Journal Pre-proof

Optimizing the Finnish colorectal cancer population screening program with decision programming

Lauri Neuvonen, Mary Dillon, Eeva Vilkkumaa, Ahti Salo,
Maija Jäntti, Sirpa Heinävaara

PII: S0377-2217(25)00287-5
DOI: https://doi.org/10.1016/j.ejor.2025.04.022
Reference: EOR 19498

To appear in: *European Journal of Operational Research*

Received date : 7 July 2023
Accepted date : 11 April 2025

Please cite this article as: L. Neuvonen, M. Dillon, E. Vilkkumaa et al., Optimizing the Finnish colorectal cancer population screening program with decision programming. *European Journal of Operational Research* (2025), doi: https://doi.org/10.1016/j.ejor.2025.04.022.

Optimizing the Finnish Colorectal Cancer Population Screening Program with Decision Programming

Lauri Neuvonen[a,*], Mary Dillon[a], Eeva Vilkkumaa[a], Ahti Salo[b], Maija Jäntti[c], Sirpa Heinävaara[c]

[a] *Department of Information and Service Management, Aalto University School of Business, Finland*
[b] *Department of Mathematics and Systems Analysis, Aalto University School of Science, Finland*
[c] *The Finnish Cancer Registry*

Abstract

In Finland, colorectal cancer (CRC) incidence rates have steadily increased over the last decades and as of 2020, CRC is the second most common cancer in both males and females. CRC is a crucial concern for the public health of Finland, highlighted by the recent implementation of a national population screening program. In this paper, we optimize the screening test positivity cut-off levels and the use of potential incentives for stratified populations to minimize cancer prevalence. The optimization results, computed with the novel Decision Programming approach for discrete multi-stage decision problems under uncertainty, show the optimal cut-off levels and uses of incentives for Finnish target groups subject to different constraints on colonoscopy capacity. The outcomes of these optimal strategies are estimated to determine the expected corresponding prevalences of CRC and required colonoscopies, and expected third-party costs. Finally, measures describing different equality perspectives are presented.

Keywords: OR in health services, colorectal cancer screening, optimization, influence diagrams, decision programming

1. Introduction

Colorectal cancer (CRC), also referred to as bowel cancer, is cancer of the colon and/or rectum. CRC incidence rates have steadily increased over the last decades with CRC currently being the second most common cancer in adults in Finland, and the third most common worldwide (Finnish Cancer Registry 2017, World Health Organization 2018). In 2020 CRC accounted for 2.3% of all deaths in Finland and was the second most diagnosed cancer in both males and females, making it a crucial concern to public health (Finnish Cancer Registry 2022, Advisory Board of OSF 2022).

Over 70% of colorectal cancers develop via the adenoma-carcinoma sequence (Hardy et al. 2000) in which adenomas (i.e., growths on the epithelial tissue in the bowel) develop into cancer. This sequence is a slow process that can take from several years to a decade (Simon 2016). Due to this, early detection and removal of pre-cancerous adenomas can prevent progression to cancer, making

*Corresponding author.
Email address: lauri@laurineuvonen.fi (Lauri Neuvonen)

CRC extremely suitable for population screening. Population screening of CRC using faecal occult blood testing has long been shown to reduce later-stage CRC incidence and improve mortality (Mandel et al. 1993, Kewenter et al. 1994, Kronborg et al. 1996, Mandel et al. 1999). Along with early detection of adenomas, early diagnosis of cancer improves the probability of successful treatments and outcomes by providing care at the earliest possible time. This has an important impact on public health strategy as it helps avoid CRC-attributed deaths and morbidity, in addition to high treatment and indirect costs associated with advanced cancer stages (World Health Organization 2017).

In 2019, Finland implemented a new CRC population screening program, which became nationwide in 2022. The program invites entire age cohorts to screen using faecal immunochemical testing (FIT) to first filter participants for further examination. Then, those program participants whose FIT result exceeds a given cut-off level (hemoglobin level in the stool sample) are invited for a colonoscopy (Finnish colorectal cancer screening expert groups 2021). Colonoscopy is an invasive procedure, which allows the visual inspection of the colon to detect abnormalities such as polyps or adenomas. If an abnormality is discovered, it can be removed via polypectomy and sent for pathology assessment. Colonoscopies are resource intensive and may cause discomfort and adverse effects (such as bleeding or perforation of the bowel) to the participant. It is therefore important to develop cost-effective screening strategies in which scarce colonoscopy resources are allocated to those participants for whom they yield the highest health benefits.

In the current program, the FIT cut-off level for a positive result is fixed based on the participant's sex. Specifically, the Finnish program employs a cut-off level of 25 μg Hb/g for both females and males, regardless of their age. In reality, however, the interpretation of the result should be differentiated by participants' risk profiles (i.e., age, sex, and family history of cancer; Selby et al. 2018, Peng et al. 2020). Moreover, because a positive result leads to an invitation to a colonoscopy, the cut-off level should also take into account the capacity to carry out colonoscopies. Consequently, the program may not be the most cost-effective approach to reduce Finnish colorectal cancer incidence and mortality.

The purpose of this paper is to build an optimization model for the Finnish CRC population screening program with the aim of minimizing cancer prevalence (i.e., the probability of a population member having CRC) in the target population with respect to a colonoscopy resource constraint. In particular, for each segment of the population specified by the participants' sex and age, we define (i) whether the segment should be invited to screen or not, (ii) what the optimal FIT cut-off levels are, and (iii) whether it pays off to use incentives to boost participation. The results of our model can be used to improve the current program in ways that help allocate scarce colonoscopy resources in a more cost-effective manner.

The rest of the paper is structured as follows. Section 2 discusses related work and our contributions to the literature. Section 3 presents an overview of the current population screening program used in Finland, and details the model used to optimize this program. The results of the model are presented and discussed in Section 4, and Section 5 concludes.

2. Related work and contributions

Traditionally, cost-effectiveness evaluations of public health programs, such as population screening, are assessed through methods of cost-benefit analysis (e.g. Ellison et al. 2002), cost-utility analysis (e.g. Gupta et al. 2011, Dillon et al. 2018), or cost-effectiveness analysis (e.g. Ladabaum et al. 2019). Health economic methodologies such as these typically compare costs and health outcomes of a given screening strategy to those of a baseline strategy (for example, no screening). Quality of life indicators, such as quality-adjusted life-years (QALYs) and life-years saved/gained (LYS/LYG), are frequently used to calculate the incremental cost-effectiveness ratio (ICER) for comparing alternative strategies. In these analyses, strategies are not optimized in the mathematical sense, but rather assessed in regard to their dominance. Consequently, such analyses tend to suggest strategies that are infeasible due to a lack of resources or suboptimal in that (i) the resources could be reallocated to achieve a better population-level health outcome or (ii) the same health outcome with less resources.

Some studies have sought to overcome the above issues by investigating the cost-effectiveness of a large set of alternative strategies so that the best-performing strategy within this set could be assumed to be close to optimal (Wilschut et al. 2011, Van Der Meulen et al. 2017). For instance, a recent study by Whyte et al. (2022) uses an existing CRC simulator to evaluate the cost-effectiveness of over 60,000 screening strategies, which together cover a wide range of possible FIT cut-off values, ages to start screening, and screening frequencies. Heinävaara et al. (2022) use a similar approach to find cost-effective screening strategies when the FIT cut-off level can differ depending on the participant's sex. In particular, they develop separate MISCAN-Colon models for Finnish men and women to evaluate the cost-effectiveness of 181 sex-specific strategies and 362 combinations thereof. These kinds of approaches cannot, however, be utilized when the goal is to find screening strategies in which FIT cut-off values as well as screening times are to be optimized for segments defined by both age and sex. This is because in such settings, the number of possible strategies becomes so large (e.g., close to 26 million in the setting described in this paper) that the probability of finding an optimal strategy within a set of even tens of thousands of predetermined strategies becomes vanishingly small.

The two most typical approaches for optimization-based design of screening programs are Partially Observable Markov Decision Processes (POMDPs) and simulation-optimization. POMDPs, in particular, have been popular in optimizing screening programs thanks to their ability to account for development and uncertainty in patients' states of health. Ayer et al. (2012) build a POMDP model to optimize patient-specific mammography screening times, and Alagoz et al. (2013) provide a tutorial in optimizing cancer screening using a POMDP approach.Cevik et al. (2018) present a constrained POMDP model to study the optimal allocation of limited mammography resources to screen a population. Lee et al. (2019) optimize the use of limited resources for the screening of a population for hepatocellular carcinoma by modeling the problem as a family of restless bandits in which each patient's disease progression is assumed to evolve as a POMDP.

From the perspective of CRC screening, a particularly relevant approach is presented by Erenay

et al. (2014), who develop a POMDP model to optimize colonoscopy screening policies for CRC prevention and surveillance in the U.S. context in view of maximizing total expected quality-adjusted life years (TQALYs). In this model both static (sex and age) and dynamic risk factors (history of CRC or adenomatous polyp) are considered in modeling disease progression. Input parameters for the model are obtained from a simulation study based on Mayo Clinic-Rochester patient records (Erenay et al. 2011), the SEER database[1], and the literature. The model is validated based on expert opinion as well as by comparing model outputs with statistics from reliable databases and established simulation models. The optimal policies resulting from this model suggest that, for instance, screening should occur more frequently than what is recommended by current guidelines, especially for younger people. Moreover, while low-risk women should be screened less frequently than low-risk men, women with a personal history of CRC should undergo colonoscopy more often than men with a personal history of CRC.

From the point of view of optimizing the Finnish CRC population screening program, the above approach (as well as other POMDP approaches) is not entirely suitable. First, this approach only considers a single screening tool (colonoscopy), and therefore cannot be used to optimize the use of a pre-screening test (such as a FIT). Yet, in CRC screening, the FIT is used precisely to avoid the unnecessary use of colonoscopies. Moreover, while the above approach accommodates uncertainties regarding cancer state transitions, it ignores uncertainties related to adherence to invitation, continued participation, and potential adverse events resulting from screening. These aspects may, however, have significant impacts on the overall performance of the screening program. The accommodation of a FIT and additional chance events in a POMDP model would likely lead to computational intractability, at least when aiming to obtain exact optimal solutions (e.g. Li et al. 2009). Moreover, the approach by Erenay et al. (2014) does not admit constraints on the capacity for carrying out colonoscopies. Such constraints are, however, relevant from the perspective of policy making and directly affect optimal FIT cut-off levels, whereby they should ideally be accommodated. Thus, for the purposes of our study, a different modeling approach is required.

Methods of simulation-optimization have been used to find optimal screening strategies even in fairly complex settings and for different forms of cancer including, e.g., cervical (McLay et al. 2010), breast (Rauner et al. 2010), and prostate (Bertsimas et al. 2018) cancer. Underwood et al. (2012), for instance, use simulation-optimization in the context of prostate cancer screening to find optimal cut-off values for prostate-specific antigen used to determine whether a prostate biopsy is necessary. Young et al. (2021) use derivative-free optimization coupled with microsimulation to find an optimal CRC screening strategy for males, when a strategy is defined by the choice of a single screening tool (FIT, sigmoidoscopy, or colonoscopy), a starting and ending age for screening, and screening frequency. Yet, the optimization models used in these types of approaches are often highly nonlinear, which typically necessitates the use of heuristics (such as a genetic algorithms used by Underwood et al. 2012) to keep the problem computationally tractable, especially if the

[1] `https://seer.cancer.gov/`

strategies themselves consist of a number of different decisions (see, e.g., Neuvonen et al. 2023). Although such heuristics have been found to perform well in specific problem instances, there are generally no guarantees for the optimality of the found solutions, or even the extent to which the objective function values of these solutions deviate from the optimum. Moreover, the performance of a heuristic is typically highly contingent on the characteristics of the optimization problem (see, e.g., Young et al. 2021), whereby the suitability of a given heuristic in the context of optimizing a given screening program may not be generalizable to other kinds of public health programs.

In this paper, we build an optimization model for the Finnish CRC population screening program with the aim of minimizing cancer prevalence in the target population with respect to a colonoscopy resource constraint. The optimization model is implemented in three levels. At the first level, we find all segment-specific (segments defined by age and sex) Pareto optimal screening strategies in view of maximizing the probability of detecting CRC, a large adenoma or a benign growth while minimizing the probability of a colonoscopy. At the second level these results are used to find sex-specific Pareto optimal screening strategies in view of minimizing the prevalence of CRC and the probability of performing a colonoscopy. Finally, at the third level, we identify the combination of segment-specific strategies that together minimize the prevalence of cancer in the target population, given a fixed constraint on the total expected number of performed colonoscopies. To maintain computational efficiency, dominated or infeasible strategies are eliminated along the way.

