Aus dem Deutschen Krebsforschungszentrum (DKFZ) Heidelberg
(Wissenschaftlicher Stiftungsvorstand: Prof. Dr. Michael Baumann)
Abteilung Gesundheitsökonomie
(Leiter: Prof. Dr. Michael Schlander)

Cost-effectiveness analysis of colorectal cancer screening strategies in Germany

Inauguraldissertation
zur Erlangung des medizinischen Doktorgrades
der
Medizinischen Fakultät Mannheim
der Ruprecht-Karls-Universität
zu
Heidelberg

vorgelegt von
Chih-Yuan Cheng

aus
Tainan, Taiwan

2022

Dekan: Prof. Dr. med. Sergij Goerdt

Referent: Prof. Dr. med. Michael Schlander

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AA	Advanced adenoma
ABC	Approximate Bayesian computation
ABC-SMC	Approximate Bayesian computation sequential Monte Carlo
ACG	American College of Gastroenterology
APMC	adaptive population Monte Carlo
CCE	Colon capsule endoscopy
CEA	Cost-effectiveness analysis
CEAC	Cost-effectiveness acceptance curve
CI	Confidence interval
CIMP	CpG island methylator phenotype
CIN	Chromosomal instability
CISNET	Cancer Intervention and Surveillance Modelling Network
CRC	Colorectal cancer
CrI	Credible interval
crSP	Clinically relevant serrated polyps
CTC	Computer tomography colonography
DECAS	Discrete Event simulation model for the natural history of Colorectal cancer from the Adenoma and Serrated neoplasia pathways
DES	Discrete event simulation
FIT	Fecal immunochemical test
FS	Flexible sigmoidoscopy
G-DRG	German diagnosis-related group
gFOBT	Guaiac-based fecal occult blood test
GGPO	German Guideline Program in Oncology
HDI	Human Development Index
HP	Hyperplastic polyps
HR	Hazard ratio
ICER	Incremental cost-effectiveness ratio
INNC	Incremental number-needed-to-colonoscope
IQR	Interquartile interval
IQWiG	*Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen* (Institute for Quality and Efficiency in Health Care)
LY	Life-year
LYG	Life-year gained

MCMC	Markov Chain Monte Carlo
mtsDNA	Multitarget stool DNA test
NHB	Net health benefit
NNC	Number-needed-to-colonoscope
PSA	Probabilistic sensitivity analysis
QALY	Quality-adjusted life-year
QALYG	Quality-adjusted life-year gained
RCT	Randomized control trial
SSL	Sessile serrated lesions
STM	State-transition model
TSA	Traditional serrated adenomas
UICC	Union for International Cancer Control
UKFSS	UK Flexible Sigmoidoscopy Screening
USPSTF	US Preventive Services Task Force
WTP	Willingness-to-pay
ZfKD	*Zentrum für Krebsregisterdaten* (German Centre for Cancer Registry Data)

LIST OF TABLES

LIST OF FIGURES

LIST OF SUPPLEMENTARY TABLES

LIST OF SUPPLEMENTARY FIGURES

1 INTRODUCTION

1.1 Colorectal cancer

1.1.1 Epidemiology

Colorectal cancer (CRC) is the third most common cancer and ranks the second among cancers with highest mortality in the world (Sung et al., 2021). Number of new CRC cases and deaths due to CRC globally were estimated to be 1.9 million and 0.9 million in 2020 respectively, namely one tenth of new cancer cases and deaths were attributed to CRC (Sung et al., 2021). In general, the CRC incidence and mortality in women are approximately 25% lower than in men (Dekker et al., 2019). The incidence of CRC is positively correlated with the socioeconomic development: CRC incidence is usually higher in the countries with higher Human Development Index (HDI), which reflects life expectancy, education and income level in a country (Fidler et al., 2016). It is therefore not surprising that Europe is among the regions with highest CRC incidence (Sung et al., 2021). In Germany, 58,100 new cases and 24,048 CRC deaths are projected in 2022, among the top three cancers in terms of the number incidence and mortality cases (Robert Koch Institut, 2021).

1.1.2 Pathogenesis and classification

Around 35-40% of CRC cases are associated with hereditary factors (Jasperson et al., 2010; Keum and Giovannucci, 2019). Majority of these hereditary CRC cases have a family history without any obvious genetic cancer syndrome, while approximately 5% are related to some inherited cancer syndromes, e.g., hereditary nonpolyposis colorectal cancer (also known as Lynch syndrome) or familial adenomatous polyposis (Jasperson et al., 2010). On the other hand, about 60-65% of CRC are sporadic cases, which most likely arise through acquired somatic genetic and epigenetic aberrations (Jasperson et al., 2010; Keum and Giovannucci, 2019). These anomalies are largely attributable to potentially modifiable risk factors, for example, diet high of red or processed meat, sedentary lifestyle which leads to decreased physical activity and excess body weight, smoking, and alcohol consumption (Keum and Giovannucci, 2019; Sung et al., 2021). Increased animal-source food intake and sedentary lifestyle likely contribute the most to the high CRC incidence in aforementioned high HDI countries and the increasing trend in those transitioning to higher HDI countries (Sung et al., 2021).

CRCs mostly arise from a benign precursor, defined as a polyp, which is an abnormal protrusion from the lining of the large bowel wall into the lumen (Conteduca et al., 2013). Special features of polyps include that they take more than a decade to progress to carcinoma, and that they are

visible under endoscopy and can be readily removed (Carethers and Jung, 2015; Keum and Giovannucci, 2019). These characteristics of polyps provide a unique window of opportunity for CRC prevention via screening and precursor removal.

According to the current understanding, there are two major distinct carcinogenesis pathways explaining the formation of sporadic CRCs: adenoma-carcinoma pathway and serrated neoplasia pathway (Dekker et al., 2019; Keum and Giovannucci, 2019). Adenoma pathway is deemed account for 70-85% of CRC development, while serrated neoplasia pathway the other 15-30% (see **Figure 1**) (Bettington et al., 2013; Ijspeert et al., 2015). Adenoma-carcinoma pathway is so far better understood, in which CRCs develop through conventional adenomas with abnormalities caused by chromosomal instability (CIN) (Bakhoum et al., 2014; East et al., 2015). On the other hand, the evidence for serrated neoplasm pathway only substantiated in the recent decade (Crockett and Nagtegaal, 2019), and it is largely manifested by CpG island methylator phenotype (CIMP) with more than half having microsatellite instability (Weisenberger et al., 2006; East et al., 2015). CIMP is a form of epigenetic modification due to hypermethylation at the CpG islands, which is mostly caused by gene mutation, like in *BRAF* or in *KRAS*, and leads to inactivation of the promotor regions of tumor suppressor genes (Weisenberger et al., 2006; Keum and Giovannucci, 2019).

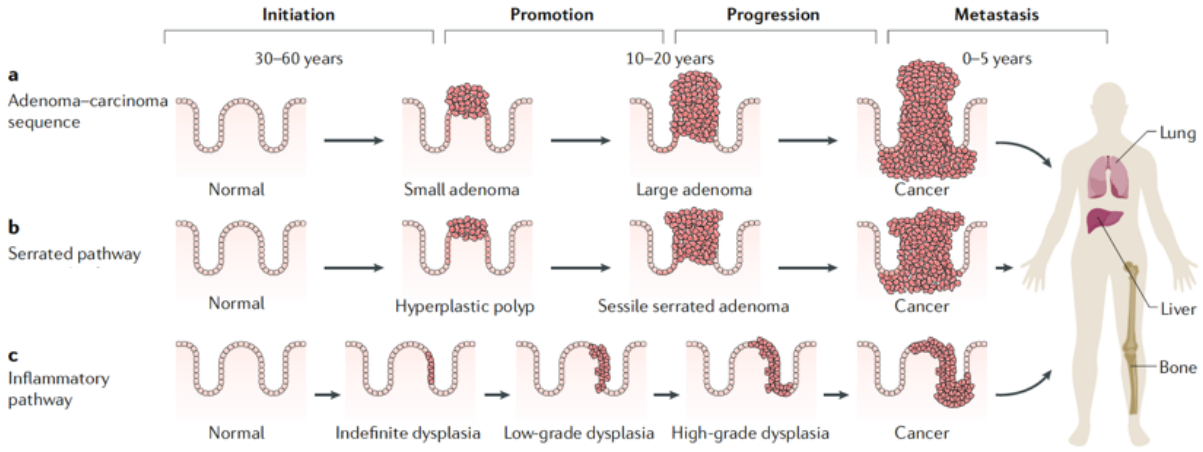


Figure 1. Pathways of CRC pathogenesis.
A | Adenoma-carcinoma pathway, accounting for 70-85% of sporadic CRC. **B** | Serrated neoplasia pathway, accounting for 15-30% of sporadic CRC. **C** | Inflammatory pathway, not a focus of present thesis. Reproduced with permission (Keum and Giovannucci, 2019).

Clinically, adenomas can be categorized into advanced adenoma if adenomas are either greater than 10mm in diameter, with villous components or high-grade dysplasia (Keum and Giovannucci,

2019). Advanced adenomas harbour, with or without multiplicity (>3 lesions), 30-50% of higher risk transitioning into CRCs compared with only 1% for non-advanced adenoma (Conteduca et al., 2013), and the transition risk increases with age at detection (Brenner et al., 2007). As for serrated lesions, according to the 5th edition of WHO classification (Nagtegaal et al., 2020), lesions can be categorized into hyperplastic polyp (HP), sessile serrated lesion (SSL) and traditional serrated adenomas (TSA). HPs are the most prevalent (60-75%) type of serrated lesions (Ijspeert et al., 2015). Majority of HPs are located in distal colon and not considered having malignant potential (Keum and Giovannucci, 2019). However, recent literature revealed that they might still progress to CRC through developing into SSL or TSA (Crockett and Nagtegaal, 2019). SSL was used to be called sessile serrated adenoma or sessile serrated polyp but became the recommended term by the WHO classification in 2019. SSLs account for around 20-35% of serrated lesions, and they are located mostly in proximal colon (Ijspeert et al., 2015). Under 10% of SSLs contain dysplasia and present higher risk to progress into CRC (Crockett and Nagtegaal, 2019). TSA is a rare subtype (<1%) and is often located in distal colon with a polypoid morphology (Ijspeert et al., 2015). Having an SSL or TSA might increase the risk of CRC by 2.5 and 1.8 folds, respectively, compared to those without history of polyps (Erichsen et al., 2016).

Other than the WHO classification, which requires additional histopathology features to definitively differentiate HPs and SSLs, Ijspeert et al. (2017) proposed a classification, clinically relevant serrated polyps (crSPs), which spares the need for an accurate histopathology differentiation and potentially allows endoscopists to better detect and remove serrated lesions with malignancy potential. Clinically relevant SPs are defined as serrated lesions ≥ 10 mm or serrated lesions > 5 mm located proximally to the splenic flexure (Ijspeert et al., 2017; Schramm et al., 2018).

1.1.3 Brief overview of disease management

Management of CRC is only briefly reviewed below. Diagnosis via colonoscopy is the method of choice. The most recent CRC staging by the American Joint Committee on Cancer (AJCC) system (American Cancer Society, 2022; Weiser, 2018) is described in **Table 1.**

Some early stage cancers (T1 cancers) can also be completely resected during colonoscopy, but surgery remains the cornerstone for any treatment with curative intent (Dekker et al., 2019). As for rectal cancers, perioperative chemoradiotherapy has been widely used, shown to downsize most of the tumors and reach complete response in 15-20% of the patients (Dekker et al., 2019).

Regarding systemic treatment, fluoropyrimidine-based chemotherapy can improve survival in resected stage III cancers as an adjuvant chemotherapy, and the addition of oxaliplatin has

become a standard treatment based on some landmark studies including MOSAIC (Yothers et al., 2011). For metastatic diseases, the systemic treatment is much more complex, and patients usually require several lines of treatment. In general, Fluoropyrimidines, oxaliplatin, and irinotecan form the chemotherapy backbone, with some biologics (e.g., anti-VEGF or anti-EGFR antibody, checkpoint inhibitor like PD-L1 inhibitor, etc.) added on top, tailored to tumor-specific and patient-specific characteristics (Dekker et al., 2019).

