2022), a colonoscopy constraint of 5% was employed to ensure sufficient colonoscopy capacity. Given that our results fall within this constraint and closely align with the target positivity rates, the existing colonoscopy capacity is likely adequate to accommodate the proposed screening strategy.

Furthermore, as CRC screening gradually reduces the prevalence of CRC by detecting cancer precursors, the prevalence of CRC will likely decrease when extending the screening to younger age groups. Consequently, FIT positivity rates may decrease, leading to fewer individuals being referred for colonoscopy. Thus, the number of colonoscopies is unlikely to be a limiting factor with the optimized thresholds in the future. It is worth considering the need for re-optimizing FIT thresholds as CRC prevalence decreases over time. This re-optimising of FIT thresholds could be formatted as an influence diagram containing multiple periods and changing CRC prevalence. This could be formulated into MILP problem and could potentially be solved with Decision Programming, providing an intriguing optimisation problem over multiple periods.

While our study provides valuable insights, it has several limitations. The evaluation of health benefit parameters was approximated, which could impact the accuracy of our results. However, our sensitivity analysis revealed that the model was not highly sensitive to parameter variations, minimizing the potential impact on our conclusions. Future research could address these limitations by improving the accuracy of health benefit parameter estimates. One approach could involve constructing a multi-objective model, eliminating the need for precise parameter values. Instead of maximizing net monetary benefit (NMB), this multi-objective model could minimize monetary costs and maximize the number of tumors detected. Given that the cost and health benefits of tumor detection are already represented as separate value nodes in our influence diagram, it would not need to be modified to cater to the multi-objective approach.

Furthermore, the model does not incorporate explicit constraints on the total number of colonoscopies performed. While our results indicate that colonoscopy capacity is unlikely to be a limiting factor with the optimal strategy, the inclusion of such constraints directly within the model can enhance its robustness. By implementing these constraints, the model can effectively identify and remove forbidden strategies that would be deemed unfeasible. If the model is further expanded, this elimination of forbidden strategies is vital to ensure the computational performance of the model.

Finally, it is worth noting that the threshold in our model was represented as a discrete variable with relatively large intervals (15 µg/g). This choice aligns with the approach taken in the study by Heinävaara et al. (2022), where FIT sensitivities were available for these specific values. While this decision was reasonable given the limitations of certain parameters, such as health benefits, attempting to decrease

the interval size revealed the model's heightened sensitivity to changes in other parameters. Consequently, in the future, particularly if more precise data for the other variables can be acquired or a more intricate model is crafted, employing finer intervals or even a continuous spectrum for the threshold value should be contemplated.

In conclusion, optimising FIT thresholds in CRC screening was successfully possible with Decision Programming framework, proving that the framework can be used in the healthcare context. Results were obtained for different age groups and sexes and it was proved that different thresholds were beneficial. Future studies should focus on expanding the model by improving parameter estimates or formulating the problem as a multi-objective optimisation model, and exploring continuous threshold options.

References

Eunate Arana-Arri, Isabel Idigoras, Begoña Uranga, Raquel Pérez, Ana Irurzun, Iñaki Gutiérrez-Ibarluzea, Callum G Fraser, and Isabel Portillo. Population-based colorectal cancer screening programmes using a faecal immunochemical test: should faecal haemoglobin cut-offs differ by age and sex? *BMC cancer*, 17(1):1–13, 2017.

Johannes Blom, Christian Löwbeer, K Miriam Elfström, Marika Sventelius, Daniel Öhman, Deborah Saraste, and Sven Törnberg. Gender-specific cut-offs in colorectal cancer screening with fit: Increased compliance and equal positivity rate. *Journal of medical screening*, 26(2):92–97, 2019.

Tor J Eide. Risk of colorectal cancer in adenoma-bearing individuals within a defined population. *International journal of cancer*, 38(2):173–176, 1986.

Niilo Färkkilä, Saku Torvinen, Harri Sintonen, Tiina Saarto, Heikki Järvinen, Juha Hänninen, Kimmo Taari, and Risto P Roine. Costs of colorectal cancer in different states of the disease. *Acta Oncologica*, 54(4):454–462, 2015.

Fin Gov. Valtioneuvoston asetus seulonnoista annetun valtioneuvoston asetuksen 2 §:n muuttamisesta, 2021. (339/2011) 2 §, Availabte at: https://finlex.fi/fi/laki/alkup/2021/20210752.

Finnish Cancer Registry, 2021. Accessed 11.6.2023. Available: https://cancerregistry.fi/statistics/cancer-statistics/.

Gurobi Optimization. Gurobi optimiser reference manual, 2023. URL http://www.gurobi.com.

Helmi Hankimaa, Olli Herrala, Fabricio Oliveira Oliveira, and Jaan Tollander de Balsch. Decisionprogramming.jl –a framework for modelling decision problems using mathematical programming. 2023. Pre-print available: https://arxiv.org/abs/2307.13299.