In building the model, we employ Decision Programming (Salo et al. 2022), a novel approach to solving discrete multi-period decision problems under uncertainty. In this approach, Influence Diagrams are built for each segment to capture decisions related to the screening strategy as well as various uncertainties regarding, e.g., adherence to screening invitations, continued participation, test results, and potential side effects from colonoscopy examination. Moreover, segment-specific transition probabilities between different bowel states (alongside screening decisions) are used to capture cancer progression. The Decision Programming framework can be used to formulate the tasks of identifying Pareto optimal screening strategies for the segment-specific Influence Diagrams as mixed-integer linear programming (MILP) problems, which are solved using a Modified Augmented Weighted Tchebychev algorithm. Importantly, the ability of the Decision Programming framework to accommodate multiple objectives – which would not be possible within a POMDP framework – enables us to efficiently eliminate infeasible or dominated strategies, which leads to improved computational performance.

Compared to existing approaches to optimizing screening programs, our approach offers several benefits. First, it employs Influence Diagrams to capture decisions and uncertainties related to screening strategies. Such diagrams resemble basic process flow charts and are, therefore, arguably more easily understood by healthcare practitioners than, e.g., large decision trees used in POMDP models. Second, our approach helps find optimal screening strategies under resource constraints even in cases where these strategies consist of a number of different decisions the outcomes of which are subject to various sources of uncertainty. Thus, the problem formulation does not need to be simplified for computational purposes, which increases the legitimacy of decision recommendations.

Finally, the Decision Programming approach is flexible in that it can be used to accommodate multiple objectives, deterministic constraints as well as probabilistic risk measures that can be treated as objectives or constraints. Thus, our approach can be augmented to fit the purposes of optimizing other kinds of screening programs as well.

In summary, the contribution of this paper to existing literature is threefold. First, we propose a novel approach to optimizing CRC screening strategies that simultaneously accommodates i) multiple objectives, ii) optimization of segment-specific invitation decisions and FIT cut-off levels, iii) detailed modeling of uncertainties related to the screening process, and iv) resource constraints. Second, we use this approach to offer insights into how the CRC population screening program in Finland could be improved. Finally, we demonstrate the usefulness of the Decision Programming methodology in the healthcare context, in which it has not been applied.

3. Optimization of the Finnish CRC population screening program

3.1. Problem and model overview

In April 2019, Finland established a new CRC population screening program in volunteering municipalities (see Finnish Cancer Registry 2019 for details). The program began with twelve out of 311 municipalities in Finland volunteering to take part. The program became nationwide in 2022, and by 2031 all people aged between 56 and 74 will be invited to screen. The program invites entire age cohorts to screen using a feacal immunochemical test (FIT), after which those participants with a positive FIT result are invited to a colonoscopy. In 2022 all 60-, 62-, 64-, 66- and 68-year-old males and females in volunteering municipalities were invited to participate in the program. Participation is free of charge for the invitee.

A central screening hub mails the screening invitations and FITs to participants. The FIT is performed at home and returned in a pre-paid envelope to the screening hub. Once the laboratory has analyzed the FIT sample, a written result is given to the participant via mail. Those with a positive result are requested to contact a screening nurse in their municipality of residence to discuss the need for further examination. This additional examination is usually a colonoscopy; however, this may vary between municipalities depending on their standard procedure. If a growth is found during the additional examination, a sample is taken and sent to a pathologist for analysis. Once the pathologist's findings have been reviewed, a decision on the need for further examination or treatment is made. If necessary, the participant is referred to surgery. Treatment may be required at this stage, but the ensuing treatment process is not within the scope of the screening program. After treatment, patients participate in a separate surveillance program that usually lasts for five years. A new invitation to the screening program is sent every two years to those participants who have not received treatment in the previous round. Screening is continued in this periodic manner as long as the invitee's age is within the program limits.

In this section, we develop a model to determine optimal screening strategies for the Finnish colorectal cancer population screening program in view of minimizing the prevalence of colorectal

cancers while adhering to a capacity constraint on the total expected number of performed colonoscopies. In this model, a screening strategy consists of decisions for each segment of the target population (defined by the participants' age and sex) on (i) whether to send an invite to participate in the program, (ii) whether to offer an incentive for participation, and (iii) which FIT cut-off level to use. The inclusion of an incentive decision in our problem does not reflect the reality of the program; however, it may be an option in other countries with similar programs.

In our model, we assume that a participant's bowel state $b \in \mathcal{B} = \{\mathrm{N}, \mathrm{B}, \mathrm{L}, \mathrm{R}\}$ is either Normal (N), Benign growth (B), Large adenoma (L) or CRC (R). Benign growths are growths that do not pose a high risk of developing into cancerous growths in the near future, whereas large adenomas do, even though they are not cancerous at the moment nor will they necessarily develop into such. This classification of growths is based on the adenoma-carcinoma sequence, which is a common assumption in the literature on CRC (Silva-Illanes & Espinoza 2018, Diedrich et al. 2023). A similar classification is used by, e.g., Gyrd-Hansen et al. (1997), Heitman et al. (2010), and Pence et al. (2013). Transition probabilities between bowel states are assumed to depend on the participant's sex and age. In this way, our model incorporates more specific information on cancer progression than most studies on CRC population screening, in which differences in transition probabilities are assumed to depend on age only (Silva-Illanes & Espinoza 2018, Lansdorp-Vogelaar et al. 2022).

In the model, we allow the strategy for any segment of the target population to be chosen independently of other segments, but assume that once chosen, the same strategy will be used for all participant groups of same sex and age, i.e., same segment. For instance, the same strategy is applied to all 62-year old males regardless of the year they enter the screening program. We also assume that from year to year a segment (e.g., 60-year-old females) contains the same number of participants. Under these simplifying yet reasonable assumptions, we can focus on finding Pareto optimal sex-specific screening strategies for only a single group of participants entering the program at the age of 60. Thus, segments can be defined by combinations (g, k), where $g \in \{\mathrm{F}, \mathrm{M}\}$ is the participant's sex (female or male) and $k \in \{1, \ldots, 5\}$ is the screening period which corresponds to a participant's age in 2-year intervals so that period $k = 1$ refers to 60-year-olds, period $k = 2$ to 62-year-olds etc. We also assume that the population segments are internally homogeneous in that participants within a given segment are not differentiated from one another based on, e.g., family history. In this sense, our model operates on average representatives for the segments. This is in line with the general idea of population screening, i.e., testing a large segment of the population regardless of their detailed state of health.

Under the above assumptions, we model the problem of finding optimal screening strategies via three levels. A schematic description of this model is illustrated in Figure 1. At level 1 (performed separately for all screening periods of both males and females), influence diagrams (IDs) are used to capture how abnormal bowel states are found, colonoscopies are performed, and costs are generated in a given period as a function of screening decisions. These IDs are used to find Pareto optimal screening strategies for each segment with respect to minimizing the expected number of colonoscopies and maximizing the expected number of detected abnormal bowel states. At level

2, these segment-specific strategies are combined to produce sex-specific strategies that minimize the prevalence of CRC in all age groups combined while minimizing the number of colonoscopies performed. At level 3, those two Pareto optimal sex-specific strategies are identified that together minimize the expected prevalence of cancer in the screened population subject to a constraint on the total expected number of colonoscopies.

Our model is implemented through five steps, as illustrated in Process 1. The first four steps correspond to level 2 (the identification of Pareto optimal sex-specific strategies) and include level 1 (the identification of Pareto optimal segment-specific strategies) as the first step. These four steps are carried out iteratively for all five screening periods so that on each iteration round k, a set of strategies $Z_{g,\to k} \in \{Z_{g,\to k}\}^{\mathrm{PO}}$ is found that are Pareto optimal up to period k. Then, level 3 (the identification of sex-specific strategies that together minimize the expected population-level CRC prevalence subject to a constraint on the total expected number of colonoscopies performed) is carried out in step 5. The tasks associated with each step are detailed in Sections 3.2-3.4. Tables of notation related to the screening process and IDs are presented in the Supplementary material.

3.2. Level 1: Identification of Pareto optimal segment-specific strategies

At level 1 (step 1 in Process 1), segment-specific ID models are used to find Pareto optimal screening strategies for each segment in view of minimizing the expected number of colonoscopies and maximizing the expected number of abnormal bowel states. In what follows, we will give a detailed description of (i) the segment-specific ID model and (ii) the Decision Programming formulation for finding Pareto optimal segment-specific strategies for the ID model.

3.2.1. Segment-specific ID model

An ID is a discrete acyclic graph constructed of three types of nodes $\mathcal{N}$ represented as sets; decision $\mathcal{D}$, chance $\mathcal{C}$ and utility nodes $\mathcal{U}$, with dependencies shown by directed arcs $\mathcal{A} \subseteq \{(i,j)|i,j \in \mathcal{N}, i \neq j\}$. Every chance node $j \in \mathcal{C}$ and decision node $j \in \mathcal{D}$ has a finite set of discrete states $s_j \in \mathcal{S}_j$. The state $s_j \in \mathcal{S}_j$ of a given node $j \in \mathcal{C} \cup \mathcal{D}$ represents a chance or decision alternative. An arc $(i,j) \in \mathcal{A}$, represented by an arrow, indicates that node i is the predecessor of node j, and that the state s_j at node j is conditionally dependent on the state s_i at the preceding node i.

The ID corresponding to the Finnish population screening program is shown in Figure 1. Here, decision nodes in the screening pathway are represented by squares, uncertainties (i.e., chance nodes) in the screening process are represented by circles, and the utilities that are to be optimized are represented by diamonds. The differences between segments in the target population are reflected in the parameter values of the segment-specific IDs (please see the Supplementary material for details on these parameters).

The decision nodes 1, 2, and 3 in the ID of Figure 1 correspond to decisions about 1) what FIT test cut-off value to use to select patients for a colonoscopy in the given target segment, 2) whether to use an incentive to boost participation among invitees in this segment (specifically, we assume that an incentive worth 10 euros halves the number of non-returned samples) and 3)

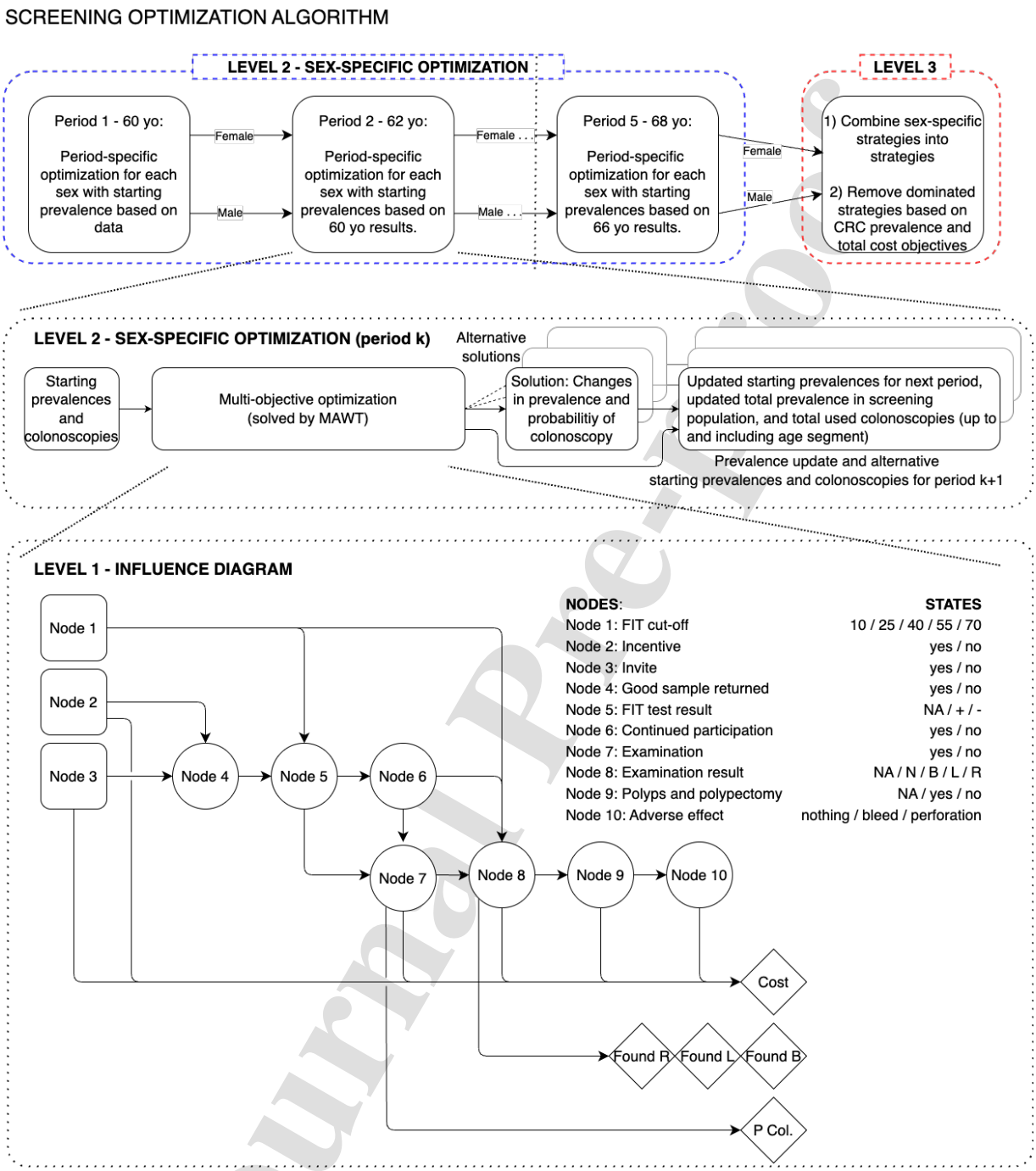


Figure 1: Schematic description of the multistage optimization approach to solve the CRC screening optimization problem.

whether to invite the target segment to the screening program. The corresponding sets of decision alternatives are $\mathcal{S}_2 = \mathcal{S}_3 = \{\text{Yes}, \text{No}\}$ and $\mathcal{S}_1 = \{10, 25, 40, 55, 70\}$ μg Hb/g of blood in the stool sample. We use binary decision variables $z(s_j)$ to indicate whether an alternative is chosen or not. A segment-specific strategy $Z_{g,k} = [z(s_j)]_{j \in \mathcal{D}}$ is defined as a vector of such decision variables.