Table 1. Colorectal cancer staging by the American Joint Committee on Cancer (AJCC) system

AJCC Stage	Stage grouping	Stage description
0	**Tis N0 M0**	The cancer is in its earliest stage. This stage is also known as carcinoma in situ or intramucosal carcinoma (Tis). It has not grown beyond the inner layer (mucosa) of the colon or rectum.
I	**T1 or T2 N0 M0**	The cancer has grown through the muscularis mucosa into the submucosa (T1), and it may also have grown into the muscularis propria (T2). It has not spread to nearby lymph nodes (N0) or to distant sites (M0).
IIA	**T3 N0 M0**	The cancer has grown into the outermost layers of the colon or rectum but has not gone through them (T3). It has not reached nearby organs. It has not spread to nearby lymph nodes (N0) or to distant sites (M0).
IIB	**T4a N0 M0**	The cancer has grown through the wall of the colon or rectum but has not grown into other nearby tissues or organs (T4a). It has not yet spread to nearby lymph nodes (N0) or to distant sites (M0).
IIC	**T4b N0 M0**	The cancer has grown through the wall of the colon or rectum and is attached to or has grown into other nearby tissues or organs (T4b). It has not yet spread to nearby lymph nodes (N0) or to distant sites (M0).
IIIA	**T1 or T2 N1/N1c M0**	The cancer has grown through the mucosa into the submucosa (T1), and it may also have grown into the muscularis propria (T2). It has spread to 1 to 3 nearby lymph nodes (N1) or into areas of fat near the lymph nodes but not the nodes themselves (N1c). It has not spread to distant sites (M0).
	T1 N2a M0	The cancer has grown through the mucosa into the submucosa (T1). It has spread to 4 to 6 nearby lymph nodes (N2a). It has not spread to distant sites (M0).
IIIB	**T3 or T4a N1/N1c M0**	The cancer has grown into the outermost layers of the colon or rectum (T3) or through the visceral peritoneum (T4a) but has not reached nearby organs. It has spread to 1 to 3 nearby lymph nodes (N1a or N1b) or into areas of fat near the lymph nodes but not the nodes themselves (N1c). It has not spread to distant sites (M0).
	T2 or T3 N2a M0	The cancer has grown into the muscularis propria (T2) or into the outermost layers of the colon or rectum (T3). It has spread to 4 to 6 nearby lymph nodes (N2a). It has not spread to distant sites (M0).
	T1 or T2 N2b M0	The cancer has grown through the mucosa into the submucosa (T1), and it might also have grown into the muscularis propria (T2). It has spread to 7 or more nearby lymph nodes (N2b). It has not spread to distant sites (M0).
IIIC	**T4a N2a M0**	The cancer has grown through the wall of the colon or rectum (including the visceral peritoneum) but has not reached nearby organs (T4a). It has spread to 4 to 6 nearby lymph nodes (N2a). It has not spread to distant sites (M0).
	T3 or T4a N2b M0	The cancer has grown into the outermost layers of the colon or rectum (T3) or through the visceral peritoneum (T4a) but has not reached nearby organs. It has spread to 7 or more nearby lymph nodes (N2b). It has not spread to distant sites (M0).
	T4b N1 or N2 M0	The cancer has grown through the wall of the colon or rectum and is attached to or has grown into other nearby tissues or organs (T4b). It has spread to at least one nearby lymph node or into areas of fat near the lymph nodes (N1 or N2). It has not spread to distant sites (M0).
IVA	**Any T Any N M1a**	The cancer may or may not have grown through the wall of the colon or rectum (Any T). It might or might not have spread to nearby lymph nodes. (Any N). It has spread to 1 distant organ (such as the liver or lung) or distant set of lymph nodes, but not to distant parts of the peritoneum (the lining of the abdominal cavity) (M1a).

IVB	**Any T** **Any N** **M1b**	The cancer might or might not have grown through the wall of the colon or rectum (Any T). It might or might not have spread to nearby lymph nodes (Any N). It has spread to more than 1 distant organ (such as the liver or lung) or distant set of lymph nodes, but not to distant parts of the peritoneum (the lining of the abdominal cavity) (M1b).
IVC	**Any T** **Any N** **M1c**	The cancer might or might not have grown through the wall of the colon or rectum (Any T). It might or might not have spread to nearby lymph nodes (Any N). It has spread to distant parts of the peritoneum (the lining of the abdominal cavity), and may or may not have spread to distant organs or lymph nodes (M1c).

Reference: (American Cancer Society, 2022; Weiser, 2018)

1.2 Colorectal cancer screening

Given the high disease burden of CRC and the unfavorable survival in the late stage disease (Siegel et al., 2020), prevention has always been advocated an important strategy to tackle the epidemic (Brenner and Chen, 2018). Primary prevention involves improvement of modifiable risk factors, including healthier life style through more physical activities and reduced consumption of red and process meat (Boyle et al., 2012; Vieira et al., 2017), cessation of tobacco smoking (Botteri et al., 2008), and restricted alcohol intake (Bagnardi et al., 2015). Secondary prevention is mainly done through screening, which will be detailed below. Similar factors which are crucial in primary prevention have also been shown to impact CRC survival, therefore, addressing the aforementioned risk factors also lies at the center for tertiary prevention (Brenner and Chen, 2018).

1.2.1 Effectiveness of various screening tools

As described, the slow progression from normal colorectal epithelium to small precursor lesions then to CRC provides a unique window of opportunity for CRC screening to early detect and remove precursor lesions. Increased uptake of screening, especially the accelerated progression in the roll-out of colonoscopy screening since early 2000 in some high-incidence countries (Siegel et al., 2012; Keum and Giovannucci, 2019), has contributed to the decline in CRC incidence in those countries (Sung et al., 2021).

There are two main types of CRC screening tests: stool-based and direct visualization tests. Stool-based tests recommended by the majority of screening guidelines (Wolf et al., 2018; Shaukat et al., 2021; US Preventive Services Task Force, 2021) include high sensitivity guaiac-based fecal occult blood test (gFOBT), fecal immunochemical test (FIT), and multitarget stool DNA test (mtsDNA). Direct visualization tests include colonoscopy, flexible sigmoidoscopy (FS) and computed tomography colonography (CTC). The American College of Gastroenterology (ACG) also recommended consider colon capsule endoscopy (CCE) as an alternative (Shaukat et al., 2021). Among all tests, colonoscopy is the only one-step approach test, while the rest are two-step approach tests, i.e., subsequent colonoscopy is required after a positive result. The two-step

approaches usually require a robust systems-based support and perform better in a population-based (or organized) screening (Senore et al., 2015; Shaukat et al., 2021). The summary of CRC screening modality recommended by selected international guidelines (European Colorectal Cancer Screening Guidelines Working Group et al., 2013; Wolf et al., 2018; GGPO, 2019; Helsingen et al., 2019; Shaukat et al., 2021; US Preventive Services Task Force, 2021) is presented in **Table 2**.

Table 2. Summary of selected major international recommendations for colorectal cancer screening

	Professional society or guidelines (Year)					
	ACG 2021	**USPSTF 2021**	**ACS 2018**	**BMJ Int'l Panel 2019**	**European Guidelines 2013**	**German GGPO 2014**
Starting age	45	45	45	50[4]	50	50
Stopping age	75	75 (85)[2]	75 (85)[3]	79	74	-
Screening tests (recommended interval in years)						
Colonoscopy	10[1]	10	10	15[4]	10-20	10
FS	5-10	5	5	15[4]	10-20	5[6]
HSgFOBT	-	1	1	-	1-2	1
FIT	1[1]	1	1	1-2[4]	1-2[5]	1[6]
CTC	5	5	5	-	NR	NR
mtsDNA	3	1-3	3	-	NR	NR
CCE	5	-	NR	-	NR	NR

Note: ACG = American; ACS = American Cancer Society; BMJ Int'l Panel = British Medical Journal International Panel; CCE = colon capsule endoscopy; CTC = computer tomographic colonography; FIT = fecal immunochemical test; FS = flexible sigmoidoscopy; GGPO = German Guideline Program in Oncology; HSgFOBT = high sensitivity guaiac-based fecal occult blood test; mtsDNA = multitarget stool DNA test; NR = not recommended; USPSTF = US Preventive Services Task Force; - = not considered in the recommendations.

(1) Recommended as primary CDC screening modalities.

(2) Individuals aged 76-85 years should be selectively recommended by clinicians.

(3) Decision for CRC screening should be individualized by clinicians for individuals aged 76-85 years.

(4) CRC screening is only recommended for individuals with an estimated 15-year risk of CRC >3%. Colonoscopy and sigmoidoscopy are recommended only once, while annual or biennial FIT limited to 15 years.

(5) FIT was deemed superior to gFOBT.

(6) FS should be accompanied with annual gFOBT. If FIT has a specificity >90% and high sensitivity, it can be used as an alternative to gFOBT.

References: (Bénard et al., 2018; Lin et al., 2021)

1.2.1.1 Stool-based tests

Four large-scale randomized control trials (RCTs) have confirmed the efficacy of annual or biennial gFOBT followed by a colonoscopy if positive results (Jodal et al., 2019), where 16-32% of CRC mortality reduction was observed after more than 15 years of follow-up (Kronborg et al., 2004; Lindholm et al., 2008; Scholefield et al., 2012; Shaukat et al., 2013). The information and results of the four RCTs are summarized in **Table 3**.

Table 3. Summary of four large gFOBT randomized control trials

	Study			
	Mandel	**Scholefield**	**Kronbrog**	**Kewenter**
Study location	Minnesota, USA	Nottingham, UK	Funen, Denmark	Gothenburg, Sweden
Design	Voluntary	Population-based	Population-based	Population-based
Screening interval	Annual[1] & biennial[2]	Biennial[3]	Biennial[4]	Biennial[5]
Study period	1975-1992	1981-1995	1985-2002	1982-1995
Age	50-80	45-74	45-75	60-64
Screening (N)	A: 15,570; B: 15,587	76,253	30,966	34,164
Control (N)	15,394	76,384	30,967	34,144
Follow-up years	30	19.5	17	15.5
HR, CRC incidence (mean, 95% CI)[6]	-	0.94 (0.85-1.05)	1.02 (0.93-1.12)	0.96 (0.86-1.06)
HR, CRC mortality (mean, 95% CI)[6]	A: 0.68 (0.56-0.82) B: 0.78 (0.65-0.93)	0.82 (0.7-0.95)	0.84 (0.73-0.96)	0.84 (0.71-0.99)

Note: A = annual; B = biennial; CI = confidence interval; NS = not statistically significant; - = not reported

(1) 11 screening rounds over 15 years; (2) 3 to 6 screening rounds over 15 years; (3) 3 to 5 screening rounds; (4) 9 screening rounds; (5) 2-3 screening rounds, interval up to 10 years; (6) Ratio between intervention group and control group.

References: (Jodal et al., 2019; Kronborg et al., 2004; Lindholm et al., 2008; Scholefield et al., 2012; Shaukat et al., 2013)

After these trials took place, newer immunological fecal occult blood tests, iFOBT of FIT, became available. FITs have similar sensitivity for CRC or advanced lesion detection as seen in newer high-sensitivity gFOBT (Whitlock et al., 2008; Imperiale et al., 2019). Moreover, FIT has several advantages over gFOBT, including the ease of sample collection, requiring only one stool sample, and no dietary or medication restrictions, as opposed to three stool samples and avoidance of diet containing animal blood prior to the test for gFOBT (Shaukat et al., 2021). Therefore, despite no RCT data so far to confirm FIT efficacy in CRC incidence and mortality reduction, FITs are now widely recommended, or already adopted, for population-wide organized screening programs (Schreuders et al., 2015; Lin et al., 2020), and the convenience of FIT has led to increased uptake (Senore et al., 2015). FIT is mostly recommended for annual or biennial screening (Schreuders et al., 2015).

The mtsDNA test is approved by the US Food and Drug Administration and is recommended for screening every three years by major US screening guidelines (Wolf et al., 2018; Shaukat et al., 2021; US Preventive Services Task Force, 2021). The stool test consists of a FIT for hemoglobin plus an assay for mutated *KRAS*, methylated *BMP3*, methylated *NDRG4,* and *β-actin* (Imperiale et al., 2014). Stool DNA test was shown to have approximately 20 percentage points higher sensitivity to detect cancer and adenomas than FIT, and it also showed a significantly increased sensitivity for SSLs (42% vs 5% in FIT). However, the specificity was lower for CRC and advanced lesions (87% vs 95% in FIT) (Imperiale et al., 2014).