Sirpa Heinävaara, Andrea Gini, Tytti Sarkeala, Ahti Anttila, Harry de Koning, and Iris Lansdorp-Vogelaar. Optimizing screening with faecal immunochemical test for both sexes-cost-effectiveness analysis from finland. *Preventive medicine*, 157: 106990, 2022.

Paul Hewitson, Paul Glasziou, Eila Watson, Bernie Towler, and Les Irwig. Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult): an update. *Official journal of the American College of Gastroenterology| ACG*, 103(6):1541–1549, 2008.

Iyad A Issa, Malak Noureddine, and et al. Colorectal cancer screening: An updated review of the available options. *World journal of gastroenterology*, 23(28):5086, 2017.

Laura Koskenvuo, N Malila, J Pitkäniemi, J Miettinen, S Heikkinen, and V Sallinen. Sex differences in faecal occult blood test screening for colorectal cancer. *Journal of British Surgery*, 106(4):436–447, 2019.

Aino Kuoppa, Tytti Sarkeala, Aku Leivonen, and Sirpa Heinävaara. Kolorektaalisyövän seulontapilotti – miten kaksi ensimmäistä vuotta sujuivat? *Duodecim*, 138 (19):1677–84, 2022.

Richard H Lash, Robert M Genta, and Christopher M Schuler. Sessile serrated adenomas: prevalence of dysplasia and carcinoma in 2139 patients. *Journal of clinical pathology*, 63(8):681–686, 2010.

Neil Murphy, Victor Moreno, David J Hughes, Ludmila Vodicka, Pavel Vodicka, Elom K Aglago, Marc J Gunter, and Mazda Jenab. Lifestyle and dietary environmental factors in colorectal cancer susceptibility. *Molecular aspects of medicine*, 69:2–9, 2019.

Sisse Helle Njor, Lennart Friis-Hansen, Berit Andersen, Bo Søndergaard, Dorte Linnemann, Jens Christian Riis Jørgensen, Ole Roikjær, and Morten Rasmussen. Three years of colorectal cancer screening in denmark. *Cancer epidemiology*, 57: 39–44, 2018.

Off J Eur Union. Council recommendation of 2. December 2003 on cancer screening (2003/878/EC), 2003. L327: 34-8.

Fabricio Oliveira, Olli Herrala, Jaan Tollander de Balsch, Helmi Hankimaa, and Juho Andelmin. DecisionProgramming.jl, October 2021. URL https://github.com/gamma-opt/DecisionProgramming.jl.

J Pitkäniemi, K Seppä, M Hakama, O Malminiemi, T Palva, MS Vuoristo, H Järvinen, H Paimela, P Pikkarainen, A Anttila, et al. Effectiveness of screening for colorectal cancer with a faecal occult-blood test, in finland. *BMJ open gastroenterology*, 2 (1):e000034, 2015.

Ahti Salo, Juho Andelmin, and Fabricio Oliveira. Decision programming for mixed-integer multi-stage optimization under uncertainty. *European Journal of Operational Research*, 299(2):550–565, 2022.

Tytti Sarkeala, Martti Färkkilä, Ahti Anttila, Marja Hyöty, Matti Kairaluoma, Tero Rautio, Markku Voutilainen, Sanni Helander, Maija Jäntti, Milla Lehtinen, et al. Piloting gender-oriented colorectal cancer screening with a faecal immunochemical test: population-based registry study from finland. *BMJ open*, 11(2):e046667, 2021.

Aasma Shaukat, Steven J Mongin, Mindy S Geisser, Frank A Lederle, John H Bond, Jack S Mandel, and Timothy R Church. Long-term mortality after screening for colorectal cancer. *New England Journal of Medicine*, 369(12):1106–1114, 2013.

Camille Thélin, Sanjay Sikka, and et al. Epidemiology of colorectal cancer—incidence, lifetime risk factors statistics and temporal trends. *Screening for colorectal Cancer with colonoscopy. London: IntechOpen Limited*, pages 61–77, 2015.

Esther Toes-Zoutendijk, Monique E van Leerdam, Evelien Dekker, Frank Van Hees, Corine Penning, Iris Nagtegaal, Miriam P van der Meulen, Anneke J van Vuuren, Ernst J Kuipers, Johannes MG Bonfrer, et al. Real-time monitoring of results during first year of dutch colorectal cancer screening program and optimization by altering fecal immunochemical test cut-off levels. *Gastroenterology*, 152(4): 767–775, 2017.

Joan L Warren, Carrie N Klabunde, Angela B Mariotto, Angela Meekins, Marie Topor, Martin L Brown, and David F Ransohoff. Adverse events after outpatient colonoscopy in the medicare population. *Annals of internal medicine*, 150(12): 849–857, 2009.

Graeme P Young, Erin L Symonds, James E Allison, Stephen R Cole, Callum G Fraser, Stephen P Halloran, Ernst J Kuipers, and Helen E Seaman. Advances in fecal occult blood tests: the fit revolution. *Digestive diseases and sciences*, 60: 609–622, 2015.