Process 1: Steps in the process of optimizing the Finnish CRC screening program

Level 2 (including level 1)

1. In the first step (**level 1**), segment-specific IDs are utilized to find strategies $Z_{g,k}$ that are Pareto optimal in view of minimizing the expected share of invitees for whom a colonoscopy is performed while maximizing the expected share of invitees with detected cancers and abnormal bowel states in segment (g, k). As input parameters, these segment-specific IDs use starting prevalences $\psi_{g,k-1,b}$ of different bowel states $b \in \mathcal{B}$, which are either estimated from data (period $k = 1$) or computed based on screening decisions corresponding to each Pareto optimal strategy $Z_{g,\to k-1} \in \{Z_{g,\to k-1}\}^{\mathrm{PO}}$ up to the previous period alongside information on natural cancer progression due to aging (periods $k = 2, \ldots, 5$; see step 3).
2. In step 2, the set of Pareto optimal strategies found in step 1 is combined with the corresponding $Z_{g,\to k-1}$ to obtain set $\{Z_{g,\to k}\}$. The vector of expected shares of invitees with detected cancers and abnormal bowel states corresponding to strategy $Z_{g,\to k}$ in this set is denoted by $\tilde{\psi}_{g,k}(Z_{g,\to k}) = [\tilde{\psi}_{g,k,\mathrm{B}}(Z_{g,\to k}), \tilde{\psi}_{g,k,\mathrm{L}}(Z_{g,\to k}), \tilde{\psi}_{g,k,\mathrm{R}}(Z_{g,\to k})]$.
3. The third step is to compute for each $Z_{g,\to k} \in \{Z_{g,\to k}\}$ (i) the updated starting prevalences $\psi_{g,k}(Z_{g,\to k})$ for the next period by accounting for the impacts of screening as well as natural cancer progression due to aging, and (ii) the combined cancer prevalence $\Psi_{g,k,\mathrm{R}}(Z_{g\to k})$ for segments comprising sex g and age groups corresponding to screening periods $1, \ldots, k$.
4. In the fourth step, those $Z_{g,\to k}$ are removed from the set $\{Z_{g,\to k}\}$ which are infeasible with respect to the constraint on the maximum number of colonoscopies, or dominated in view of minimizing i) the expected prevalence of benign growths in the current segment, ii) the expected prevalence of large growths in the current segment, iii) the total expected prevalence of CRC in the sex-specific population screened up to and including this period, and iv) the total expected number of colonoscopies required by segments corresponding to sex g up to and including this period. The remaining solutions form the set $Z^{\mathrm{PO}}_{g,\to k}$ of Pareto optimal strategies for sex g from period 1 to period k.

These four steps are repeated for both sexes $g \in \{\mathrm{F}, \mathrm{M}\}$ until all periods have been optimized. This results in sets $\{Z_g\}^{\mathrm{PO}} = \{Z_{g,\to K}\}^{\mathrm{PO}}$ of Pareto optimal sex-specific strategies.

Level 3

5. The fifth step finds the pair of sex-specific strategies from sets $\{Z_{\mathrm{F}}\}^{\mathrm{PO}}$ and $\{Z_{\mathrm{M}}\}^{\mathrm{PO}}$ that minimizes the expected population-level cancer prevalence subject to a constraint on the total expected number of performed colonoscopies.

Chance nodes in the ID correspond to returning a FIT sample, FIT results, continued participation, the discovery of polyps and removal of polyps via polypectomy during a colonosopy, and adverse effects from the colonoscopy (such as perforation and bleeding). The conditional probabilities for these chance nodes have been obtained from the literature (see the Supplementary material for details).

Utility nodes in Figure 1 correspond to costs, health outcomes and treatment decisions resulting from the screening process of a single participant. Specifically, the values at these utility nodes

are captured by variables $U_{\mathrm{C}},\ U_{\mathrm{R}},\ U_{\mathrm{L}},\ U_{\mathrm{B}}$, and U_{PCol}, where $U_{\mathrm{C}} \in \mathbb{R}^{+}$ refers to the direct costs incurred by the participant's screening process (see the Supplementary material for details on costs), $U_{\mathrm{R}},\ U_{\mathrm{L}},\ U_{\mathrm{B}} \in \{0,1\}$ to detected cancers and growths so that each variable obtains a value of 1 if and only if cancer (R), large growth (L), or benign growth (B) is detected, and $U_{\mathrm{PCol}} \in \{0,1\}$ to performed colonoscopies so that the variable obtains a value of 1 if and only if a colonoscopy is performed.

The values of the utility node variables depend on the adopted strategy $Z_{g,k}$ (i.e., decision node alternatives) as well as the realizations of different chance events that are compatible with this strategy (i.e., chance node alternatives). A combination of uniquely defined values for all chance and decision node alternatives $s_j \in S_j\ \forall j \in \mathcal{C} \cup \mathcal{D}$ is called a *scenario path* $s \in \mathcal{S}$, where $\mathcal{S}$ is the set of all possible scenario paths. For example, one possible scenario path is

$$s = (40, \text{No, Yes, Yes, Positive, Yes, Colonoscopy, Benign Growth, Polypectomy, None}).$$

The above scenario path describes a sequence of events in which a person is invited (s_3 = Yes) to the screening program without monetary incentive (s_2 = No) using a FIT with a cut-off level value of 40 μg/Hb g ($s_1 = 40$). The invitee returns a usable sample (s_4 = Yes) that tests over 40 μg/Hb g and is thus scored as positive (s_5 = Positive), and contact is established with the local nurse (s_6 = Yes). A further examination is chosen to be a colonoscopy (s_7 = Colonoscopy), the result of which indicates a benign growth in the bowel (s_8 = Benign Growth). During the colonoscopy, the growth is removed from the bowel (s_9 = Polypectomy) and no adverse event occurs (s_{10} = None). This path results in utility node values $U_{\mathrm{C}} = 585.37$ €, $U_{\mathrm{R}} = U_{\mathrm{L}} = 0,\ U_{\mathrm{B}} = 1$, and $U_{\mathrm{PCol}} = 1$.

3.2.2. Decision Programming formulation for optimizing the ID model

The aim of the ID optimization model is to find Pareto optimal strategies $Z_{g,k}$ for each segment-specific ID in view of minimizing the expected number of performed colonoscopies and maximizing the expected shares of the target segment for whom a benign growth, large adenoma, or cancer has been detected. Thus, the performance of a given strategy $Z_{g,k}$ depends on (i) the values of utility nodes $U_i,\ i \in \{\mathrm{PCol}, \mathrm{R}, \mathrm{L}, \mathrm{B}\}$ corresponding to different scenario paths that are compatible with this strategy, and (ii) the probabilities of such scenario paths. To compute these probabilities, we define an *information set* $\mathcal{I}(j)$ as the set of direct predecessors of a given node j, i.e. $\mathcal{I}(j) = \{i \in \mathcal{N} | (i,j) \in \mathcal{A}\}$ (Salo et al. 2022). For instance, in Figure 1 the information set of node $j = 5$ (FIT result) consists of nodes $j = 1$ (FIT cut-off level) and $j = 4$ (usable sample is returned).

An *information state* $s_{\mathcal{I}(j)} \in \mathcal{S}_{\mathcal{I}(j)}$ is defined as the combination of states s_i for all nodes of the information set $i \in \mathcal{I}(j)$. In our previous example, the set of possible information states for node $j = 5$ is $\mathcal{S}_{\mathcal{I}(5)} = \prod_{i \in \mathcal{I}(5)} \mathcal{S}_i = \mathcal{S}_1 \times \mathcal{S}_4$. For chance nodes, the probability of outcome s_j depends on these information states. Specifically, the conditional probability of $s_j \in \mathcal{S}_j$ (where $j \in \mathcal{C}$) occurring is $\mathbb{P}[X_j = s_j | X_{\mathcal{I}(j)} = s_{\mathcal{I}(j)}]$, where $X_{\mathcal{I}(j)}$ are random variables X_i denoting the values of nodes in the information set $i \in \mathcal{I}(j)$. For example, in Figure 1 the conditional probabilities for different outcomes in node 4 (i.e., whether a usable sample is returned) depend on the decision in node 3

(i.e., whether the person has been invited to screen) and decision in node 2 (i.e., whether they were offered an incentive). The conditional probability tables for each chance node in Figure 1 can be found in the Supplementary material.

For the optimization model, we define auxiliary variables $\pi(s)$, which refer to probabilities of scenario paths s that are compatible with strategy $Z_{g,k} = [z(s_j)]_{j\in\mathcal{D}}$. These auxiliary variables are linked to the decision variables $z(s_j)$ through

$$\pi(s) = \begin{cases} p(s) = \prod\limits_{j\in C} \mathbb{P}(X_j = s_j | X_{\mathcal{I}(j)} = s_{\mathcal{I}(j)}), & \text{if } z(s_{\mathcal{I}(j)}) = s_j \ \forall j \in \mathcal{D} \\ 0, & \text{otherwise,} \end{cases} \tag{1}$$

where s_j and $\mathcal{I}(j)$ are taken from path s. For example, if on scenario path s' the decision alternatives are $s_1 = 10, s_2 = \text{No}$, but strategy $Z_{g,k} = [z(s_j)]_{j\in\mathcal{D}}$ corresponds to decision alternatives $s_1 = 40, s_2 = \text{No}$, then $\pi(s') = 0$, as the first decision for s' does not match the current strategy. These auxiliary variables can be used to define the four objectives of the optimization problem, of which the first is to be minimized and the rest are to be maximized:

$$\text{Expected share of target segment who undergo colonoscopy:} \quad P^{\text{col}} = \sum_s \pi(s) U_{\text{PCol}}(s) \tag{2}$$

$$\text{Expected share of target segment with detected benign adenoma:} \quad \tilde{\psi}_B = \sum_s \pi(s) U_{\text{B}}(s) \tag{3}$$

$$\text{Expected share of target segment with detected large adenoma:} \quad \tilde{\psi}_L = \sum_s \pi(s) U_{\text{L}}(s) \tag{4}$$

$$\text{Expected share of target segment with detected cancer:} \quad \tilde{\psi}_R = \sum_s \pi(s) U_{\text{R}}(s). \tag{5}$$

We formulate this multiobjective optimization problem of finding Pareto optimal strategies $Z_{g,k}$ for the ID of segment (g, k) as a Decision Programming model (Salo et al. 2022). Our formulation is both discrete and linear in outcomes, linear in the auxiliary decision variables $\pi(s)$, and utilizes a *Modified Augmented Weighted Tchebychev* (MAWT) norm approach (Holzmann & Smith 2018) to convert a multiobjective mixed-integer linear programming (MOMILP) problem into a single-objective problem. Technically, the problem is formulated as follows:

$$\min_{z(s_j)} \quad \mu \tag{6}$$

subject to

$$\mu \geq w_O |O - \phi_O^{utopia}| + \epsilon \sum_{O'\in\mathcal{O}} w_{O'} |O' - \phi_{O'}^{utopia}|, \qquad \forall\, O \in \mathcal{O} = \{P^{\text{col}}, \tilde{\psi}_B, \tilde{\psi}_L, \tilde{\psi}_R\} \tag{7}$$

$$\sum_s z(s_j) = 1, \qquad \forall\, j \in \mathcal{D},\ s_{\mathcal{I}(j)} \in \mathcal{S}_{\mathcal{I}(j)} \tag{8}$$

$$0 \leq \pi(s) \leq p(s), \qquad \forall\, s \in \mathcal{S} \tag{9}$$

$$\pi(s) \leq z(s_j), \qquad \forall\, s \in \mathcal{S}, j \in \mathcal{D} \tag{10}$$

$$\pi(s) \geq p(s) + \sum_{j \in \mathcal{D}} z(s_j) - |\mathcal{D}|, \qquad \forall\ s \in \mathcal{S} \tag{11}$$

$$z(s_j),\ U_i(s), \in \{0, 1\}, \qquad \forall\ j \in D,\ s \in \mathcal{S},\ i \in \{\text{PCol}, \text{B}, \text{L}, \text{R}\} \tag{12}$$

$$\pi(s) \in \mathbb{R} \qquad \forall\ s \in \mathcal{S}. \tag{13}$$

In this formulation, the single objective function in Equation (6) is represented by variable μ. This variable together with constraints (7) specifies the MAWT norm that is used to measure the distance of the original objective function vector $[P^{\text{col}}, \tilde{\psi}_B, \tilde{\psi}_L, \tilde{\psi}_R]$ from a utopian point ϕ^{utopia} consisting of objective function values that could be obtained by optimizing each objective individually. A more detailed description of the MAWT norm and the computation of parameters $\mathbf{w}$ and ϵ is presented by Holzmann & Smith (2018). Constraints (8)-(13) correspond to the Decision Programming formulation, and their purpose is to give a linear representation of Equation (1) linking the auxiliary decision variables $\pi(s)$ with the actual decision variables $z(s_j),\ j \in \mathcal{D}$. For details on these constraints, we refer the reader to Salo et al. (2022).