1.2.1.2 Direct visualization tests

Colonoscopy is the gold standard for colorectal lesion diagnosis by directly visualizing the whole colon, and it offers advantages of early-stage cancer or polyp detection and removal in one step. This confers a long-term protection in terms of CRC incidence and mortality reduction by 69% and 68%, respectively, as reported in a pooled analysis from a systematic review of six observational studies (Brenner et al., 2014b). Although there is no RCT evidence to date, several RCTs are ongoing and results are expected in the coming decade (see **Table 4a**) (Kaminski et al., 2012; Quintero et al., 2012; Dominitz et al., 2017). Colonoscopy is also used as a single screening modality with an interval of 10 years in several countries (Schreuders et al., 2015).

FS also allows directly visualization, but only for the left side of the colon, and it requires referral to colonoscopy if lesions are found during FS screening. Four RCTs evaluating one-time or 3-5 yearly FS screening have provided up to 17-years of follow-up data (see **Table 4b**) (Atkin et al., 2017; Holme et al., 2017), and they reported 20-30% of CRC incidence and mortality reduction. However, not many countries adopt FS in the screening program as it requires similar infrastructure as colonoscopy, but it does not inspect the entire colon which require further colonoscopy examination for positive findings (Schreuders et al., 2015; Shaukat et al., 2021).

Table 4. Summary of clinical studies evaluating direct visualization tests as screening tool
Table 4a. Three ongoing colonoscopy randomized control trials

	Study		
	NordICC	**COLOPREV4**	**CONFIRM**
Study location	NL, NO, SE, PO	Spain	US
Design	Vs. no screening	Vs. biennial FIT	Vs. annual FIT
Screening interval	Single colonoscopy	Single colonoscopy	Single colonoscopy
Enroll period	2009-2014	2008-2021	2012-2018
Age	55-64	50-69	50-75
Intervention (N)	31,420	26,703	25,000
Comparison (N)	62,974	26,599	25,000
Anticipated follow-up years	15	10	10

Table 4b. Four large flexible sigmoidoscopy randomized control trials

	Study			
	UKFSS	**PLCO**	**SCORE**	**NORCCAP**
Study location	UK	US	Italy	Norway
Design	Voluntary	Voluntary	Voluntary	Population-based
Screening interval	Once	3-5 years, twice	Once	Once
Study period	1994-1999	1993-2001	1995-1999	1999-2001
Age	55-64	55-74	55-64	50-64
Screening (N)	57,254	77,443	17,148	20,780
Control (N)	113,178	77,444	17,144	79,430
Follow-up years	17.1	16.8	11.4	14.8
HR, CRC incidence (mean, 95% CI)[1]	0.65 (0.59-0.71)	0.79 (0.72-0.85)	0.82 (0.69-0.96)	0.8 (0.7-0.92)
HR, CRC mortality (mean, 95% CI)[1]	0.59 (0.49-0.7)	0.74 (0.63-0.87)	0.78 (0.56-1.08)	0.73 (0.56-0.94)

Note: CI = confidence interval; FIT = fecal immunochemical test; NL = Netherlands; NO = Norway; SE = Sweden; PO = Poland

(1) Ratio between intervention group and control group.

References: (Dominitz et al., 2017; Jodal et al., 2019; Kaminski et al., 2012; Quintero et al., 2012) (Atkin et al., 2017; Holme et al., 2017; Jodal et al., 2019)

CTC was shown to have 68-98% sensitivity and 80-93% specificity for lesions ≥6mm (Pickhardt et al., 2011). However, some concerns for CTC to detect right-sided and flat lesions exist, and the sensitivity for SSLs was only 0.8%, which was significantly lower when compared with colonoscopy (3.1%) in the same study (IJspeert et al., 2016).

CCE achieved some improvement in screening in the recent years due to both software and hardware enhancement, and it demonstrated a sensitivity at 81% and specificity at 93% for lesions ≥6mm, while the completion rate was 79% in a prospective study (Rex et al., 2015). ACG recommended both CTC and CCE as alternatives for those who are unable to undergo colonoscopy or FIT screening (Shaukat et al., 2021).

1.3 Cost-effectiveness of colorectal cancer screening

1.3.1 Cost-effectiveness analysis

Resources are not unlimited, and the same holds true within the healthcare systems. Therefore, healthcare policy makers often need to make decisions to fund healthcare services under budgetary constraints. One way to guide the decision is based on how can we allocate available resources to maximize health of the population. This requires the comparison of costs and effectiveness of outcomes among alternative healthcare interventions (e.g., for detection, prevention, or treatment). This type of study for such comparison is called cost-effectiveness analysis (CEA) (Drummond et al., 2015). Between alternatives, an incremental cost-effectiveness ratio (ICER) can be derived by taking the difference in effectiveness over the difference in costs. When an ICER is under certain willingness-to-pay (WTP) threshold, an alternative intervention might be deemed cost-effective (Drummond et al., 2015).

By definition, the effectiveness of outcomes used in CEAs is measured by natural units, e.g., per life saved, per case averted, or per symptom-free day (Drummond et al., 2015). However, we will face difficulty when we need to compare alternatives comprised of different outcomes measures. To overcome this, we often use quality-adjusted life year (QALY) as the common outcome

measures, which is the product of a quality of life value in one's health state (utility) and the years the person stays in that health state. The type of CEAs which uses QALY as the outcome measure is also referred to as cost-utility analysis (CUA) (Drummond et al., 2015).

1.3.2 Cost-effectiveness of various screening strategies

To better understand the state of the art of the cost-effectiveness of CRC screening, a systematic review was jointly conducted with Dr. Tao Ran as part of a broader overarching project related to this present doctoral thesis (Ran et al., 2019). The review extended the findings from a previous work (Lansdorp-Vogelaar et al., 2011) to systematically search and analyze cost-effectiveness analyses (CEAs) of CRC screening published between 1st of January 2010 and 31st of December 2017. The fundamental question was to review the cost-effectiveness of six established screening strategies, which included annul and biennial gFBOT, annul and biennial FIT, 5-yearly FS, and 10-yearly colonoscopy, in comparison to no-screening. All six strategies had incremental cost-effectiveness ratios (ICER) under USD 50,000 per life-year gained (LYG) compared with no screening in 22 out of the 23 studies reviewed, and majority of the US studies and some of the European studies even deemed these strategies dominant (cost-saving and more effective). When applying the same willingness-to-pay (WTP) threshold to all screening strategies, 10-year colonoscopy appeared to be the most cost-effective strategy among the established ones, especially in the US studies (Ran et al., 2019).

The review also analyzed the cost-effectiveness on three alternative screening modalities: mtsDNA, CTC, and CCE. However, at the time of review, only 5- and 10-yearly CTC were found to have consistent cost-effectiveness results compared with no-screening. There were not abundant studies to verdict the cost-effectiveness of the other two strategies (Ran et al., 2019).

1.3.3 Models and challenges for cost-effectiveness analyses

The discrepancy on the most cost-effective CRC screening strategies observed in the systematic review is most likely linked to the tremendous uncertainty and variations surrounding the modeling study evidence (Ran et al., 2019), given that economic evaluations generating cost-effectiveness evidence heavily depend on modeling. Uncertainty related to modeling studies are usually categorized into four types: structural uncertainty, parameter uncertainty, stochastic uncertainty, and heterogeneity (Briggs et al., 2012). Structural uncertainty derives from the inherent model structure assumptions. Parameter uncertainty arises from the estimation of individual input parameters. Stochastic uncertainty and heterogeneity both refer to the variability related to simulated individuals or patients, but they distinct from each other that the former

describes the random variability happens within the outcomes between identical individuals, while the latter denotes the variability found between different individuals due to personal characteristics (Briggs et al., 2012). The structural uncertainty among different models, specifically the different structures and assumptions for CRC natural history modeling, may explain the large part of the disagreements on cost-effectiveness of CRC screening. On top of that, uncertainty surrounding the parameters input to the natural history model also contributes to a fraction of the variations.

1.3.3.1 Markov vs microsimulation models

Concerning the structure variations of natural history model, one of the non-negligible topics is the choice of model type. Among the CRC screening models, a great majority of the models are state-transition models (STMs). Most of the STMs are modeled at the cohort level, also commonly known as "Markov models," while the other small proportion of models take an individual-level modeling approach, so called "microsimulation models." (Siebert et al., 2012; Ran et al., 2019) (Hereafter, cohort STM and Markov model, as well as individual-level STM and microsimulation model will be used interchangeably). It is worth noting that three models (MISCAN-Colon, SimCRC and CRC-SPIN 1.0/CRC-SPIN 2.0), under the scheme of the US National Cancer Institute, Cancer Intervention and Surveillance Modelling Network (CISNET), so far dominate the CRC screening modeling realm (Kuntz et al., 2011; van Ballegooijen et al., 2011), and they have been extensively used to inform several important CRC screening recommendations (Wolf et al., 2018; Shaukat et al., 2021; US Preventive Services Task Force, 2021).

Both cohort or individual-level STMs simulate a single closed cohort without interpersonal interaction and allow the transition to happen at a specific point of time. STMs can capture many features, including health state changes and disease risk over time, which gives them the flexibility to model complex research questions required to reflect time (Siebert et al., 2012). The advantages of cohort STMs are that they are relatively simple to develop, debug, communicate, and implement on commonly used software, e.g., Microsoft Excel, if not too complex. However, cohort STMs also bear a critical disadvantage: the transition probabilities into next states do not depend on history (i.e., occupation of past states or the time spent on the current state). This memory-less characteristic is called "Markovian property", and it poses a limitation for modeling the clinical questions that require the consideration of individual history along the disease pathway. The problem can indeed be solved by creating more states to store the history, yet this might lead to an undesired large number of states (which is called "state explosion.") (Siebert et al., 2012)

On the other hand, individual-level STMs focus on individuals along the disease pathway, keep track of their history, and, therefore, reflect individual heterogeneity. Hence, this type of model can be free of the memory-less issue faced by cohort STMs (Briggs et al., 2016). Although individual-level STMs can be more flexibly used for simulation, given their complexity, they usually require more input parameters and computational power and are more difficult to debug (Siebert et al., 2012).

1.3.3.2 Challenges of disease natural history modeling

Other than the choice of model type, variations in the assumptions for the CRC natural history introduced additional uncertainty as well as challenges. As discussed in Chapter 1.1.2, the understanding of CRC carcinogenesis pathways only advanced beyond the conventional adenoma-carcinoma sequence in the recent two decades, and only until recent years has the consensus on the serrated neoplasia pathway been reached (Crockett and Nagtegaal, 2019). This led to the situation that, previous modeling works mostly concentrated on simulating CRC development through adenoma-carcinoma pathway. For lesions potentially arising from serrated pathway and unable to be explained by the conventional sequence, a few previous models attempted to capture those by adding extra natural history assumptions, e.g., "de novo" cancers without pre-existing lesions or fast-growing adenomas (Silva-Illanes and Espinoza, 2018).

With different assumptions on the CRC natural history pathways, input parameters required also vary consequently, which further amplify the parameter uncertainty among models and add to the conundrum of performing economic evaluations for CRC screening. Among the challenges, CRC natural history parameters that are not directly observable (so-called *deep parameters*) but critical to modeling the progression of precancerous and pre-clinical (or asymptomatic) cancer lesions within the disease natural history lie at the center (Silva-Illanes and Espinoza, 2018). The most crucial deep parameters in CRC natural history are believed to be precancerous lesion dwell time (hereafter, dwell time) and cancer sojourn time (hereafter, sojourn time) (Silva-Illanes and Espinoza, 2018). Dwell time represents the time from the precancerous lesion occurrence until transforming into a pre-clinical cancer, while sojourn time indicates the time from the onset of a pre-clinical cancer to the transition into a clinical (or symptomatic) cancer and being detected (Kuntz et al., 2011; van Ballegooijen et al., 2011). Together, they provide the lead time for any lesions to be detected early before surfacing as clinical cancers.

Both dwell and sojourn times are random variables with unknown distribution in the population (Kuntz et al., 2011; Silva-Illanes and Espinoza, 2018). Given the enormous challenge to estimate the two parameters, most CRC models turned away from using the two times directly. Instead, CRC STMs seek an indirect way by using transitional probabilities between adenoma or cancer

states to simulate the disease progression, and the dwell time and sojourn time in turn become the outputs of the overall time in which individuals stayed in those health states (Kuntz et al., 2011; Silva-Illanes and Espinoza, 2018). Albeit modelers find a way to circumvent the unobservable dwell and sojourn times, transition probabilities themselves are also not directly observable. Nevertheless, it is relatively feasible to estimate the transition probabilities via calibration (Silva-Illanes and Espinoza, 2018) or statistical estimation using observable epidemiological data (Brenner et al., 2011, 2013, 2014a). As a result, the challenge of parameter uncertainty surrounding CRC natural history modeling is slightly mitigated but remains a non-negligible one.

1.4 Colorectal cancer screening in Germany

According to the German S3 Guidelines for Colorectal Cancer published by the German Guideline Program in Oncology (GGPO) (2019), which outlines the recommendations across prevention, screening, diagnosis and management of CRC, and the latest recommendations for screening were made in 2013. CRC screening is recommended to average-risk individuals aged greater than 50 years old, and the GGPO working group does not recommends upper age limit and states that screening in older age should be individually considered depending on their "biological age." (GGPO, 2019) Colonoscopy is recommended as "gold standard" and to be repeated every 10 years. If individuals refuse, 5-yearly FS with annual gFOBT or annual gFOBT is recommended. Additionally, if FIT has a specificity >90% and high sensitivity, it can be used as an alternative to gFOBT (GGPO, 2019). The comparison of German recommendations with those from other professional societies, see **Table 2**.

These recommendations, however, differ from the CRC screening program currently implemented and reimbursed by the German health insurance, in which screening colonoscopy and FIT are included (Gemeinsamer Bundesausschuss (G-BA), 2018). At the time when the GGPO working group made the recommendation for gFOBT in 2013, they deemed that the sensitivity and specificity of FIT were not stable enough to be recommended for screening purposes (GGPO, 2019), which has actually been improved in the past years (Imperiale et al., 2019). Moreover, the GGPO already recognized in the guidelines that FS is not covered by the German health insurance's catalogue of benefit and cannot be billed, let alone that there is no quality assurance measures for FS in place (GGPO, 2019). Thus, these considerations reflect in the currently implemented and reimbursed screening modalities in Germany.

In practice, CRC screening with gFOBT in Germany was already offered to men and women older than 45 years old starting in 1977, which was more than 20 years before the publications of the protective effects from gFOBT RCTs. However, the information regarding the gFOBT screening back then was not well documented (Haug, 2018).

In October 2002, a nationwide screening colonoscopy program was introduced in Germany, in which gFOBT and screening colonoscopies are covered by statutory health insurance (SHI), accounting for around 90% of overall insured population in Germany. Anyone aged older than 50 years old is entitled to have annual gFOBT screening followed by a screening colonoscopy when they turn age 55 years old. If the first colonoscopy result is negative and participants receive the first screening colonoscopy before age of 65, they will be offered a second screening colonoscopy with a 10-year interval (Pox et al., 2012; Haug, 2018).

After 10 years (2003-2012) into the new screening offer, Brenner et al. estimated that the CRC age-standardized incidence decreased by 14% in both men and women, which was an inversed trend comparing with the incidence prior to 2002 (Brenner et al., 2016). The CRC mortality also decreased by 21% in men and 27% in women during the same period, continuing the trend prior to the introduction of screening colonoscopy (Brenner et al., 2016).

Despite the program showing its effectiveness in reducing CRC incidence and mortality, the participation rates remain low. According to the statistics from the Federal Health Monitoring (Gesundheitsberichterstattung des Bundes, GBE) (2021), the uptake rate for annual FOBT in 2017 was 7% for men and 25% for women, while the uptake rate for biennial FOBT in 2016-2017 was 16% for men and 24% for women. A similarly low rate was observed in screening colonoscopy: cumulative participation rate between 2008 and 2017 was around 18% for both sexes. The higher annual FOBT uptake rate observed in women aged 50-54 years old was likely due to the encouragement from gynecologists (Haug, 2018). Overall, the low participation rates might be attributed to the fact that the German CRC screening program was essentially an opportunistic screening program without centralized coordination to identify and invite eligible participants (Haug, 2018; Senore et al., 2019).

In recent years, there were changes to improve the German CRC screening program. In 2016, Federal Joint Committee (Gemeinsamer Bundesausschuss, G-BA) reached the decision to replace gFOBT with FIT for screening starting in April 2017 (G-BA, 2016). Besides, with the passing of the new laws for cancer screening and registry in 2013, the CRC screening program was set to be implemented in an organized manner, whereby centralized invitation systems should be established and systematic evaluation and improvement of the screening program should be executed (Haug, 2018). In April 2019, the organized screening program officially kicked off, with

personal invitation letters sent by SHI companies when eligible individuals reach age 50, 55, 60, and 65 years old. Alongside the change, a sex-differentiated strategies was introduced: the age entitled to receive the first screening colonoscopy decreased to 50 years old for men, while the screening offer to women remained unchanged (G-BA, 2018). The current German CRC screening program is summarized in **Table 5**.

With the transition from the usage of gFOBT to FIT and from opportunistic to organized program, both mechanisms bear the objective to improve the screening participation rates in Germany (Hol et al., 2010; Levin et al., 2018), which might further translate into even more significant CRC incidence and mortality reduction.

Table 5. Current colorectal cancer screening approach in Germany

	Screening test	
	Colonoscopy	**FIT[1]**
Men	Starting eligible age: 50 years old Entitled to 2 screening colonoscopies if the first was done before age of 65 years	If no screening colonoscopy is used: ▪ Age 50-54 years: annually ▪ Age ≥ 55 years: biennially
Women	Starting eligible age: 55 years old Entitled to 2 screening colonoscopies if the first was done before age of 65 years	Age 50-54 years: annually If no screening colonoscopy is used: ▪ Age ≥ 55 years: biennially

Note: FIT = fecal immunochemical test

(1) Individuals with a positive FIT result will receive a follow-up diagnostic colonoscopy.

Reference: (Gemeinsamer Bundesausschuss (G-BA), 2018)

1.5 Rationale, aims and objectives of the thesis

1.5.1 Rationales

The primary motivation of the present doctoral thesis is to conduct a CEA of current or alternative CRC screening strategies to understand the cost-effectiveness of the German organized CRC screening program and to inform future CRC screening policy, by using a newly developed individual-level CRC screening model. After reviewing the literature on the topic of CRC screening, as laid out in Chapter 1, several reasons justify the motivation of the present thesis.

To start with, the cost-effectiveness of the current organized screening program is not yet explored. To date, there is only one CEA on this topic conducted in Germany using a Markov model approach. The evaluation was performed during the opportunistic screening program era, where annual gFOBT/FIT at 50-54 years old followed by either biennial gFOBT/FIT at 55-75 years old or 10-yearly colonoscopy at 55 and 65 years old were evaluated. Under the assumption of perfect uptake and adherence, the study found that all above mentioned strategy increased life-years and

cost less, while the pure FIT and FIT-colonoscopy combined strategy dominated other strategies (Ladabaum et al., 2014). Adherence to the screening was found to have a significant impact on the cost-effectiveness results (Ladabaum et al., 2014), and similar findings were also reported in another cohort STM study predicting the effects of screening colonoscopy (Chen et al., 2018). However, the adherence rates examined were based on assumptions instead of being directly informed by real-world data. Therefore, there is a need to examine the cost-effectiveness under the uptake rates achieved by the current organized program to obtain a more realistic picture. Moreover, the cost-effectiveness of the new sex-differentiated strategy which offers men colonoscopy screening from age of 50 has not been evaluated in any of the German modeling studies.

Secondly, several recent studies analyzing the CRC incidence trend in younger adults aged 20 to 49 years reported that the incidence trend in high HDI countries has been increasing since early 2000s (Siegel et al., 2019; Vuik et al., 2019). In Germany, age-standardized CRC incidence among young adults 20-49 years old in 2008-2012 was 7.7 per 100,000, increasing 1.3% annually in the last decade. In the same period, the CRC incidence among people aged 50 years and older in Germany declined by 1% per year (Siegel et al., 2019). The reason for this rising trend in CRC incidence in younger adults is still not fully understood but is likely linked to the change in diet habits and life style (Siegel et al., 2020).

Given the notable upward CRC incidence trend, American Cancer Society set the first example and deemed screening from age of 45 years as a qualified recommendation (Wolf et al., 2018), based on the modeling results from two of the three CISNET models (Peterse et al., 2018). It showed that screening from age of 45 years with 10-yearly colonoscopy may result in 6% more LYGs with 17% more colonoscopies, which was one of the most efficient strategies examined in the models. Other screening modalities starting at age 45 years old also appeared efficient (Peterse et al., 2018). More recently, the updated recommendations from the US Preventive Services Task Force (USPSTF) (2021) and recommendations from the American College of Gastroenterology (Shaukat et al., 2021) also made the same decision citing updated modeling works (Knudsen et al., 2020; Ladabaum et al., 2019). Since the rising CRC incidence trend in younger adults is also noted in Germany, without the clinical evidence on the effect of screening in people aged 45 to 49 years, modeling study tailored to German context is warranted to evaluate the incremental benefits, harms and burden of screening starting at 45 years old in comparison to age of 50.

Thirdly, cost-effective evidence on CRC screening in Germany from an individual-level or microsimulation model is lacking. The two aforementioned German modelling studies were both based on Markov models (Chen et al., 2018; Heisser et al., 2021a; Ladabaum et al., 2014). Given the advantages of individual-level modeling elucidated in Chapter 1.3.2.1, individual-level models

appear to be a more suitable method if individuals' history is to be considered to account for the heterogeneity within the screening population. Consequently, modeling evidence should be explored from an individual-level modeling approach.

Fourthly, as illustrated in Chapter 1.3.2.2, almost all CRC screening models to date, including the CISNET models, assume CRCs develop through the conventional adenoma-carcinoma pathway exclusively (Kuntz et al., 2011; Silva-Illanes and Espinoza, 2018). Only one model from the Netherlands, ASCCA (Greuter et al., 2014), and its Australian variant, policy-1 bowel (Lew et al., 2017), account for the serrated neoplasia pathway, which explains 15-30% of the CRC development (Bettington et al., 2013; Ijspeert et al., 2015). Ignoring the serrated neoplasia pathway might gravitate the structural uncertainty of CRC screening modeling. Thus, it is warranted to investigate the impact on the prediction of CRC screening modeling by incorporating the serrated pathway.

Lastly, to address another challenge laid out in Chapter 1.3.2.2 regarding the deep parameters, one common approach is through calibration, i.e., a process of seeking good fit between the model output and real-world observable data (or called calibration targets) by randomly adjusting input parameters (Rutter et al., 2011; Vanni et al., 2011). Among the existing CRC individual-level models, most models took grid search optimization approaches to calibrate the model parameters, and most commonly, the downhill simplex method (also known as Nelder-Mead algorithm) (Loeve et al., 1999; Knudsen et al., 2012; Greuter et al., 2014; Lew et al., 2017; Prakash et al., 2017). However, the downhill simplex method outputs only one best-fit parameter set, which foregoes capturing the uncertainty around fitted parameters (Vanni et al., 2011).

Bayesian approach can be seen as a remedy to capture the uncertainty around calibrated parameters, and a few CRC screening modelers have trialed this approach: one of the CISNET microsimulation models, CRC-SPIN 1.0 (Rutter et al., 2009), and a cohort STM model from Whyte et al. (Whyte et al., 2011) took the route using Markov Chain Monte Carlo (MCMC) to calibrate their parameters, and more recently, CRC-SPIN 2.0 (Rutter et al., 2018) attempted to use a modified method of approximate Bayesian computation (ABC). Bayesian calibration methods, including MCMC and ABC, allow model calibration to use known information as priors and the observed data as benchmark, through the likelihood function, to gradually converge to the unknown "true" parameter posterior distributions. This way, it naturally addresses the parameter uncertainties in the form of distributions (Rutter et al., 2009). However, MCMC requires a specification of the likelihood in a closed-form function, which is analytically intractable for complex and high-dimensional models (Beaumont, 2019). Against this background, the ABC approach is designed to tackle the issue by circumventing the necessity of specifying the likelihood function. Instead, it directly compares the distance between outputs from proposed

random parameters and the observed data to determine if newly proposed parameters can be a good fit (Beaumont, 2019). This unique property of the ABC approach makes it particularly suitable to approximate the parameters of complex models (Beaumont, 2019), which is also the case for the CRC natural history model. Hence, it would be of high methodological interest to explore the application of ABC calibration approach for a new CRC screening model.

Summing up the last three reasons, there is currently no microsimulation accounting for both CRC tumorigenic pathways and calibrated by Bayesian methods, which drives the motivation to develop a new model specifically for Germany instead of adapting existing models. (Detailed comparison between DECAS and other existing models are described in Chapter 4.1.1.)

1.5.2 Aims and objectives

Summarizing the rationale justifying the motivation mentioned above, there are two major aims of the present doctoral thesis:

1) To construct an individual-level model reflecting new evidence in CRC disease development and to explore new model calibration method for CRC disease modeling,

2) To conduct an up-to-date CEA evaluating the cost-effectiveness of various CRC screening strategies in the current German organized CRC screening program and to inform future CRC screening policy in Germany.