The optimization problem (6)-(13) is solved iteratively as summarized in the Supplementary material and detailed in Holzmann & Smith (2018), and produces the complete set of Pareto optimal solutions $Z_{g,k}$ for segment (g, k). For brevity, we denote the set $\{Z_{g,k}\}$ of Pareto optimal solutions corresponding to a given strategy $Z_{g,\to k-1}$ up to period $k-1$ as

$$\{Z_{g,k}\} = \text{MAWT}(ID(g, k, \psi_{g,k-1}(Z_{g,\to k-1})), \tag{14}$$

where $\psi_{g,k-1}(Z_{g,\to k-1})$ is the starting prevalence vector corresponding to strategy $Z_{g,\to k-1}$. Each solution $Z_{g,k} \in \{Z_{g,k}\}$ is combined with $Z_{g,\to k-1}$ to obtain a set of Pareto optimal strategies $Z_{g\to k}$ up to and including period k. The objective function values corresponding to these strategies are denoted by $P^{\text{col}}_{g,k}(Z_{g,\to k})$ and $\tilde{\psi}_{g,k,b}(Z_{g,\to k}),\ b \in \{\text{B}, \text{L}, \text{R}\}$.

3.3. Level 2: Identification of Pareto optimal sex-specific strategies

Level 2 combines the segment-specific strategies $Z_{g,k} \in \{Z_{g,k}\},\ k \in \{1, \ldots, K\}$ found on level 1 to produce sex-specific strategies $Z_{g,\to K}$ that minimize the prevalence of cancer in all age groups combined, while minimizing the number of colonoscopies performed to ensure that the population-level strategies identified on level 3 use colonoscopy resources efficiently. Additional objectives of minimizing the segment-specific prevalences of benign growths and large adenomas are included to ensure that the natural progression between different bowel states does not lead to dominated strategies in terms of total cancer prevalence at the end of Level 2. Moreover, a constraint on the expected number of performed colonoscopies is imposed to prune out strategies that would be infeasible on level 3. Taken together, the aim of level 2 for sex $g \in \{\text{F}, \text{M}\}$ can be formulated as a multiobjective optimization problem:

$$\min_{Z_{g,\to K}} .\quad \Psi_{g,R}(Z_{g,\to K}),\ \psi_{g,L}(Z_{g,\to K}),\ \psi_{g,B}(Z_{g,\to K}),\ N^{\text{col}}_{g,K}(Z_{g,\to K}) \tag{15}$$

$$\text{subject to} \quad N^{\text{col}}_{g,K}(Z_{g,\to K}) \leq N^{\text{col,max}}, \tag{16}$$

where $\Psi_{g,R}(Z_{g,\to K})$ is the combined cancer prevalence in all age groups, $\psi_{g,b}(Z_{g,\to K})$, $b \in \{\text{B}, \text{L}\}$ is the prevalence of bowel state b in segment (g, K), $N^{\text{col}}_{g,K}(Z_{g,\to K})$ is the expected total number of colonoscopies performed in all age groups, and $N^{\text{col,max}}$ is the maximum number of colonoscopies that can be carried out.

Problem (15)-(16) is solved by iterating through steps 1-4 in Process 1 for all $k \in \{1, \ldots, K\}$. The algorithm for carrying out this task is presented and discussed in Section 3.3.1. Section 3.3.2 presents the equations through which the prevalences $\psi_{g,k-1,b}$ of different bowel states $b \in \{\text{N}, \text{B}, \text{L}, \text{R}\}$ are updated between screening periods $k-1$ and k to account for the impact of screening as well as natural cancer progression due to aging (step 3 in Process 1).

3.3.1. Algorithm

The algorithm for solving problem (15)-(16) is presented in pseudocode in Algorithm 1. As main inputs, the algorithm requires 1) the ID describing the screening process for each segment (g, k) defined by age and sex, 2) vectors $\psi_{g,0} = [\psi_{g,0,b}]_{b \in \mathcal{B}}$ of starting prevalences for the first age segments (one per sex) and 3) K, the number of periods to be included. Algorithm 1 provides two sets of sex-specific strategies (one for each sex $g \in \{\text{F}, \text{M}\}$) that are Pareto optimal in view of minimizing the prevalence of cancer in all age groups combined as well as the number of colonoscopies performed.

The algorithm starts by initialising empty sets of efficient solutions (one per sex). The sexes are handled separately (line 2). The algorithm then proceeds on line 3 to compute the set of Pareto optimal strategies $\{Z_{g,\to 1}\}^{\text{PO}}$ for sex g for the first period (cf. Equation (15)) based on starting prevalences $\psi_{g,0}$ estimated from data. For each strategy $Z_{g,\to 1} \in \{Z_{g,\to 1}\}^{\text{PO}}$, the updated starting prevalences $\psi_1(Z_{g,\to 1})$ are computed on line 4 by accounting for the impact of screening as well as natural cancer progression due to aging (see Section 3.3.2 for a detailed description of prevalence update). On line 6, the total cancer prevalence $\Psi_{g,\to 1}(Z_{g,\to 1})$ corresponding to strategy $Z_{g,\to 1}$ in age groups up to and including period 1 is defined simply as the cancer prevalence in the current segment. Then, the total expected number $N^{\text{col}}_{g,\to 1}(Z_{g,\to 1})$ of colonoscopies carried out so far is computed on line 6 for each $Z_{g,\to 1} \in \{Z_{g,\to 1}\}^{\text{PO}}$ as the product of the expected share $P^{\text{col}}_{g,1}(Z_{g,1})$ of target segment who undergo colonoscopy in the given strategy and the number $N_{g,1}$ of invitees in the target segment.

The computation of strategies for the remaining periods $k \in \{2, \ldots, 5\}$ starts on line 7. For each Pareto optimal strategy $Z_{g,\to k-1} \in \{Z_{g,\to k-1}\}^{\text{PO}}$ up to the previous period, the set $\{Z_{g,k}\}$ of Pareto optimal strategies for the current segment (g, k) is solved using the MAWT algorithm with starting prevalences $\psi(Z_{g,\to k-1})$ on line 9 (cf. Equation (15)). Then, each strategy $Z_{g,k}$ in this set is combined with $Z_{g,\to k-1}$ on line 10 to obtain a strategy up to and including period k. The prevalences of different bowel states are then updated on line 12 as in the first period. On line 13, the total cancer prevalence $\Psi_{g,k,\text{R}}(Z_{g,\to k})$ is computed in segments corresponding to sex g and age groups up to and including period k (see Section 3.3.2 for details). The total expected number of colonoscopies carried out so far for sex g is updated on line 14 by adding to the total expected

Inputs : Process model as Influence Diagram (ID), prevalence vectors $\psi_{\mathrm{F},0}, \psi_{\mathrm{M},0}$, screening periods $k \in \{1 \dots K\}$

Output: Sex-specific Pareto optimal sets $\{Z_{\mathrm{F}}\}^{\mathrm{PO}}$ and $\{Z_{\mathrm{M}}\}^{\mathrm{PO}}$, and the related objective values $\Psi_{\mathrm{R},g}(Z_g), g \in \{\mathrm{F}, \mathrm{M}\}$.

$\{Z_{\mathrm{F},\to k}\}^{\mathrm{PO}}, \{Z_{\mathrm{M},\to k}\}^{\mathrm{PO}} = \varnothing \quad \forall k \in [1 \dots K]$; `// Initialize set of efficient solutions as empty sets.`

for $g \in \{\mathrm{F}, \mathrm{M}\}$ **do**

  `// First period,` $k = 1$`, handled separately.`

  `// Compute Pareto optimal solutions for first period`

  $\{Z_{g,\to 1}\}^{\mathrm{PO}} \leftarrow$ `MAWT` $(ID(g,1), \psi_{g,0}))$; `//` $\psi_{g,k} = [\psi_{g,k,b}]_{b \in \mathcal{B}}$

  $\psi_{g,1}(Z_{g,\to 1}) \leftarrow$ `UpdatePrevalences`$(Z_{g,1}, \psi_{g,0}) \quad \forall Z_{g,1} \in \{Z_{g,1}\}^{\mathrm{PO}}$;

  $\Psi_{g,\to 1,R}(Z_{g,\to 1}) \leftarrow \psi_{g,1,R}(Z_{g,1}) \quad \forall Z_{g,1} \in \{Z_{g,1}\}^{\mathrm{PO}}$;

  $N^{\mathrm{col}}_{g,\to 1}(Z_{g,1}) \leftarrow N_{g,1} P^{\mathrm{col}}_{g,1}(Z_{g,1}) \quad \forall Z_{g,1} \in \{Z_{g,1}\}^{\mathrm{PO}}$;

  for $k \in \{2, \dots, K\}$ **do**

    for $Z_{g,\to k-1} \in \{Z_{g,\to k-1}\}^{\mathrm{PO}}$ **do**

      `// Compute Pareto optimal solutions for period` k `based on preceding strategy` $Z_{g,\to k-1}$

      $\{Z_{g,k}\} \leftarrow$ `MAWT` $(ID(g,k,\psi_{g,k-1}(Z_{g,\to k-1})))$;

      for $Z_{g,k} \in \{Z_{g,k}\}$ **do**

        $Z_{g,\to k} \leftarrow (Z_{g,\to k-1}, Z_{g,k})$; `//Extend` $Z_{g,\to k-1}$ `by` $Z_{g,k}$ `into a new strategy` $Z_{g,\to k}$`, covering steps until` k

        $\psi_{g,k}(Z_{g,\to k}) \leftarrow$ `UpdatePrevalences`$(Z_{g,\to k}, \psi_{g,k-1})$;

        $\Psi_{g,\to k,\mathrm{R}}(Z_{g,\to k}) \leftarrow$ `TotalPrevalences` $(\psi_{g,k,\mathrm{R}}(Z_{g,\to k}), \Psi_{g,\to k-1,\mathrm{R}}(Z_{g,\to k-1}))$;

        `// Add new colonoscopies from` $Z_{g,k}$ `to preceding colonoscopies.`

        $N^{\mathrm{col}}_{g,\to k}(Z_{g,\to k}) \leftarrow N^{\mathrm{col}}_{g,\to k-1}(Z_{g,\to k-1}) + N_{g,k} P^{\mathrm{col}}_{g,k}(Z_{g,k})$;

        if $N^{\mathrm{col}}_{g,\to k}(Z_{g,\to k}) \le N^{\mathrm{col,max}}$ **then**

          $\{Z_{g,\to k}\}^{\mathrm{PO}} \leftarrow \{Z_{g,\to k}\}^{\mathrm{PO}} \bigcup Z_{g,\to k}$; `// Collect found strategies for step k.`

        end

      end

    end

    $\{Z_{g,\to k}\}^{\mathrm{PO}} \leftarrow$ `RemoveDom'd`$(\{Z_{g,\to k}\}^{\mathrm{PO}}, \Psi_{g,\to k,\mathrm{R}}(\cdot), \psi_{g,k,\mathrm{B}}(\cdot), \psi_{g,k,\mathrm{L}}(\cdot),\ N^{\mathrm{col}}_{g,\to k}(\cdot))$;

  end

end

$\{Z_{\mathrm{F}}\}^{\mathrm{PO}} \leftarrow \{Z_{\mathrm{F},\to K}\}^{\mathrm{PO}}, \{Z_{\mathrm{M}}\}^{\mathrm{PO}} \leftarrow \{Z_{\mathrm{M},\to K}\}^{\mathrm{PO}}$;

Algorithm 1: Level 2 algorithm

number of colonoscopies carried out up to the previous period the product of the expected share $P^{\mathrm{col}}_{g,k}(Z_{g,\to k})$ of the target segment who undergo colonoscopy in the given strategy and the number $N_{g,k}$ of invitees in the current target segment.

On lines 14-16, those strategies that are feasible with respect to a constraint on the total expected number of colonoscopies (line 15) are collected in set $\{Z_{g,\to k}\}^{\mathrm{PO}}$. Once the feasible, Pareto optimal strategies $Z_{g,\to k}$ have been found corresponding to all Pareto optimal strategies $Z_{g,\to k-1}$ from the previous period, those strategies are removed (using pairwise dominance checks) on line 20 from the set $\{Z_{g\to k}\}^{\mathrm{PO}}$ of Pareto optimal strategies which are dominated in view of minimizing (i) the expected prevalence $\psi_{g,k,\mathrm{B}}$ of benign growths in the current segment, (ii) the expected prevalence $\psi_{g,k,\mathrm{L}}$ of large growths in the current segment, (iii) the total expected prevalence $\Psi_{g,\to k,\mathrm{R}}$ of CRC in the population of sex g screened up to and including period k, and (iv) the total expected number $N^{\mathrm{col}}_{g,\to k}$ of colonoscopies required by segments corresponding to sex g up to and including period k. This removal of strategies that would most likely become infeasible or dominated in the upcoming periods helps to maintain computational efficiency. It can also be motivated by these plausible requirements concerning the performance of strategies up to and including period k. Yet, it is, in principle, possible that some of these strategies could be parts of feasible and Pareto-optimal strategies in the final period K. The characterization of those combinations of parameter values for which this may occur is left as a topic for further work.

3.3.2. Prevalence update

Step 3 in Process 1 (i.e., lines 4, 5, 12 and 13 in Algorithm 1) corresponds to computing the starting prevalences $\psi_{g,k,b}$ of different bowel states $b \in \mathcal{B}$ for sex g in period k. To do this, one must account for both the impact of screening as well as natural cancer progression through aging. To accommodate the impact of screening, we assume that any benign growth, large adenoma, or CRC found during the screening pathway is removed and that the bowel returns to a normal state[2]. The updated prevalence of bowel state $b \in \{\mathrm{N}, \mathrm{B}, \mathrm{L}, \mathrm{R}\}$ corresponding to screening strategy $Z_{g,k} \in \{Z_{g,k}\}$ is denoted by $\tilde{\psi}_{g,k,b}(Z_{g,k})$. Natural cancer progression due to aging is reflected in our model by transition probability $\mathbb{T}^{g,k}_{b,b'}$, i.e., the probability that the bowel state of a participant of sex g is b' in period $k+1$ given that it was b in period k. We assume that this progression follows the adenoma-carcinoma sequence, meaning the transition through bowel states can be represented by a linear recurrence relation. Based on discussions with the Finnish cancer registry, these assumptions can be deemed acceptable for the purposes of this paper. However, more refined transition models could be integrated into this modeling approach to improve accuracy.

Taking into account both the impact of screening as well as natural cancer progression, the starting prevalences $\psi_{g,k,b}$ of different bowel states $b \in \mathcal{B}$ for sex g in period k can be computed

[2]An individual receiving treatment will move on to a separate surveillance program and will no longer be a part of the population screening program. The assumption of the bowel returning to a normal state after the removal of a benign growth, large adenoma, or CRC can be seen to reflect the average dynamics of disease progression in the target population of the screening program.

using the following difference equations:

$$\psi_{g,k,B}(Z_{g,\to k}) = (\psi_{g,k-1,\mathrm{B}} - \tilde{\psi}_{g,k,B}(Z_{g,\to k}))(1 - \mathbb{T}^{g,k}_{\mathrm{B,L}}) + \psi_{g,k-1,\mathrm{N}}\mathbb{T}^{g,k}_{\mathrm{N,B}} \tag{17}$$

$$\psi_{g,k,\mathrm{L}}(Z_{g,\to k}) = (\psi_{g,k-1,\mathrm{L}} - \tilde{\psi}_{g,k,\mathrm{L}}(Z_{g,\to k}))(1 - \mathbb{T}^{g,k}_{\mathrm{L,R}}) + (\psi_{g,k-1,\mathrm{B}} - \tilde{\psi}_{g,k,\mathrm{B}}(Z_{g,\to k}))\mathbb{T}^{g,k}_{\mathrm{B,L}} \tag{18}$$

$$\psi_{g,k,\mathrm{R}}(Z_{g,\to k}) = \psi_{g,k-1,\mathrm{R}} - \tilde{\psi}_{g,k,\mathrm{R}}(Z_{g,\to k}) + (\psi_{g,k-1,\mathrm{L}} - \tilde{\psi}_{g,k,\mathrm{L}}(Z_{g,\to k}))\mathbb{T}^{g,k}_{\mathrm{L,R}} \tag{19}$$

$$\psi_{g,k,\mathrm{N}}(Z_{g,\to k}) = 1 - \sum_{b\in\{\mathrm{B,L,R}\}} \psi_{g,k,b}(Z_{g,\to k}) \quad . \tag{20}$$

For instance, in Equation (17) the prevalence estimate for benign growth (B) is updated by 1) computing the remaining prevalence after applying strategy $Z_{g,\to k}$ by subtracting the expected share of found benign growths $\tilde{\psi}_{g,k,\mathrm{B}}(Z_{g,\to k})$ from the previous prevalence estimate $\psi_{g,k-1,\mathrm{B}}$, 2) multiplying this remaining prevalence with the share of participants whose benign adenomas do not develop into large adenomas ($1 - \mathbb{T}^{g,k}_{\mathrm{B,L}}$), and finally 3) adding the share of participants in this segment who will develop benign adenomas from previously normal bowels. The logic for bowel states L (large growths) and R (cancer) is similar, with the exception of there not being further bowel states to which to develop from state R. Equation (20) simply states that participants with normal bowel states are those whose bowel sates are not B, L, or R. Equations (17)-(20) are compactly represented on line 12 of Algorithm 1 by function

$$\mathrm{UpdatePrevalences}(Z_{g,\to k}, \psi_{g,k-1}(Z_{g,\to k-1})) \tag{21}$$

The total prevalence $\Psi_{g,\to k+1,\mathrm{R}}$ of cancer in the target population is updated on line 13 of Algorithm 1 through function TotalPrevalences$(\cdot)$, defined as

$$\begin{aligned}\Psi_{g,\to k,\mathrm{R}}(Z_{g,\to k}) &= \mathrm{TotalPrevalences}(\psi_{g,k,\mathrm{R}}(Z_{g,\to k}), \Psi_{g,\to k-1,\mathrm{R}}(Z_{g,\to k-1})) \\ &= \frac{\Psi_{g,\to k-1,\mathrm{R}}(Z_{g,\to k-1}) \cdot (\sum_{i=1}^{k-1} N_{g,i}) + \psi_{g,k,\mathrm{R}} N_{g,k}}{\sum_{i=1}^{k} N_{g,i}}. \end{aligned} \tag{22}$$

The first term in the numerator is the expected number of cancer cases in segments corresponding to sex g and age groups up to and including period $k-1$, while the second term is the expected number of cancer cases in the current segment (g,k). The denominator is the total number of participants in segments corresponding to sex g and all age groups up to and including period k.

3.4. Level 3: Identification of an optimal screening strategy.

The aim of level 3 (step 5 in Process 1) is to choose the two Pareto optimal sex-specific strategies $Z_{F,\to K} \in \{Z_{\mathrm{F}}\}^{\mathrm{PO}}$ and $Z_{\mathrm{M},\to K} \in \{Z_{\mathrm{M}}\}^{\mathrm{PO}}$ identified on level 2 that together minimize the expected CRC prevalence in the entire screening population subject to a constraint on the expected total number of colonoscopies performed. Let $J_{\mathrm{F}} = |\{Z_{\mathrm{F}}\}^{\mathrm{PO}}|$ and $J_{\mathrm{M}} = |\{Z_{\mathrm{M}}\}^{\mathrm{PO}}|$ be the number of Pareto optimal strategies for females and males, respectively. Let $x_{\mathrm{F},j} \in \{0,1\},\ j \in \{1,\ldots,J_{\mathrm{F}}\}$ and $x_{\mathrm{M},j} \in \{0,1\},\ j \in \{1,\ldots,J_{\mathrm{M}}\}$ be binary decision variables each of which obtains a value of

1 if and only if the j-th strategy is selected from the set $\{Z_F\}^{\mathrm{PO}}$ or $\{Z_M\}^{\mathrm{PO}}$, respectively. Let $\Psi^j_{g,\to K,R}$ denote the total cancer prevalence in all age groups of sex g corresponding to strategy j. The task of identifying the optimal pair of sex-specific screening strategies can now be formulated as the binary linear programming problem

$$\min_{x_{\mathrm{F},j},\, x_{\mathrm{M},j}} \quad \Psi_R = \frac{N_{\mathrm{F}}}{N} \cdot \Psi^j_{\mathrm{F},\to K,R} \cdot x_{\mathrm{F},j} + \frac{N_{\mathrm{M}}}{N} \cdot \Psi^j_{\mathrm{M},\to K,R} \cdot x_{\mathrm{M},j} \tag{23}$$

$$\text{subject to} \quad N^{\mathrm{col}}_{\mathrm{F},j} x_{\mathrm{F},j} + N^{\mathrm{col}}_{\mathrm{M},j} x_{\mathrm{M},j} \leq N^{\mathrm{col,max}}, \tag{24}$$

$$\sum_j x_{g,j} = 1, \qquad \forall g \in \{\mathrm{F}, \mathrm{M}\} \tag{25}$$

$$x_{\mathrm{M},j} \in \{0,1\}, \qquad \forall j \in \{1, \ldots, J_{\mathrm{M}}\} \tag{26}$$

$$x_{\mathrm{F},j} \in \{0,1\}, \qquad \forall j \in \{1, \ldots, J_{\mathrm{F}}\}, \tag{27}$$

where $N_g = \sum_{k=1}^{K} N_{g,k}$ is the number of persons of sex g in the total population, N is the total population size, $N^{\mathrm{col}}_{g,j}$ is the expected number of colonoscopies performed in strategy j for sex g, and $N^{\mathrm{col,max}}$ is the maximum number of colonoscopies that can be performed for the entire population.

4. Results and recommendations for improving the Finnish CRC screening program

In this Section, we present optimal CRC screening strategies for the Finnish target population with and without the possibility of using monetary incentives to boost participation in the screening program. The case in which incentives are not available reflects the current Finnish CRC screening program. Hence, results corresponding to this case will provide recommendations for improving the current program by suggesting optimal FIT cut-off levels for different age and sex groups. Furthermore, these results will reveal information on the expected costs of both the current screening strategy as well as the optimized strategies. To our knowledge, no other cost analysis for the current Finnish CRC screening program has been published in the public domain. Results corresponding to the hypothetical case in which monetary incentives are available help assess the potential benefits as well as costs of such incentives, if they were to be utilized in the future. To examine the implications of different colonoscopy resource constraints on optimal screening strategies, health outcomes and costs, we compute the results by varying the maximum number of colonoscopies per screening round between 3,000 and 12,000. Strategies corresponding to these different values of colonoscopy capacity are compared to the baseline case of no screening as well as the current screening program. The input parameters used to compute the results (including target segment sizes, participation rate, FIT and colonoscopy sensitivities and specificities, probabilities of adverse events, prevalences of different bowel states, and costs) are in the Supplementary material.

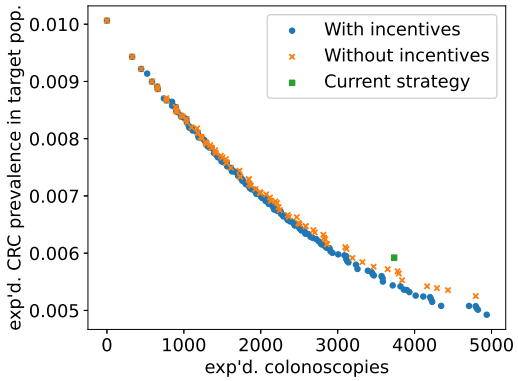
The number of potential screening strategies was $11^{10} \approx 29.5$ billion[3]. Pareto optimal sex-specific strategies were computed using a Julia-based algorithm on Aalto University's Triton HPC cluster. Computation time for males was 3 days and 21 hours, and for females over 6 days due to a higher number of feasible strategies. The memory requirements were approximately 4.4 and 4.8 Gb for males and females, respectively. Altogether 6,389 and 15,194 Pareto-optimal strategies were identified for males and females, respectively. The identification of those two Pareto-optimal sex-specific strategies that together minimized the expected CRC prevalence subject to a colonoscopy resource constraint (i.e., level 3) consumed significantly less resources and could be performed on a regular laptop computer.