To achieve the aims, the following objectives are planned. The thesis will start with the development of a new individual-level CRC natural history model, a Discrete Event simulation model for the natural history of Colorectal cancer from the Adenoma and Serrated neoplasia pathways (DECAS)**,** which considers both adenoma-carcinoma and serrated neoplasia pathways for CRC tumorigeneses. A detailed explanation of how input parameters for DECAS are calibrated using an ABC approach will be laid out.

Screening component of DECAS will then be superimposed on the well-calibrated model, and external validation against multiple RCTs or cohort studies as well as cross validation with CISNET models will be presented. Lastly, the calibrated DECAS with screening component will be utilized to evaluate the cost-effectiveness of the current German organized CRC screening program, together with the exploration of other alternative screening strategies (including the screening starting from age 45 years old), in the aim of informing future CRC screening policies.

Based on these research objectives, Chapter 2 Material and Methods and Chapter 3 Results will be further divided to sub-chapters to include the methods and results of CRC treatment cost analysis

(Chapter 2.1 and 3.1), DECAS development and calibration (Chapter 2.2 and 3.2), DECAS screening validation (Chapter 2.3 and 3.3), and CEA of German CRC screening program using DECAS (Chapter 2.4 and 3.4). It is then followed by the discussion of the findings in Chapter 4. Finally, in Chapter 5, the doctoral thesis will conclude by rounding up all the results and discussions.

2 MATERIAL AND METHODS

2.1 CRC treatment cost analysis

(Part of Chapter 2.1 has been published (Cheng et al., 2021)*.)*

As a first step to allow the CRC treatment input into the CEA, a cost analysis using a large German SHI administrative database was conducted. Due to the fact that this analysis was a followed-up study of the survival analysis on colon cancer patients by Trautmann et al. (2018), the following cost analysis only focused on the cost data collected from colon cancer patients. Given that literature showed the overall treatment costs of colon cancer and rectal cancer are not significantly different (Bradley et al., 2016; Haug et al., 2014), the results of this cost analysis of colon cancer treatment will be used in the following CEA as the CRC treatment costs. The cost analysis followed the national standards for Good Practice in Secondary Data Analysis (Swart et al., 2015). The costs were discounted with a 3% annual rate (Basu and Ganiats, 2016), and all costs were inflated using the Health Consumer Price Index for Germany (Statistisches Bundesamt, 2021) to 2021 Euro. The analysis was performed using R software, version 4.0.4.

2.1.1 Data source

The data source for the colon cancer treatment cost analysis made used an administrative database of a large SHI company (AOK PLUS) covering approximately 2 million people (approximately 50% of the general population) in the federal state of Saxony, Germany. It contained pseudonymized inpatient (diagnosis, medical procedures, treatment time), outpatient (diagnoses, medical procedures, healthcare providers, and drug prescriptions), and individual information (age, sex, ZIP-code, date of death, date of leaving insurance) at the patient-level in the years 2005-2015 (Trautmann et al., 2018).

2.1.2 Net colon cancer treatment costs

2.1.2.1 Phase of cancer care

The analysis followed previously described methods (Yabroff et al., 2008; Haug et al., 2014; Laudicella et al., 2016) to estimate the costs of cancer care by phases – the initial, continuing and terminal phase. The initial phase comprises the first 12 months post-diagnosis, whereas the terminal phase consists of the last 12 months before death. The continuing phase is the remaining period in between the initial and terminal phases. A “U-shaped” distribution of the costs is

expected (Haug et al., 2014; Laudicella et al., 2016; Yabroff et al., 2008), namely high costs are to be observed in the initial and terminal phases while low costs are expected in the continuing phase. Assuming not every patient survived more than 24 months to be included in all three phases, patients' follow-up periods were allocated sequentially first to the terminal, then to the initial, and lastly to the continuing phase (Laudicella et al., 2016; Yabroff et al., 2008).

2.1.2.2 Patient identification

Following the previous study (Trautmann et al., 2018), 6,186 patients were identified with incident colon cancers using the following criteria: (1) continuously insured by AOK PLUS throughout the study period or until death; (2) with an inpatient diagnosis of malignant neoplasm at the colon/rectosigmoid junction (ICD-10-GM C18/C19) and at least one hospitalized surgical treatment between 1 January 2008 and 31 December 2014; (3) no inpatient diagnosis (C18/C19) within 3 years prior to the diagnosis; (4) no outpatient visits with diagnosis C18/C19 prior to 1 year before the diagnosis. The diagnosis date was defined as the first hospital admission due to the diagnosis C18/C19.

Given that some patients survived beyond the observational period (31 December 2014), making it impossible to assign the terminal phase. Therefore, for the analysis of terminal phase costs, only the patients who died before 31 December 2014 were selected (*N*=1,827).

For the analysis of the initial and continuing phases, to ensure at least 12-month follow-up, only the patients diagnosed between 1 January 2008 and 31 December 2013 were included (*N*=4,438). Among this cohort, some also survived beyond the observational period. Therefore, the use of cost information for this subgroup of survivors (*N*=3,011) was censored until 31 December 2013. This guaranteed the remaining costs for the survivors could be properly allocated to the initial and continuing phases. See **Figure 2** for the patient selection flow chart.

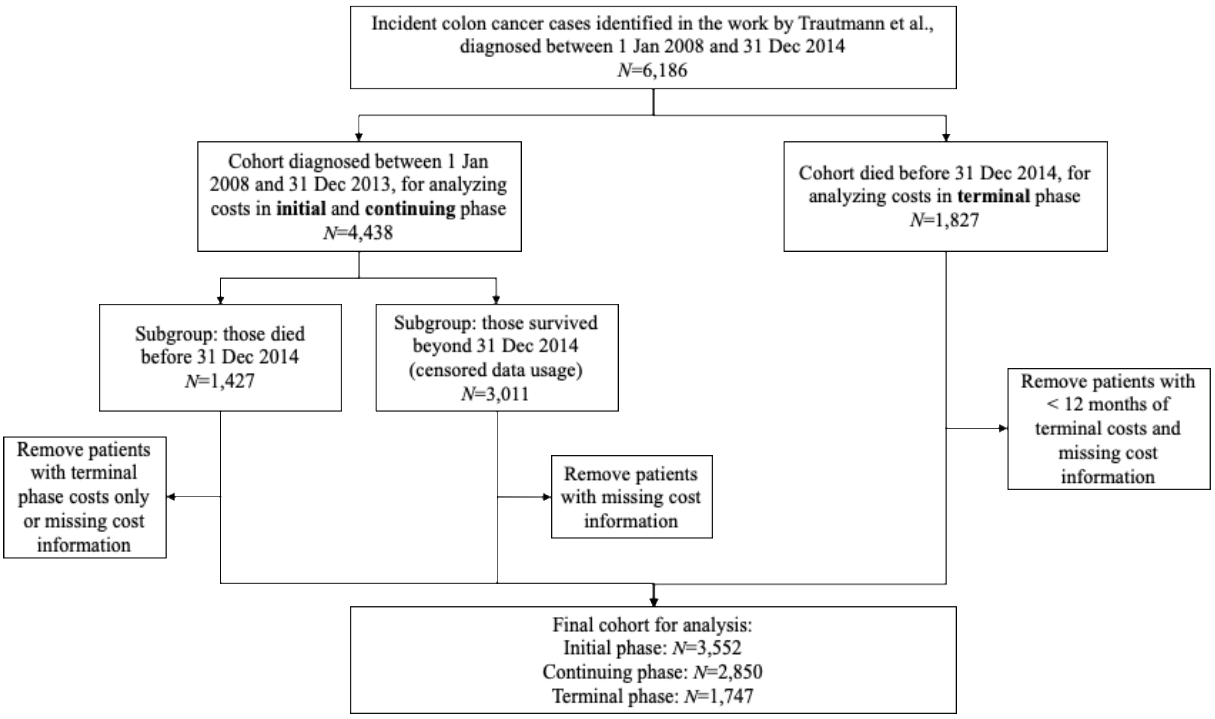


Figure 2. Flow chart of colon cancer cost analysis patient selection.
Adapted from (Cheng et al., 2021)

2.1.2.3 Control group selection

Two separate control groups were established through 1:10 propensity score matching with a standard caliper width of 0.05 (Rubin, 2005) for cases in (1) the initial and continuing phase and (2) the terminal phase, respectively. General matching criteria included birth year (grouped by 10-year interval), sex, and four selected comorbidities (See **Appendix A Supplementary Table 1** for further details). The year and quarter of death and hospital admission within one year before death were additionally used to match terminal phase controls. Controls for cases in initial and continuing phases were matched by year from 2008 to 2013 and combined afterwards. For each year, only controls with at least one hospital admission (of any cause) were included. For each of the matched controls in the initial and continuing phase cohort, a pseudo-diagnosis date which corresponded to the diagnosis date of a patient was assigned (Yabroff et al., 2008).

2.1.2.4 Cost calculation

All the inpatient and outpatient (including consultation, procedures and medications, informed by Uniform Value Scale (*Einheitlicher Bewertungsmaßstab, EBM*) and prescription costs) costs as well as the entire follow-up months for each subject by the allocated phase were summed up. It was followed by the calculation of the mean monthly costs by phase for the patient and control

groups, respectively. To prevent skewness from extreme outliers (especially in the control group), before calculating the means, winsorization of the extreme cost estimates for each phase in both patient and control groups was performed – the procedure replaced the extreme figures by the top and bottom 5% values within the cohort (Haug et al., 2014). The average monthly net colon cancer treatment costs for each phase were the difference of the mean monthly phased costs between the patient and control groups, and they were then annualized. Subgroup analysis by cancer severity was also performed following the definition of cancer stage proxies (see **Appendix B Supplementary Table 1**). All net treatment costs were calculated with 95% confidence intervals (CI).

2.2 DECAS – a discrete event simulation model for CRC natural history

(Part of Chapter 2.2 has been submitted and currently under review (Cheng et al., 2022)*)*

As the core of this doctoral thesis, an individual-level model simulating the CRC natural history was developed. Different from most of the microsimulation models for CRC screening, a discrete event simulation (DES) approach was decided as the fundamental modeling technique to build up the CRC model.

2.2.1 Discrete event simulation

DES was originally developed in 1960s for the purpose of industrial engineering and operations research to analyze and improve industrial and business processes. It is a flexible modeling method to study complex behaviors within individuals, populations or environments, or even the interactions between them (Karnon et al., 2012). As given away in its name, DES simulates discrete (or mutually exclusive) events at discrete time intervals, which gives versatile abilities to be applied to analyze a wide range of problems. In the last four decades, the researchers in biology-related and healthcare research fields adopted this modeling approach to simulate the reactions within cells, conduct trial design or evaluate health policies (Karnon et al., 2012).

Standard DES models consist of core concepts including entities, attribute, events, resources, queues and time (Karnon et al., 2012). Entities are the main objects moves through DES model, which possess attributes, experience events, and utilize resources over time. Attributes describe the specific features for the entities, which can include age, sex, health status and history, personal risk profile, quality of life, consumed resources and accumulated costs. Events are something happens to entities, e.g., occurrence of a disease or receiving treatment, and they drive the

transition of DES models and can compete with each other. Resources are usually the services provided to entities. When resources are limited, entities will enter queues to wait for available resources to be released. Finally, the fundamental part of DES is time, which is kept tract of during the simulation (Karnon et al., 2012).

When it comes to cancer screening modeling used in economic evaluation, individual-level STMs (microsimulation models) are more commonly used. There are some similarities between individual-level STMs and DES models, for example, they both model the research questions at individual-level and serve as useful tools to capture heterogeneity among individuals. However, the event-based characteristics gives DES models advantages over microsimulation models with fixed-cycle transitions. Specifically, the events in DES can happen any time, comparing to the fixed time interval set for microsimulation models (Karnon et al., 2012). This time-to-event approach can provide an efficient solution for the problems needed to be followed up for longer term, as one can avoid checking periodically when none of the events happen in most of the cycles (Caro and Möller, 2016). Moreover, the events can compete with each other, which is very difficult for STMs to model. Also, if a resource-constraint environment is to be modeled, DES is inherently designed to model such a problem (Karnon et al., 2012).