4.1. Pareto optimal screening strategies

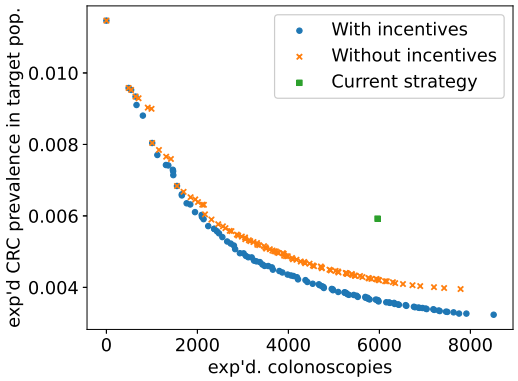
The sex-specific Pareto optimal screening strategies (i.e., the results of level 2) with and without monetary incentives are presented for females and males in Figures 2a and 2b respectively. In both figures, the current strategy (in which a FIT cut-off level of 25 μg Hb/g is used for all segments and no monetary incentives are used) is depicted with a green square. For both sexes the incentivized strategies dominate the non-incentivized, except for cases in which very few colonoscopies are performed. This is because incentives help boost participation in those segments where colonoscopies yield the highest health benefits: in particular, being able to detect abnormal bowel states in younger segments not only decreases cancer prevalence in these segments, but also decreases the risk of benign growths and large adenomas developing into cancers in subsequent periods. For females, the current screening strategy is practically Pareto optimal. Yet, for males, the current strategy is clearly dominated, implying that a lower CRC prevalence could be obtained with the same number of colonoscopies performed.

The expected CRC prevalences for the optimal population-level screening strategies (i.e., the outcomes of level 3) for different values of the maximum expected number of colonoscopies performed are depicted in Figure 3 by blue (incentivized case) and yellow circles (non-incentivized case). The optimal strategies did not change after increasing the maximum expected number of colonoscopies beyond 14,000, implying that this amount of colonoscopy resources would be sufficient for minimizing cancer prevalence in the population. The current strategy corresponding to 9,693 expected colonoscopies and an expected cancer prevalence of 0.51% is marked by a green square. The outcome of not screening at all (zero colonoscopies and a cancer prevalence of 1.07%) is depicted by a red cross. A comparison between the current strategy and Pareto optimal strategies suggests that the use of sex- and age-specific FIT cut-off levels could result in (i) an equal level of cancer prevalence with significantly fewer colonoscopies (i.e., 8,000 vs. 9,693) or (ii) a significant reduction in cancer prevalence with approximately the same number of colonoscopies, especially if measures can be taken to increase participation in the screening program.

[3]There are altogether 10 segments (g, k) corresponding to five screening periods k for both sexes g. For each segment, there are 11 possible strategies comprising a no-screening strategy and 10 screening strategies (five FIT cut-off levels, each with an option to use or not use incentives). Together, this results in 11^{10} potential strategies.



(a) Incentivized and non-incentivized Pareto optimal strategies for females.



(b) Incentivized and non-incentivized Pareto optimal strategies for males.

Figure 2: Sex-specific outcomes for Pareto-optimal strategies for both incentivized and non-incentivized cases.

In the Supplementary material, we present the results of sensitivity analyses examining the impacts of varying the values of all model parameters on expected CRC prevalence, the expected total number of colonoscopies, and expected total costs for optimal strategies corresponding to 3000, 10,000, and 14,000 colonoscopies. Expected cancer prevalence as well as expected total costs are most sensitive to changes in bowel state transition probabilities – faster transitions result in more cancers and, thus, higher treatment costs within the program. The expected number of colonoscopies is most sensitive to FIT specificity; the lower the specificity (i.e., the higher the chance of a false positive FIT result), the more colonoscopies will be carried out. Special attention should thus be paid to obtaining accurate estimates for FIT specificity to mitigate the risk of exceeding colonoscopy capacity.

Figure 4 shows the total direct costs together with the total number of colonoscopies for a case of no screening, the current strategy, and optimized strategies corresponding to different values of the maximum expected number of colonoscopies. Here, the direct costs are incurred by, e.g., the preparation and analysis of FITs, colonoscopies, and treatment (see the Supplementary material for details). These costs were not explicitly minimized in the analysis but may nevertheless play an important role in decision making due to, e.g., decision-makers wanting to avoid significant cost increases compared to current practices. Figure 4a shows that while the optimal use of incentives provides benefits in terms of the selected objectives, it also leads to significantly higher costs compared to the current strategy, except in cases in which the expected number of performed colonoscopies is low. This is due to the extra costs incurred by the incentives, the increased number of colonoscopies (and adverse events related to these colonoscopies) resulting from higher participation rate, and higher treatment costs due to a higher number of detected cancers. In the case of non-incentivized strategies, the total costs of an optimized strategy with a maximum of 7,000 expected colonoscopies coincide with those of the current strategy, in which only 6,482 are to be expected. This suggests that the use of sex- and age-specific FIT cut-off levels could help decrease the

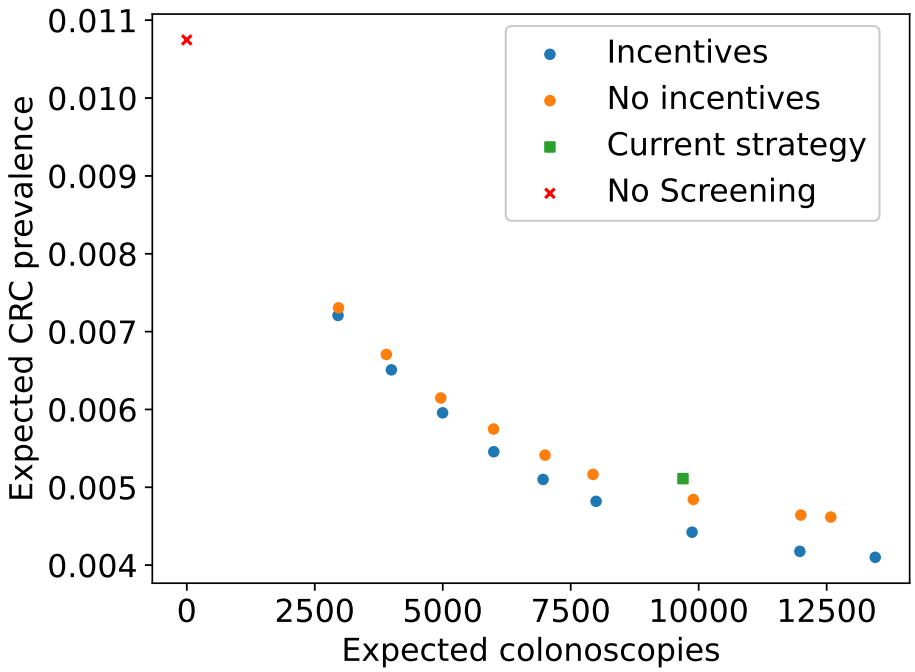
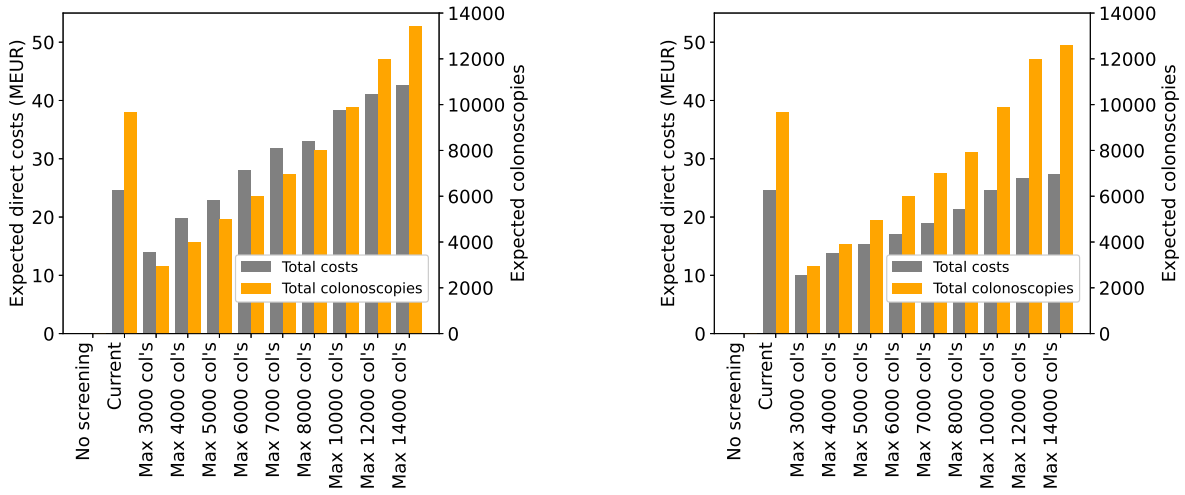


Figure 3: Outcomes of strategies that minimize cancer prevalence in the target population depending on available colonoscopies.

total costs even if the number of colonoscopies would remain the same. Finally, Figure 4 shows that the marginal increase in total costs decreases as the maximum number of expected colonoscopies increases. This can be explained by the fact that a higher level of colonoscopy resources helps detect a larger number of non-cancerous growths early on, which results in a decrease in cancer-related costs.

Finally, Tables 1 and 2 show the optimal segment-specific screening strategies for the non-incentivized and incentivized case, respectively. In both cases, a higher level of colonoscopy resources translates to lower FIT cut-off values. This is to be expected, since lower FIT cut-off values help detect and prevent more cancers due to the increased test sensitivity, but require more colonoscopies to be performed. By comparing Tables 1 and 2, it can be seen that boosting participation pays off in almost all target segments. Encouraging younger age segments to participate in the screening program is particularly effective: screening a higher number of people in the younger segments decreases not only cancer prevalence in these segments, but also the prevalence of earlier-stage growths and adenomas, which translates into lower cancer prevalence in later periods as well. Nevertheless, determining whether the use of incentives is cost-effective would necessitate a more thorough comparison between the added health benefits (Figure 3) and the added costs (Figure 4) of such incentives.

Examining the optimal, non-incentivized strategy corresponding to 10,000 colonoscopies provides interesting insights into how the current screening strategy with 9,693 colonoscopies and a fixed FIT



(a) Incentives used.

(b) Incentives not used.

Figure 4: Development of direct costs vs. number of colonoscopies.

cut-off level of 25 μg Hb/g for both sexes could be improved by adjusting the FIT cut-off levels. In particular, the optimized strategy suggests that the cut-off level for females should be decreased across all age groups. Consequently, to satisfy the constraint on the maximum expected number of colonoscopies, the cut-off level should be increased in older age segments for males. In summary, all of the above results suggest that efforts should be taken to detect abnormal bowel states in as many people and as early as possible to help minimize overall cancer prevalence in the population.

Table 1: Optimized screening policies with no option of incentive: FIT cut-off level in μg Hb/g, no invite represented by -, colonoscopy performed when positive FIT.

Maximum colonoscopies	Sex	Age 60	Age 62	Age 64	Age 66	Age 68
3,000	Male	70	70	70	-	-
	Female	10	55	-	-	-
4,000	Male	70	70	70	70	-
	Female	10	55	70	-	-
5,000	Male	55	70	70	70	-
	Female	10	10	40	-	-
6,000	Male	40	40	55	70	-
	Female	10	10	10	-	-
7,000	Male	25	40	70	70	-
	Female	10	10	10	10	-
8,000	Male	25	25	55	70	70
	Female	10	10	10	10	-
10,000	Male	10	25	25	55	70
	Female	10	10	10	10	10
12,000	Male	10	10	10	10	40
	Female	10	10	10	10	10
14,000	Male	10	10	10	10	10
	Female	10	10	10	10	10

Table 2: Optimized screening policies with option of incentive (+i): FIT cut-off level in μg Hb/g, no invite represented by -, colonoscopy performed when positive FIT.