Given the above-mentioned advantages, DECAS was developed, which simulates the projection of the natural history of colorectal cancer at individual-level in a birth cohort from the age of 20 until 90 or death. The events (lesion initiation, progression, and death) compete with each other only if the preceding states are occupied, so that it assures the lesions progress through the randomly assigned pathway (Wilkinson, 2018). Good practices for modeling, such as recommendations from ISPOR-SMDM Modeling Good Research Practice Task Force, were followed to build DECAS (Eddy et al., 2012; Karnon et al., 2012).

Slightly different to the main interest of most DES models on the occurrence of events, events were treated as driving force to make the state transition happen in a Markov process (illustrated in **Figure 3** below). To simulate the evolution of the system in continuous time, an algorithm known as first reaction method was adopted. This method was introduced by Gillespie (1976) in the field of chemical system simulation, taking advantage of a mathematical system representation comparable to that of a stochastic Petri net (Wilkinson, 2018). The pseudocodes of DECAS are detailed in **Algorithm 1**.

Algorithm 1. DECAS Pseudocodes

Set the state vector M and the transition matrix S, which defines the change of the states counts after each event
Set the rates of E events (other than the initiation of pre-cancer lesions) $R \in \{r_1, \dots, r_E\}$ and life tables
Generate N individual characters, including personal risk profile for both type of pre-cancer lesions and sex
For $i = 1$ to N **do**
  Set time tracker t_{cur} = 0
  Generate individual i's time to death from non-cancer cause t_{nonca}
  Repeat
    Multiply M and R to limit the transitions to the events where the previous state is occupied
    If number of total pre-cancer lesion is below 20
      Generate time to initiation of pre-cancer lesions $T_{init} \in \{t_a, t_s\}$
    Generate time to E events $T \in \{t_1, \dots t_E\}$ according to R
    Select the minimum among $\{T, T_{init}, t^*_{nonca}\}$, determine the next event to happen
    Update t^*_{cur}
    If initiation of adenoma or serrated polyp
      Assign location in colon and rectum
      Update M
    If progression to advanced states of pre-cancer lesions
      Update M
    # check if death, progression to cancer or 20 lesions already
    If non-cancer caused death
      Update M
      End individual i, start individual $i + 1$
    Else if progression to pre-clinical cancer
      Update t^*_{nonca}
      Update M
      Generate sojourn time t_{soj}, cancer stage at detection st and cancer survival time t_{sur}
      Repeat
        Select the minimum among $\{t_{soj}, t_{sur}, 10, t_{nonca}\}$, determine the next event to happen
        Update t^*_{cur}
        If cancer-caused death
          Update M
          End individual i, start individual $i + 1$
        If non-cancer caused death
          Update M
          End individual i, start individual $i + 1$
        If individual i survives over 10 years
          Update $t^*_{cur} = t_{nonca}$
          Update M
          End individual i, start individual $i + 1$
        Update t^*_{nonca}
    Update t^*_{nonca}
End For
Output M with respective time interval when events happened
Calculate the summary statistics

2.2.2 CRC natural history assumptions

DECAS simulates two CRC tumorigenesis pathways: the adenoma-carcinoma pathway and the serrated neoplasia pathway (see **Figure 3**). The precancerous lesions progress from non-advanced to advanced stages, to pre-clinical (asymptomatic), and then to clinical (symptomatic) cancers. Each simulated individual may have up to 20 of adenomas or serrated polyps in total throughout the simulated time horizon (Greuter et al., 2016), and only the first one transitioning to pre-clinical cancer will be followed up thereafter (van Ballegooijen et al., 2011). DECAS does not consider lesion regression. We defined advanced adenoma (AA) as adenomas >10mm, with villous components or high-grade dysplasia (Keum and Giovannucci, 2019) and clinically relevant serrated polyps (crSP) as serrated polyps ≥ 10 mm or serrated polyps > 5 mm located proximally to the splenic flexure (Schramm et al., 2018). The current version of DECAS assumed that serrated polyps contribute to 15% of CRC for easier comparison with other studies (Brenner et al., 2013; Greuter et al., 2014).

DECAS takes input of 21 transition-related parameters to randomly generate time to events, driving the transitions of the model (see **Table 6**). All the priors of the parameters were assumed to be uniformly distributed (as required by the calibration algorithm followed (Lenormand et al., 2013)), and the ranges of the priors are elicited in the following sub-chapters. The model was programed and calibrated in R software (version 4.0.4).

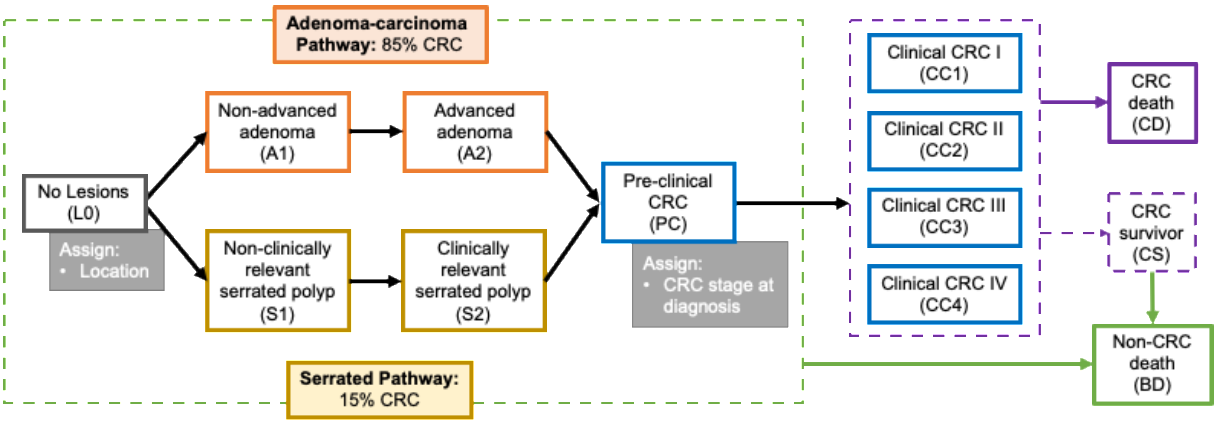


Figure 3. DECAS schematic model structure.

2.2.3 DECAS model components

2.2.3.1 Precancerous lesion initiation

The model starts with lesion occurrence, either as a non-AA or non-crSP (state L0 to A1 or S1 in **Figure 3**). DECAS follows the method by Rutter et al. (2009) to model the occurrence of adenoma or non-crSP with a nonhomogeneous Poisson process (**Equation 1**). The risk was assumed to be a function of the baseline individual risk, sex and piece-wise age effects (Brenner et al., 2014a).

Let $\psi_{1li}(t)$ denote the ith individual's personal risk of developing a pre-cancer lesion at time t, where l denotes the pathway, $l = \{adenoma, serrated\ polyp\}$. The log-risk is a linear function of the individual's baseline risk, sex and increases piece-wisely with age:

$$\psi_{1li}(t) = \exp\left\{\alpha_{0li} + \alpha_{1l}sex_i + \sum_{k=1}^{3}\delta(Y_k < age_i(t) \le Y_{k+1})\left\{age_i(t)\alpha_{2lk} + \sum_{j=2}^{k} Y_j(\alpha_{2lj-1} - \alpha_{2lj})\right\}\right\} \quad \text{(E.1)}$$

In the formula, α_{0li} represents an individual's baseline log-risk. α_{1l} represents the additional risk of developing adenomas and serrated polyps in women ($sex_i = 1$) relative to men ($sex_i = -1$) in both type of lesions. α_{2lk} represents the changes in risk with age in the kth interval, $k \in \{1,2,3\}$, for both adenomas and serrated polyps. $\delta(\cdot)$ is an indicator function, which $\delta(x) = 1$ when x is true and $\delta(x) = 0$ otherwise.

It was assumed that the individual's baseline risks for adenoma and serrated polyps to be independently and identically distributed across individuals, and to follow a log-normal distribution (Rutter et al., 2009). The lognormal assumption allowed the majority of the population to be free from precancerous lesions and a minority to be prone to developing one or multiple lesions, as seen in the literature (Pox et al., 2012). Additionally, the two baseline risks were assumed to share the same magnitude of variation. Therefore, an individual's baseline log-risk for developing a precancerous lesion $\alpha_{0li} \sim N(\mu_l, \sigma_0)$. Further assumption included the risk of developing precancerous lesions to be starting at the age of 20 years old. The model specifies three fixed age-risk intervals: [20, 50), [50, 70), [70, 90). Therefore, $Y_1 = 20, Y_2 = 50, Y_3 = 70, and\ Y_4 = 90$. Risk increases log-linearly within these age-risk intervals.

The cumulative risk (Rutter et al., 2009) for the ith individual to develop precancerous lesions by time t is then given as (**Equation 2**):

$$\Psi_{1li}(t) = \int_{20}^{age(t)} \psi_{1li}(u)du$$
$$= e^{\alpha_{0li}+\alpha_{1l}\mathrm{sex_i}} \sum_{k=1}^{3} \left\{ \delta(\mathrm{age_i(t)} > Y_{\mathrm{k}}) \left(\frac{e^{\alpha_{2lk}\min(Y_{k+1},\mathrm{age_i(t)})} - e^{\alpha_{2lk}Y_k}}{\alpha_{2lk}} \right) exp\left(\sum_{j=2}^{k} Y_j \left(\alpha_{2_{lj-1}} - \alpha_{2_{lj}} \right) \right) \right\} \quad \text{(E.2)}$$

The prior ranges for adenoma related parameters were informed by Rutter et al. (Rutter et al., 2009); for serrated polyp parameters, analogous ranges were assumed (Greuter et al., 2014; Lew et al., 2017).

2.2.3.2 Progression to the advanced stage of precancerous lesion

At this stage, non-advanced lesions progress to advanced lesions (state A1 to A2 or S1 to S2 in **Figure 3**). Once a lesion, either non-AA or non-crSP, appears in the model, locations (proximal, distal colon or rectum) will be assigned to the lesion with probabilities according to the proportions observed in literature (Greuter et al., 2014; Schramm et al., 2018). A constant risk was assumed for non-AA progressing to AA (Brenner et al., 2013) and non-crSP to crSP, respectively. The prior ranges of the parameters were informed by the annual transition probability estimated by Brenner et al. (2013) and Policy1-Bowel (Lew et al., 2017).

2.2.3.3 Progression to pre-clinical cancer

In DECAS, only AA and crSP can progress to pre-clinical CRC (A2 or S2 to PC in **Figure 3**). In analogy with the precancerous lesion initiation, the progression was designed to be a nonhomogeneous Poisson process with a piece-wise age effect and location effect. Sex was not considered in the function describing the progression to pre-clinical cancer, following the literature that analyzed German data (Brenner et al., 2013). Additionally, the sex-specific risk should already be accounted in the lesion initiation function (**Equation 1**).

It was assumed that the baseline risk of pre-clinical cancer transition for the *i*th individual at age of 20, and, for each pathway, the risk to progression was deemed as a function of location and a piece-wise constant with change-points at the age 50 and 70 (Rutter et al., 2009). The risks for adenomas and serrated polyps have the form (**Equation 3**):

$$\psi_{3li}(t) = \exp\left\{ \beta_{0\mathrm{li}} + \beta_{1\mathrm{l}} location_{li} + \sum_{k=2}^{3} \delta(Z_k < age_i(t) \leq Z_{k+1})\, \beta_{2\mathrm{lk}} \right\} \quad \text{(E.3)}$$

In the above functions, $\beta_{0\text{li}}$ is the baseline log-risk to become preclinical colon cancer for male at age of 20. β_{1l} represents the additional risks for lesions located at rectum ($location_{li} = 1$) versus colon ($location_{li} = 0$). Similar to the risk of lesion initiation, β_{2lk} represents the changes in risk with age in the kth interval, $k \in \{2,3\}$ as compared to the reference interval $k = 1$, for both adenomas and serrated polyps. The age-risk intervals are again: [20, 50), [50, 70), and [70, 90); therefore, $Z_2 = 50,\ Z_3 = 70, and\ \ Z_4 = 90$.

The choice of the factors determining the risk was supported by the literature (Brenner et al., 2013). The ranges of priors were again informed by Brenner et al. (2013) and Policy1-Bowel (Lew et al., 2017).

2.2.3.4 Cancer detection

A pre-clinical cancer may progress to clinical cancer at this stage (PC to CC in **Figure 3**). It was assumed that the first lesion becoming pre-clinical cancer will determine the cancer stage at detection and the respective stage-specific 10-year CRC survival, which are based on the CRC stage distribution and survival data from the literature (Kubisch et al., 2016; Human Mortality Database, 2018; Robert Koch Institute, 2019). When pre-clinical cancers become symptomatic, they will be detected as clinical cancers. The time from the start of being a pre-clinical cancer until clinical cancer detection, defined as sojourn time, was randomly drawn from a Weibull distribution with shape and scale parameters equal to 5.4 and 5.1, respectively. The choice of the Weibull parameters yields a mean of 4.7 and standard deviation of 1, which covers the range of sojourn time estimated in the literature (Lansdorp-Vogelaar et al., 2009; Brenner et al., 2011).

2.2.3.5 Death

Individuals could die from non-cancer causes at any time in the model, whereas only persons with a clinical cancer are subject to cancer-specific death (any to BD or CC to CD in **Figure 3**). People with a clinical cancer who survive more than 10 years were assumed to only be at risk of non-cancer cause mortality in the rest of the life course (Greuter et al., 2014).

Table 6. Summary of DECAS parameters and the priors

Model parameters	Prior distribution	Reference
Adenoma		
Baseline log-risk, mean	~ U(-9, -4.6)	(Rutter et al., 2009; Greuter et al., 2014; Brenner et al., 2014a)
Baseline log-risk, standard deviation	~ U(0.8, 5.4)	
Sex effect	~ U(-0.65, 0)	
Age effect, 20 ≤ age < 50 years	~ U(-0.06, 0.1)	
Age effect, 50 ≤ age < 70 years	~ U(-0.1, 0.15)	
Age effect, age ≥ 70 years	~ U(-0.1, 0.2)	
Serrated polyp		
Baseline log-risk, mean	~ U(-9.8, -5.4)	Assumption based on (Greuter et al., 2014; Lew et al., 2017)
Baseline log-risk, standard deviation	Same as in adenoma	
Sex effect	~ U(-0.65, 0.25)	
Age effect, 20 ≤ age < 50 years	~ U(-0.12, 0.06)	
Age effect, 50 ≤ age < 70 years	~ U(-0.12, 0.15)	
Age effect, age ≥ 70 years	~ U(-0.12, 0.2)	
Hazard of non-AA progressing to AA	~ U(0.002, 0.3)	Assumption based on (Greuter et al., 2014; Lew et al., 2017)
Hazard of non-crSP progressing to crSP	~ U(0.002, 0.6)	
Adenoma		
Base risk of colonic lesion progressing to pre-clinical cancer, male at age 20 years	~ U(0.002, 0.3)	Assumption based on (Rutter et al., 2009; Brenner et al., 2013; Greuter et al., 2014)
Location effect, rectum	~ U(2, 30)	
Age effect, 50 ≤ age < 70 years	~ U(1, 5)	
Age effect, age ≥ 70 years	~ U(1.2, 10)	
Serrated polyp		
Base risk of colonic lesion progressing to pre-clinical cancer, male at age 20 years	~ U(0.002, 0.6)	Assumption based on (Greuter et al., 2014; Lew et al., 2017)
Location effect, rectum	~ U(4, 50)	
Age effect, 50 ≤ age < 70 years	~ U(1, 5)	
Age effect, age ≥ 70 years	~ U(1.2, 10)	

Note: CI, credible interval; ~ U(a, b) denotes the uniform distribution bounded by (a, b); L0: no lesions; A1: non-advanced adenoma; S1: non-clinically relevant serrated polyp; A2: advanced adenoma; S2: clinically relevant serrated polyp; PC: pre-clinical cancer

2.2.4 Model calibration with an ABC approach

2.2.4.1 ABC rejection sampler

Let θ be the parameter set to be estimated, $\pi(\theta)$ be its prior distribution and $f(s|\theta)$ be the likelihood function of θ for a set of summary statistics s, representing a reduction of the data y to a lower dimensional set. The aim of ABC is to approximate the posterior distribution, $\pi(\theta|s) \propto f(s|\theta)\pi(\theta)$, while avoiding direct computation of the likelihood $f(s|\theta)$ (Beaumont, 2019). This is accomplished by repeatedly drawing samples θ^* from the prior, simulating summary statistics s according to the model $f(s|\theta^*)$, and retaining the proposed samples if the simulated output s and the observed data summary statistics s_y have distance $d(s_y, s) < \epsilon$, for a pre-specified distance measure d and threshold ϵ (Toni et al., 2009). After a suitable number of iterations, N samples of

the parameter set θ are obtained from the distribution $\pi(\theta|d(s_y, s) \leq \epsilon)$, which should be a good approximation for the posterior distribution $\pi(\theta|s_y)$ if ϵ is small enough (Toni et al., 2009; Beaumont, 2019). The simplest and generic form of ABC algorithm is the ABC rejection sampler proposed by Pritchard et al. (Pritchard et al., 1999; Toni et al., 2009), and Euclidean distance is frequently used to determine the distance $d(s_x, s_y)$ between s_x (the summary statistics from the simulated samples) and s_y (**Algorithm 2**).

Algorithm 2. ABC rejection sampler (adapted from Lenormand et al.(2013))

Given N as the number of samples
For $i = 1$ to N **do**
  Repeat
    Generate $\theta^* \sim \pi(\theta)$
    Simulate a dataset x according to $f(x|\theta^*)$, and compute the corresponding summary statistics $s_x = s(x)$
  Until $d(s_x, s_y) < \epsilon$, where $\epsilon \geq 0$
  Set $\theta_i = \theta^*$
End For

2.2.4.2 Adaptive population Monte Carlo

Since the emergence of the ABC rejection sampler, research has been conducted intensively to improve its efficiency, particularly for models with high dimensional parameter spaces, where it is very computationally demanding to sample the whole space (Lenormand et al., 2013; Beaumont, 2019). To improve efficiency for DECAS parameter calibration, we chose to use one algorithm from the ABC sequential Monte Carlo (ABC-SMC) family (Sisson et al., 2007; Toni et al., 2009; Beaumont, 2019): the adaptive population Monte Carlo (APMC) (Lenormand et al., 2013). ABC-SMC describes a group of algorithms which are based on sequential importance sampling (Del Moral et al., 2006) by sampling multiple parameter sets of $\theta_i^*, \dots \theta_N^*$ first from the prior distribution $\pi(\theta)$. The samples are weighted according to the importance to form a new proposal distribution for the next draws. Therefore, the initial samples propagate through a sequence of intermediate distributions, $\pi(\theta|d(s_y, s) \leq \epsilon_i), i = 1, \dots, T-1$, until the samples converge with the target distribution $\pi(\theta|d(s_y, s) \leq \epsilon_T)$. The level of tolerance decreases such that $\epsilon_1 > \dots > \epsilon_T \geq 0$, so the distributions evolve towards the target distribution (Toni et al., 2009).

APMC follows the principles of ABC-SMC general algorithm. While automatically downward adjusting ϵ in each step, APMC keeps a pre-specified α proportion of the samples, and it stops

when reaching the pre-defined threshold of proposed samples acceptance rate $P_{acc_{min}}$ (Lenormand et al., 2013), as described in the steps below:

1. At cycle $T = 1$, obtain N particles by running Algorithm C and keep αN particles with the shortest distances with a distance threshold ϵ_T.
2. For $T \geq 1$ and if $p_{acc} > p_{acc_{min}}$,
   2.1. Randomly generate new $(N - \alpha N)_{T-1}$ particles from the weighted αN_{T-1} particles using a Gaussian kernel.
   2.2. Update p_{acc}: the proportion of $(N - \alpha N)_{T-1}$ newly generated particles below ϵ_{T-1}.
   2.3. Combine the new $(N - \alpha N)_{T-1}$ particles with the αN_{T-1} particles to calculate the distances, so that we have αN_T new particles with the shortest distances together with their associated weights and distance tolerance ϵ_T.
3. Repeat steps 2.1 to 2.3 until $p_{acc} \leq p_{acc_{min}}$.

More detailed APMC algorithm is described in **Algorithm 3**. APMC has been demonstrated to converge to the target distribution faster than some other well-known ABC-SMC algorithms while maintaining the quality of posterior approximation for complex models (Lenormand et al., 2013; Bonassi and West, 2015).

Algorithm 3. APMC detailed algorithm (adapted from Lenormand et al.(2013))

Given N, $N_\alpha = \alpha N$ the number of particles to keep at each iteration among the N particles ($\alpha \in [0,1]$), $p_{acc_{min}}$ the minimal acceptance rate and s_y the summary statistics of dataset y.
For $T = 1$ **do**
  For $i = 1$ to N **do**
    Simulate $\theta_i^{(0)} \sim \pi(\theta)$ and $x \sim f(x|\theta_i^{(0)})$
    Set $d_i^{(0)} = d(s_x, s_y)$
    Set weight $\omega_i^{(0)} = 1$
  End For
  Let $\epsilon_1 = Q_{d^{(0)}}(\alpha)$ the first α-quantile of $d^{(0)}$ where $d^{(0)} = \left\{d_i^{(0)}\right\}_{1 \leq i \leq N}$
  Let $\left\{\left(\theta_i^{(1)}, \omega_i^{(1)}, d_i^{(1)}\right)\right\} = \left\{\left(\theta_i^{(0)}, \omega_i^{(0)}, d_i^{(0)}\right) | d_i^{(0)} \leq \epsilon_1, 1 \leq i \leq N\right\}$
  Take σ_1^2 as twice the weighted empirical variance of $\left\{\theta_i^{(1)}, \omega_i^{(1)}\right\}_{1 \leq i \leq N_\alpha}$
  Set $p_{acc} = 1$
  $T \leftarrow T + 1$
End For
While $p_{acc} > p_{acc_{min}}$ **do**
  For $i = N_\alpha + 1$ to N **do**
    Select θ_i^* from $\theta_j^{(T-1)}$ with probability $\frac{\omega_j^{(T-1)}}{\sum_{k=1}^{N_\alpha} \omega_k^{(T-1)}}$, $1 \leq j \leq N_\alpha$
    Generate $\theta_i^{(T-1)} | \theta_i^* \sim \mathcal{N}(\theta_i^*, \sigma_{(T-1)}^2)$ and $x \sim f(x|\theta_i^{(T-1)})$
    Set $d_i^{(T-1)} = d(s_x, s_y)$

Set $\omega_i^{(T-1)} = \frac{\pi(\theta_i^{(T-1)})}{\sum_{j=1}^{N_\alpha}\left(\omega_j^{(T-1)}/\sum_{k=1}^{N_\alpha}\omega_k^{(T-1)}\right)\sigma_{T-1}^{-1}\varphi\left(\sigma_{T-1}^{-1}\left(\theta_i^{(T-1)}-\theta_j^{(T-1)}\right)\right)}$

End For

Set $p_{acc} = \frac{1}{N-N_\alpha}\sum_{k=N_\alpha+1}^{N}\delta(d_i^{(T-1)} < \epsilon_{T-1})$

Let $\epsilon_T = Q_{d^{(T-1)}}(\alpha)$ where $d^{(T-1)} = \left\{d_i^{(T-1)}\right\}_{1\le j\le N}$

Let $\left\{\left(\theta_i^{(T)}, \omega_i^{(T)}, d_i^{(T)}\right)\right\} = \left\{\left(\theta_i^{(T-1)}, \omega_i^{(T-1)}, d_i^{(T-1)}\right) | d_i^{(T-1)} \le \epsilon_T, 1 \le i \le N\right\}$

Take σ^2 as twice the weighted empirical variance of $\left\{\theta_i^{(T)}, \omega_i^{(T)}\right\}_{1\le i\le N_\alpha}$

$T \leftarrow T+1$

End While

Where $\forall u \in [0,1]$ and $X = \{x_1, \dots, x_n\}$, $Q_X(u) = inf\{x \in X | F_X(x) \ge u\}$ and $F_X(x) = \frac{1}{n}\sum_{k=1}^{n}\delta(x_k \le x)$

Where $\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$

2.2.4.3 Calibration set-up

For DECAS calibration, a population of 30,000 was simulated, which was similar to the size of each five-year age group in the screening registry data (Kretschmann et al., 2020). We first used priors informed by literature and ran a series of test runs to gauge the ranges of priors. After confirming sensible prior ranges from test runs, pilot run to sample 50,000 parameter sets using Latin Hypercube sampling (Vanni et al., 2011) was conducted to explore the prior spaces efficiently (**Table 6**). Given the pilot results, ABC rejection sampler was applied to select 1,000 samples with the shortest standard deviation-weighted Euclidean distance. Based on the selected samples, the prior ranges were extended upon observing evidence of potential prior-data conflict, and the ranges were narrowed when regions of the parameter space had very small probability mass (Fearnhead and Prangle, 2012). Next, with the fine-tuned priors, the APMC algorithm was implemented with $\alpha = 0.1$ and $P_{acc_{min}} = 0.05$ (as suggested by Lenormand et al. (2013)) with 10,000 simulations in each cycle. The calibration with APMC algorithm was performed via the R package EasyABC (Jabot et al., 2013) and parallel computing using a 60-core cluster computer.