Maximum colonoscopies	Sex	Age 60	Age 62	Age 64	Age 66	Age 68
3,000	Male	70	70+i	70+i	-	-
	Female	25	55	-	-	-
4,000	Male	55+i	55+i	70+i	-	-
	Female	10+i	70	70	-	-
5,000	Male	55+i	55+i	70+i	70	-
	Female	10+i	70	25	40	-
6,000	Male	40+i	70+i	70+i	70+i	-
	Female	10+i	10+i	25	-	-
7,000	Male	40+i	55+i	70+i	70+i	-
	Female	10+i	10+i	10+i	55	-
8,000	Male	40+i	25+i	55+i	70+i	-
	Female	10+i	10+i	10+i	25	-
10,000	Male	25+i	25+i	25+i	55+i	70
	Female	10+i	10+i	10+i	10+i	-
12,000	Male	10+i	10+i	25+i	25+i	70
	Female	10+i	10+i	10+i	10+i	40
14,000	Male	10+i	10+i	10+i	25+i	10
	Female	10+i	10+i	10+i	10+i	40

4.2. Distribution of resources and benefits between segments

We also study the distribution of resources and benefits between segments to understand whether the proposed strategies might pose a problem from an equality perspective. Toward this end, Figures 5 and 6 show the distribution of (i) colonoscopies, (ii) the expected share of remaining cases of CRC in the target population after screening, and (iii) the expected share of reduction in cancers compared to no screening by segment for non-incentivized and incentivized strategies, respectively. Differences between age groups are not considered relevant from an equality perspective since all participants are assumed to proceed from the first age group to the last. Therefore, we limit our discussion on equality issues to differences between the sexes.

In both incentivized and non-incentivized strategies, males receive a slightly higher share of colonoscopies and a larger share of reduction in cancers. This is because the FIT sensitivity for males is much higher for all abnormal bowel states and all FIT cut-off levels compared to females. Thus, males receive a positive FIT result more often than females, whereby they are also more likely to undergo a colonoscopy examination and be treated for cancer. Consequently, the share of remaining cancers is higher for females, especially when incentives are used to boost participation. This is explained by our assumption that using an incentive for a given segment halves the number of non-adherent persons in that segment. Because the share of non-adherent persons in male segments is larger than that in female segments, incentives have a stronger effect in reducing cancers in males.

By comparing the current strategy with 9,693 colonoscopies to the optimal non-incentivized strategy with 10,000 colonoscopies, it can be seen that the use of optimized segment-specific FIT cut-off levels would (i) increase the share of colonoscopies carried out for females from 38.5% to 48.5%, (ii) decrease the share of remaining cancers in females from 59.0% to 55.5%, and (iii) increase the share of reduction in all cancers in females from 37.8% to 41.8%. Thus, it can be concluded

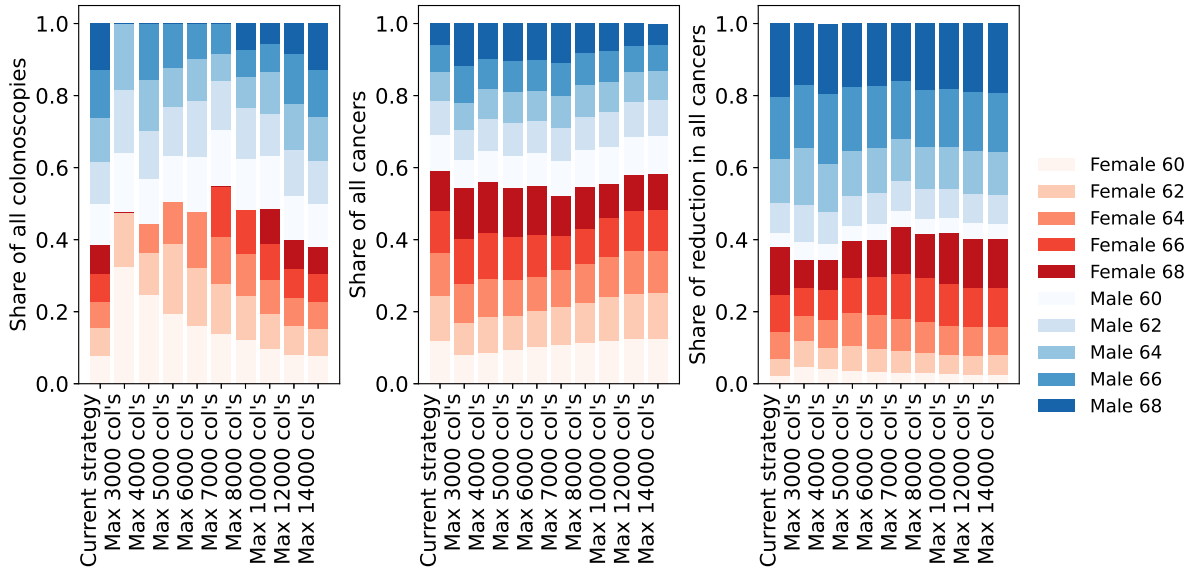


Figure 5: Distribution of expected colonoscopies, cancer prevalence and health benefits between segments, when no incentives are allowed.

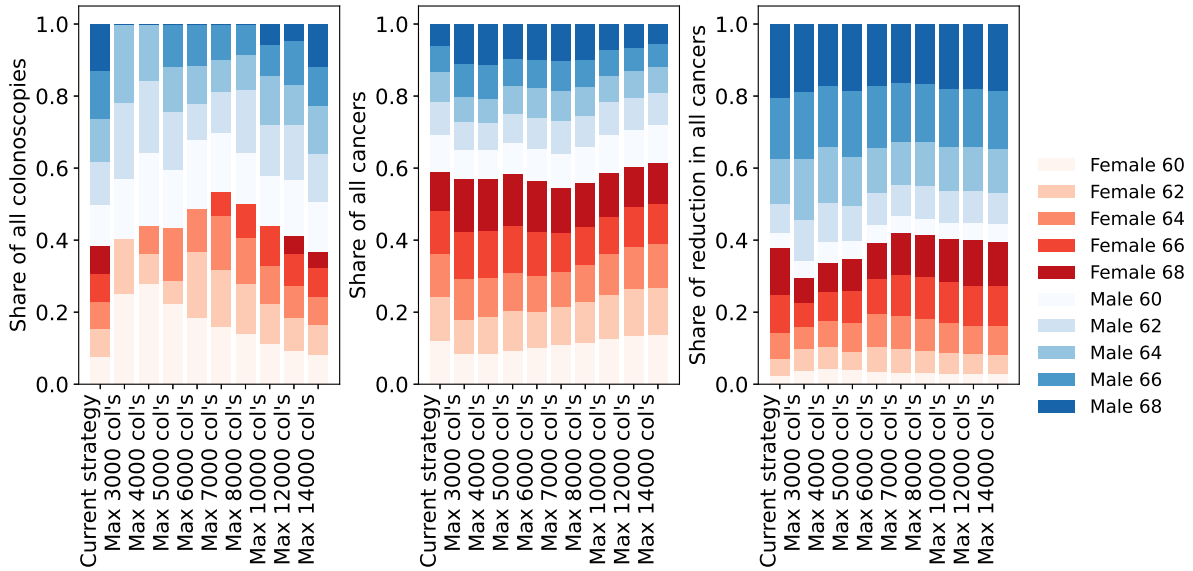


Figure 6: Distribution of expected colonoscopies, cancer prevalence and health benefits between segments when incentives can be used to increase participation rate.

that the use of optimal (non-incentivized) strategies could help improve the current strategy from an equality perspective as well.

5. Discussion and Conclusions

We have proposed a novel multilevel optimization model for improving the Finnish CRC program. In the model, influence diagrams are utilized to capture the decisions, uncertainties, and outcomes related to the screening program for each segment of the target population defined by the participants' sex and age. Then, an algorithm based on Decision Programming is used to identify Pareto optimal screening strategies for both sexes in view of minimizing the prevalence of CRC and the probability of performing colonoscopies. Finally, a binary linear programming problem is formulated to identify the combination of sex-specific strategies that together minimize the prevalence of CRC in the entire target population subject to a constraint on the total expected number of colonoscopies performed.

Compared to existing approaches, our model enables the identification of optimal screening strategies under resource constraints even in cases where these strategies consist of a number of different decisions the outcomes of which are subject to several chance events. In particular, our model makes it possible to determine for each segment separately (i) the optimal FIT cut-off levels for performing a colonoscopy and (ii) whether it pays off to use incentives to boost participation in the program. In this way, the model can help allocate scarce colonoscopy resources in a more cost-effective way.

Our results offer several insights into how the Finnish CRC population screening program could be improved. First, the FIT cut-off level for females should be decreased, especially in younger age groups. A similar result is obtained by Heinävaara et al. (2022), who use simulation to study the cost-effectiveness of different CRC screening strategies in the Finnish context. Yet, to satisfy the constraint on the maximum expected number of colonoscopies, our model suggests that the cut-off level should be increased in older age segments, especially for males. Moreover, boosting participation pays off in almost all target segments, and encouraging younger age segments in particular to participate in the screening program is highly effective. This is also in line with the results presented by Güneş et al. (2015) and Ladabaum et al. (2019), who suggest that increasing screening compliance is beneficial. In summary, efforts should be taken to detect abnormal bowel states in as many people and as early as possible to help minimize overall cancer prevalence in the population.

At present, our model has some limitations. First, our assumptions on disease progression are relatively strong. Improvements in the ways in which these processes are modeled could lead to more accurate estimates of the model parameters and, thereby, better strategy recommendations. Second, due to lack of detailed data on the connection between CRC-related deaths and prevalences of the different adenoma-carcinoma stages, we did not include the minimization of mortality as an objective. Nevertheless, taking this objective into account (or even demonstrating the impacts of optimal strategies on CRC-related mortality, as is done by Erenay et al. 2014) would be relevant

and could have an impact on the recommended strategies. We therefore suggest that the model is adjusted accordingly when the required data becomes available.

A third limitation of our proposed model is that we have assumed that no information about a participant's risk level for abnormal bowel states is available beyond their sex and age. Yet, this risk level is known to be elevated for people with a history of CRC or adenomatous polyps. This issue is not highly relevant for the current paper, which focuses on population screening and not surveillance. If screening and surveillance were to be optimized as a whole, our model could be extended to cover different risk levels for each segment, as has been done by Erenay et al. (2014). Risk levels for abnormal bowel states are also known to depend on factors such as family history, comorbidity, and the presence of high- or moderate-impact pathogenic variants in CRC susceptibility genes (Fuchs et al. 1994, Boakye et al. 2018, Tamlander et al. 2024). Population screening programs have traditionally been purposefully designed not to require such detailed information on the participants' risk levels – rather, increased risk due to factors such as family history or comorbidity is to be detected via other health pathways, which may result in suggestions for more frequent FIT or even colonoscopies. Yet, if the model were to be used to develop a targeted, risk-based screening program (as has been advocated in recent studies by, e.g., Lansdorp-Vogelaar et al. 2022 and Tamlander et al. 2024), information on different risk factors could be integrated by introducing new dimensions to the segmenting of the target population. Computationally, however, the models resulting from such an integration would be less tractable.

Fourth, we have used a limited, discrete scale for possible FIT cut-off values. The use of a denser discretization or even a continuous scale could suggest strategies that offer greater health benefits with the same capacity for colonoscopies. Yet, these kinds of modifications would require algorithmic efforts to keep the computation time manageable or, in the case of continuous cut-off levels, modifications to the Decision Programming framework. Finally, our proposed multilevel optimization model is computationally quite intensive as it, in practice, generates a multi-periodic decision tree to be solved. However, due to the nature of the algorithm, the main memory requirement is set by the structure of the influence diagram, i.e., the segment-specific model, whereas the addition of periods increases only the computation time. Furthermore, the introduction of a more detailed segment structure (to accommodate, e.g., the impacts of family history or comorbidity on a participant's risk for abnormal bowel states) would not increase the computation time either, if multiple computer units or nodes are available for computation. This is because the Pareto optimal strategies can be computed separately for all segments over all screening periods, after which the optimal combination of these segment-specific strategies can be determined on level 3 by solving a binary linear programming problem. Nevertheless, if computation time becomes an issue, the introduction of problem-specific constraints that would cut off infeasible or suboptimal branches from the decision tree can be useful. In the context of this paper, for instance, we have eliminated strategy branches that exhaust capacity constraints for colonoscopies before the final period, or lead to prevalences of abnormal bowel states that would almost certainly lead to suboptimal outcomes in terms of cancer prevalence in the final period.

This paper opens several avenues for future research. First, despite our efforts to find reliable sources for the assumptions underlying our model and its input parameters, continued work is needed to validate it further. At the time of writing this paper, the screening program had been running only for two years, and consequently we were only able to use data on the first age segment (60-year-olds) to validate whether our model produces realistic results on the number of colonoscopies and identified cancers. For a more thorough validation, data on all segments and bowel state prevalences would be needed to ensure that the predictions made by our model coincide with observed evidence. Such data would be useful in calibrating the model, particularly in terms of obtaining more accurate information on FIT sensitivities and transition probabilities, which, based on our sensitivity analyses, are the key parameters affecting the performance of the optimized strategies.

In the meantime, our results could be validated by using, e.g., the MISCAN model, which is widely used in cancer screening research. Whereas running the strategies identified in this paper in MISCAN should be relatively straightforward, implementing similar optimization capabilities directly in MISCAN would likely require the use of heuristic approaches, such as genetic algorithms. This could prove beneficial if there is a large difference in the performance of the found strategies given by our model and the MISCAN model. Moreover, a similar modeling approach to the one proposed in this paper could be applied to the screening of CRC as well as other cancers and slowly developing diseases in Finland or in other countries. Here the Decision Programming framework can prove to be quite flexible as defining the process is relatively intuitive using IDs and the optimization algorithms require updates only to objectives and constraints. It also allows flexible objective and constraint definitions. Finally, fully leveraging the probabilistic constraint capabilities of Decision Programming over periods and segments still lacks a rigorously justified approach, which could prove useful in, e.g., risk mitigation.

6. Acknowledgements

The calculations presented above were performed using computer resources within the Aalto University School of Science "Science-IT" project.

This research has been supported by The Foundation for Economic Education (grant number 16-9442) and The Paulo Foundation.

References

Advisory Board of OSF (2022). Official Statistics of Finland: Causes of deaths (web publication). Helsinki. URL: `https://pxweb2.stat.fi/PxWeb/pxweb/en/StatFin/StatFin__ksyyt/statfin_ksyyt_pxt_11br.px/` referred September 26, 2022.

Alagoz, O., Chhatwal, J., & Burnside, E. S. (2013). Optimal policies for reducing unnecessary follow-up mammography exams in breast cancer diagnosis. *Decision Analysis*, *10*, 200–224.

Ayer, T., Alagoz, O., & Stout, N. K. (2012). OR Forum—A POMDP approach to personalize mammography screening decisions. *Operations Research*, *60*, 1019–1034. URL: http://pubsonline.informs.org/doi/abs/10.1287/opre.1110.1019. doi:10.1287/opre.1110.1019.

Bertsimas, D., Silberholz, J., & Trikalinos, T. (2018). Optimal healthcare decision making under multiple mathematical models: Application in prostate cancer screening. *Health Care Management Science*, *21*, 105–118.

Boakye, D., Rillmann, B., Walter, V., Jansen, L., Hoffmeister, M., & Brenner, H. (2018). Impact of comorbidity and frailty on prognosis in colorectal cancer patients: a systematic review and meta-analysis. *Cancer Treatment Reviews*, *64*, 30–39.

Cevik, M., Ayer, T., Alagoz, O., & Sprague, B. L. (2018). Analysis of mammography screening policies under resource constraints. *Production and Operations Management*, *27*, 949–972.

Diedrich, L., Brinkmann, M., Dreier, M., Rossol, S., Schramm, W., & Krauth, C. (2023). Is there a place for sigmoidoscopy in colorectal cancer screening? A systematic review and critical appraisal of cost-effectiveness models. *Plos one*, *18*, e0290353.

Dillon, M., Flander, L., Buchanan, D. D., Macrae, F. A., Emery, J. D., Winship, I. M., Boussioutas, A., Giles, G. G., Hopper, J. L., Jenkins, M. A. et al. (2018). Family history–based colorectal cancer screening in Australia: A modelling study of the costs, benefits, and harms of different participation scenarios. *PLoS Medicine*, *15*, e1002630.

Ellison, L., Cheli, C. D., Bright, S., Veltri, R. W., & Partin, A. W. (2002). Cost-benefit analysis of total, free/total, and complexed prostate-specific antigen for prostate cancer screening. *Urology*, *60*, 42–46.

Erenay, F. S., Alagoz, O., Banerjee, R., & Cima, R. R. (2011). Estimating the unknown parameters of the natural history of metachronous colorectal cancer using discrete-event simulation. *Medical Decision Making*, *31*, 611–624.

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

Finnish Cancer Registry (2017). Cancer in Finland. URL: https://cancerregistry.fi/statistics/cancer-in-finland/ accessed on 2019-12-04.

Finnish Cancer Registry (2019). Colorectal cancer screening. URL: https://cancerregistry.fi/screening/colorectal-cancer-screening/ accessed on 27.08.2019.

Finnish Cancer Registry (2022). Cancer statistics. URL: https://tilastot.syoparekisteri.fi/syovat data from 2022-07-04, version 2022-08-16-001.

Finnish colorectal cancer screening expert groups (2021). Finnish colorectal cancer screening protocol. URL: https://syoparekisteri.fi/assets/files/2021/11/Protocol-for-and-tests-used-in-colorectal-cancer-screening.pdf.

Fuchs, C. S., Giovannucci, E. L., Colditz, G. A., Hunter, D. J., Speizer, F. E., & Willett, W. C. (1994). A prospective study of family history and the risk of colorectal cancer. *New England Journal of Medicine*, *331*, 1669–1674.

Güneş, E. D., Örmeci, E. L., & Kunduzcu, D. (2015). Preventing and diagnosing colorectal cancer with a limited colonoscopy resource. *Production and Operations Management*, *24*, 1–20.

Gupta, N., Bansal, A., Wani, S. B., Gaddam, S., Rastogi, A., & Sharma, P. (2011). Endoscopy for upper GI cancer screening in the general population: a cost-utility analysis. *Gastrointestinal Endoscopy*, *74*, 610–624.

Gyrd-Hansen, D., Søgaard, J., & Kronborg, O. (1997). *Colorectal Cancer Screening-Efficiency and Cost-effectiveness*. Centre for Health and Social Policy, Odense University.

Hardy, R. G., Meltzer, S. J., & Jankowski, J. A. (2000). Molecular basis for risk factors. *BMJ*, *321*, 886–889.

Heinävaara, S., Gini, A., Sarkeala, T., Anttila, A., de Koning, H., & Lansdorp-Vogelaar, I. (2022). Optimizing screening with faecal immunochemical test for both sexes – Cost-effectiveness analysis from Finland. *Preventive Medicine*, *157*, 106990.

Heitman, S. J., Hilsden, R. J., Au, F., Dowden, S., & Manns, B. J. (2010). Colorectal cancer screening for average-risk north americans: an economic evaluation. *PLoS Medicine*, *7*, e1000370.

Holzmann, T., & Smith, J. C. (2018). Solving discrete multi-objective optimization problems using modified augmented weighted Tchebychev scalarizations. *European Journal of Operational Research*, *271*, 436–449.

Kewenter, J., Brevinge, H., Engaras, B., Haglind, E., & Ährén, C. (1994). Results of screening, rescreening, and follow-up in a prospective randomized study for detection of colorectal cancer by fecal occult blood testing: Results for 68,308 subjects. *Scandinavian Journal of Gastroenterology*, *29*, 468–473.

Kronborg, O., Fenger, C., Olsen, J., Jørgensen, O. D., & Søndergaard, O. (1996). Randomised study of screening for colorectal cancer with faecal-occult-blood test. *The Lancet*, *348*, 1467–1471.

Ladabaum, U., Mannalithara, A., Meester, R. G., Gupta, S., & Schoen, R. E. (2019). Cost-effectiveness and national effects of initiating colorectal cancer screening for average-risk persons at age 45 years instead of 50 years. *Gastroenterology*, *157*, 137–148.

Lansdorp-Vogelaar, I., Meester, R., de Jonge, L., Buron, A., Haug, U., & Senore, C. (2022). Risk-stratified strategies in population screening for colorectal cancer. *International Journal of Cancer*, *150*, 397–405.

Lee, E., Lavieri, M. S., & Volk, M. (2019). Optimal screening for hepatocellular carcinoma: A restless bandit model. *Manufacturing & Service Operations Management*, *21*, 198–212.

Li, X., Cheung, W. K., & Liu, J. (2009). Improving pomdp tractability via belief compression and clustering. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *40*, 125–136.

Mandel, J. S., Bond, J. H., Church, T. R., Snover, D. C., Bradley, G. M., Schuman, L. M., & Ederer, F. (1993). Reducing mortality from colorectal cancer by screening for fecal occult blood. *New England Journal of Medicine*, *328*, 1365–1371.

Mandel, J. S., Church, T. R., Ederer, F., & Bond, J. H. (1999). Colorectal cancer mortality: effectiveness of biennial screening for fecal occult blood. *Journal of the National Cancer Institute*, *91*, 434–437.

McLay, L. A., Foufoulides, C., & Merrick, J. R. (2010). Using simulation-optimization to construct screening strategies for cervical cancer. *Health Care Management Science*, *13*, 294–318.

Neuvonen, L., Wildemeersch, M., & Vilkkumaa, E. (2023). Supporting strategy selection in multiobjective decision problems under uncertainty and hidden requirements. *European Journal of Operational Research*, *307*, 279–293.

Pence, B. C., Belasco, E. J., & Lyford, C. P. (2013). Combination aspirin and/or calcium chemoprevention with colonoscopy in colorectal cancer prevention: cost-effectiveness analyses. *Cancer Epidemiology, Biomarkers & Prevention*, *22*, 399–405.

Peng, L., Balavarca, Y., Niedermaier, T., Weigl, K., Hoffmeister, M., & Brenner, H. (2020). Risk-adapted cutoffs in colorectal cancer screening by fecal immunochemical tests. *Official Journal of the American College of Gastroenterology*, *115*, 1110–1116.

Rauner, M. S., Gutjahr, W. J., Heidenberger, K., Wagner, J., & Pasia, J. (2010). Dynamic policy modeling for chronic diseases: Metaheuristic-based identification of Pareto-optimal screening strategies. *Operations Research*, *58*, 1269–1286.

Salo, A., Andelmin, J., & Oliveira, F. (2022). Decision programming for mixed-integer multi-stage optimization under uncertainty. *European Journal of Operational Research*, *299*, 550–565.

Selby, K., Jensen, C. D., Lee, J. K., Doubeni, C. A., Schottinger, J. E., Zhao, W. K., Chubak, J., Halm, E., Ghai, N. R., Contreras, R. et al. (2018). Influence of varying quantitative fecal immunochemical test positivity thresholds on colorectal cancer detection: A community-based cohort study. *Annals of Internal Medicine*, *169*, 439–447.

Silva-Illanes, N., & Espinoza, M. (2018). Critical analysis of Markov models used for the economic evaluation of colorectal cancer screening: a systematic review. *Value in Health*, *21*, 858–873.

Simon, K. (2016). Colorectal cancer development and advances in screening. *Clinical Interventions in Aging*, *11*, 967–976.

Tamlander, M., Jermy, B., Seppälä, T. T., Färkkilä, M., FinnGen, Widén, E., Ripatti, S., & Mars, N. (2024). Genome-wide polygenic risk scores for colorectal cancer have implications for risk-based screening. *British Journal of Cancer*, (pp. 1–9).

Underwood, D. J., Zhang, J., Denton, B. T., Shah, N. D., & Inman, B. A. (2012). Simulation optimization of PSA-threshold based prostate cancer screening policies. *Health Care Management Science*, *15*, 293–309.

Van Der Meulen, M. P., Kapidzic, A., Van Leerdam, M. E., Van Der Steen, A., Kuipers, E. J., Spaander, M. C., De Koning, H. J., Hol, L., & Lansdorp-Vogelaar, I. (2017). Do men and women need to be screened differently with fecal immunochemical testing? A cost-effectiveness analysis. *Cancer Epidemiology Biomarkers and Prevention*, *26*, 1328–1336.

Whyte, S., Thomas, C., Chilcott, J., & Kearns, B. (2022). Optimizing the design of a repeated fecal immunochemical test bowel cancer screening programme with a limited endoscopy capacity from a health economic perspective. *Value in Health*, *25*, 954–964.

Wilschut, J. A., Habbema, J. D. F., van Leerdam, M. E., Hol, L., Lansdorp-Vogelaar, I., Kuipers, E. J., & van Ballegooijen, M. (2011). Fecal occult blood testing when colonoscopy capacity is limited. *Journal of the National Cancer Institute*, *103*, 1741–1751.

World Health Organization (2017). Early diagnosis and screening. URL: `https://www.who.int/cancer/prevention/diagnosis-screening/en/` last accessed on 04/11/2019.

World Health Organization (2018). Cancer fact sheet. URL: `https://www.who.int/news-room/fact-sheets/detail/cancer` last accessed on 18/103/2020.

Young, D., Haney, W., & Cremaschi, S. (2021). Derivative-free optimization of combinatorial problems–A case study in colorectal cancer screening. *Computers & Chemical Engineering*, *145*, 107193.

Highlights

- We develop a model to optimize strategies for two-stage colorectal cancer screening
- The model accounts for different cutoffs for first-stage tests based on age and sex
- Optimal strategies under colonoscopy constraints are found by Decision Programming
- Identified Pareto-optimal strategies minimize cancer prevalence and colonoscopies
- The results offer insights into improving the Finnish cancer screening program