2.2.5 Data sources for calibration targets and parameters

Some parameters were not calibrated but directly input to the model: adenoma and serrated polyp location in colon and rectum were informed by screening colonoscopy studies (Greuter et al., 2014; Schramm et al., 2018) (**See Appendix B Supplementary Table 2**). Clinical cancer stage distribution and stage-specific CRC mortality were input from the Bavarian Cancer Registry data (Kubisch et al., 2016; Munich Canncer Registry, 2018) (**See Appendix B Supplementary Table 3**). The background mortality was taken from the German life table 2010-2014 (Human Mortality

Database, 2018), and it was adjusted by removing CRC-specific mortality from the German Centre for Cancer Registry Data (ZfKD) in the same period (Robert Koch Institute, 2019) (**See Appendix B Supplementary Table 4**). Lastly, the mean sojourn time was taken as 4.7 years (95% CI 4.5-4.9) from a study using German screening colonoscopy registry data (Brenner et al., 2011).

In total, 74 CRC epidemiological data points were used as calibration targets to calibrate the model (**See Appendix C Supplementary Table 5**). An important data source for the calibration targets was the data from German screening colonoscopy registry, and its details have been described elsewhere (Pox et al., 2012). The calibration of adenoma prevalence was from the registry data of 3.3 million first-time average-risk participants aged 55 years old and older in the period of 2007-2014. For CRC prevalence, we used the screening colonoscopy registry data in the period of 2003-2006, when the effect of screening colonoscopy to CRC incidence was still minimal. The serrated polyp prevalence and the proportion of multiple lesions were calibrated with the data from a study, which included 4,161 screening colonoscopies among average-risk individuals aged 50 years old and older in North Rhine-Westphalia, Germany during the period of 2012-2016 (Schramm et al., 2018). The prevalence data for the age group 40-49 years old were derived by applying the proportion from a meta-analysis of screening colonoscopy studies (Leshno et al., 2016), and all target prevalence were upward-corrected considering colonoscopy miss rate from a meta-analysis (Zhao et al., 2019). Regarding the 15% assumption for serrated neoplasia pathway, the proportions (85% from AA and 15% from crSP) were applied to the target CRC prevalence data to derive the prevalence of CRCs developing from adenomas or serrated lesions, respectively, for model calibration.

2.2.6 DECAS natural history validation

Natural history validation for DECAS was performed according to the modeling guidelines (Eddy et al., 2012), which includes internal and external validation. Concerning external validity, DECAS predicted age- and sex-specific CRC incidence were validated against the data from ZfKD in 2003-2006 (Robert Koch Institute, 2019) (**See Appendix C Supplementary Table 6**).

2.2.7 DECAS natural history dwell time

As mentioned in Chapter 1.3.2.2, dwell time and sojourn time are among the most critical unobservable parameters in CRC natural history modeling. Sojourn time is a directly input from a Weibull distribution in DECAS (Chapter 2.3.3.4). On the other hand, although dwell time is not a direct input or target for calibration in DECAS, it can be regarded as an output which is estimated indirectly through the calibration with age-specific prevalence data. Knowing the period of dwell

time will help analyze the screening effects predicted by DECAS screening model (Rutter et al., 2016). Therefore, in the same simulation exercise generating the outputs for incidence validation with ZfKD, the adenoma and serrated dwell time in DECAS were recorded. The dwell time here is defined as time from the first non-AA or non-crSP appearance in DECAS until the occurrence of first pre-clinical cancer.

2.3 DECAS screening model

(Part of Chapter 2.3 has been submitted and currently under review (Cheng et al., 2022)*)*

2.3.1 General assumptions of DECAS screening model

DECAS screening component is superimposed on the natural history model. If an individual participates in the screening, the time to next screening will compete with all other time to events in the natural history model. If the time to next screening is the shortest, the individual will receive a screening test. Depending on the type of screening test and test characteristics, it will determine what action to be taken for this individual. The same works for the time to surveillance if individuals has been found to have lesions in the screening and participate in the subsequent surveillance. It is assumed in DECAS that if the participants participate in the first screening or surveillance round, they will adhere to the participation in all the following rounds.

In general, before any lesion reaches the clinical cancer state within an individual, all lesions in an individual are subject to detection by screening or surveillance tests and will be removed upon detection by screening, follow-up or surveillance colonoscopy. Screening-detected cancer stage distribution was applied based on the German screening colonoscopy registry data (Pox et al., 2012) (**Appendix B Supplementary Table 3**). In the following sections, different actions following different screening tests will be described.

2.3.1.1 Screening with stool-based tests

If an individual participates in the stool-based test screening (e.g., gFOBT or FIT), DECAS will first determine if the test can successfully detect the lesion present in the individual. If the stool-based test detects the lesion, the individual will be referred to a follow-up colonoscopy. Next, if the individual participates in the follow-up colonoscopy and the colonoscopy successfully detects the lesions, all lesions will be removed. If positive stool tests followed by a negative colonoscopy finding, the stool test results are deemed false positive and next stool-based screening will be in

10 years (Gemeinsamer Bundesausschuss (G-BA), 2018). The screening management flow of stool-based test (FIT as an example) is shown in **Figure 4**.

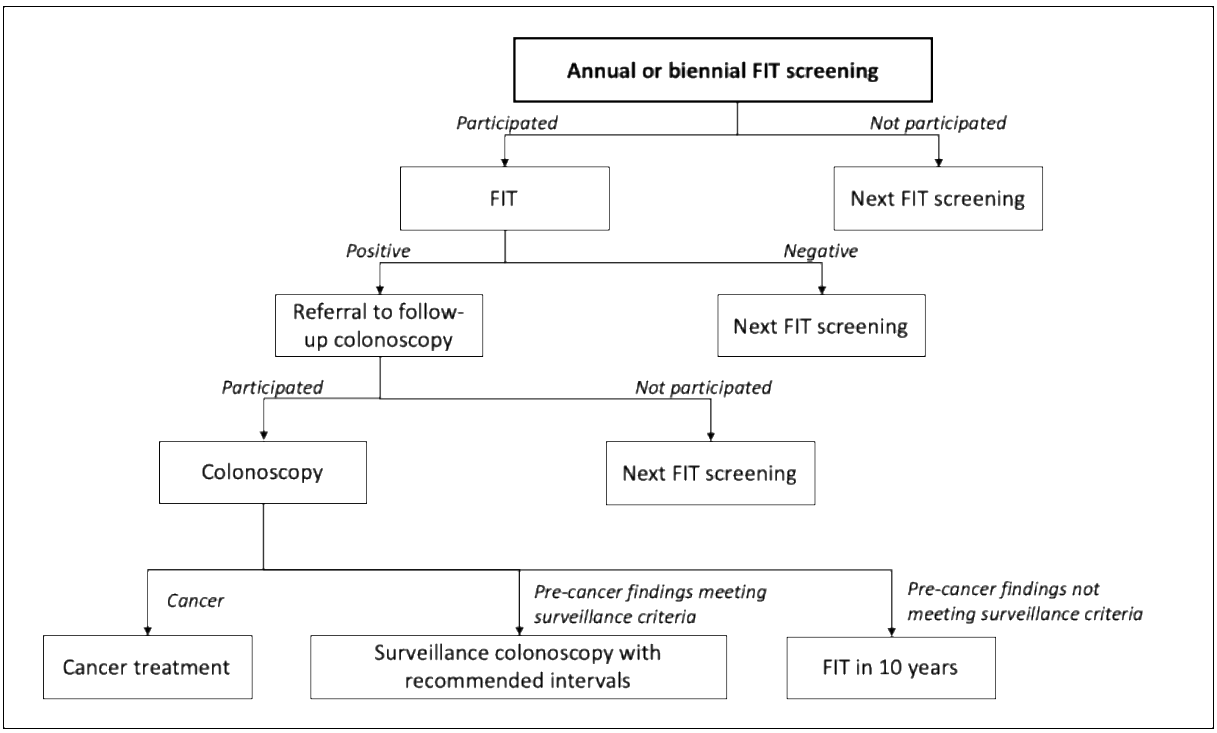


Figure 4. Management flow of screening with annual or biennial fecal immunochemical tests
FIT = fecal immunochemical test. Surveillance colonoscopy management for Germany is detailed in Table 14.

2.3.1.2 Screening with direct visualization tests

Two direct visualization tests are discussed here: FS and colonoscopy. In DECAS screening model, FS was assumed to visualize only rectum and sigmoid colon, while colonoscopy was assumed to visualize the entire colon. Both were assumed to have 100% reach rate for the parts of colon they can visualize.

For FS screening in DECAS, if an individual participates in the screening and FS detects lesions, the lesions in rectum and sigmoid colon will be immediately removed; furthermore, the individual will be referred to a follow-up colonoscopy. If one participates in follow-up colonoscopy and lesions are detected, rest of the lesions will be completely removed.

For colonoscopy screening, if an individual participates in colonoscopy screening, upon detection, the lesions in the entire colon will be taken out. The screening flows of one-time FS and 10-yearly colonoscopy screening are in **Figure 5** and **Figure 6**, respectively.

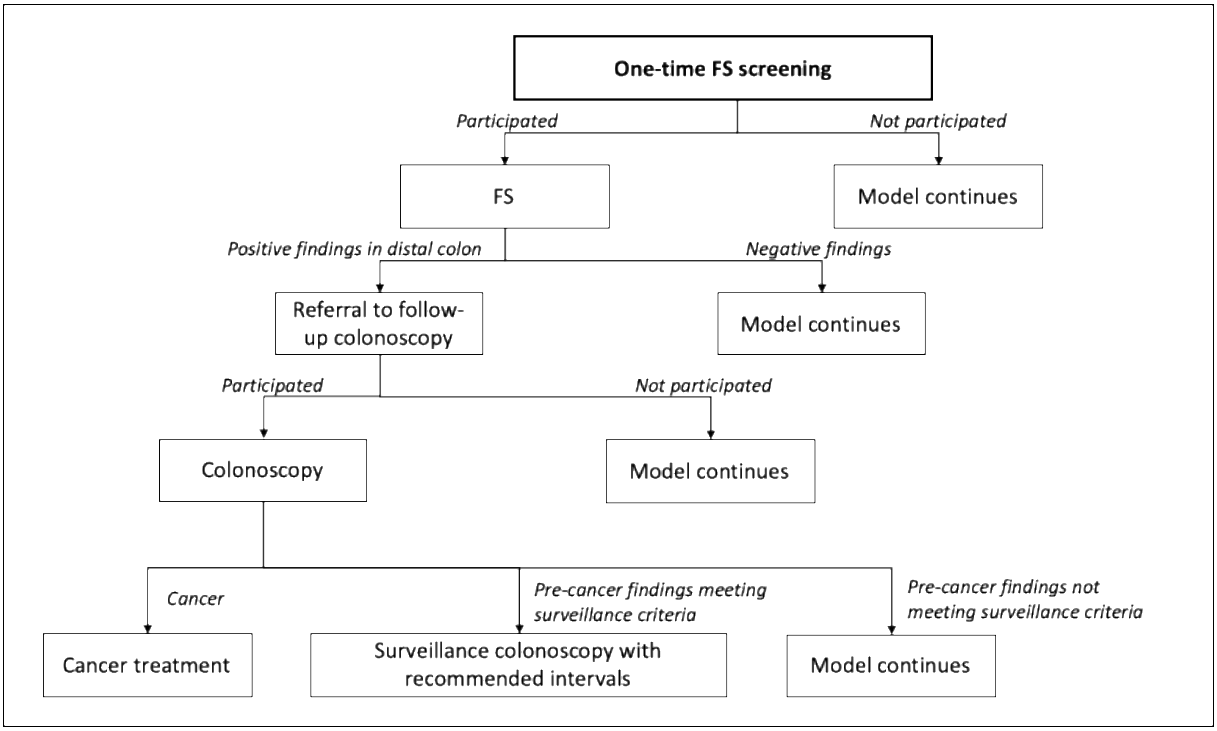


Figure 5. Management flow of screening with one-time flexible sigmoidoscopy
FS = flexible sigmoidoscopy. Surveillance colonoscopy management for Germany is detailed in Table 14.

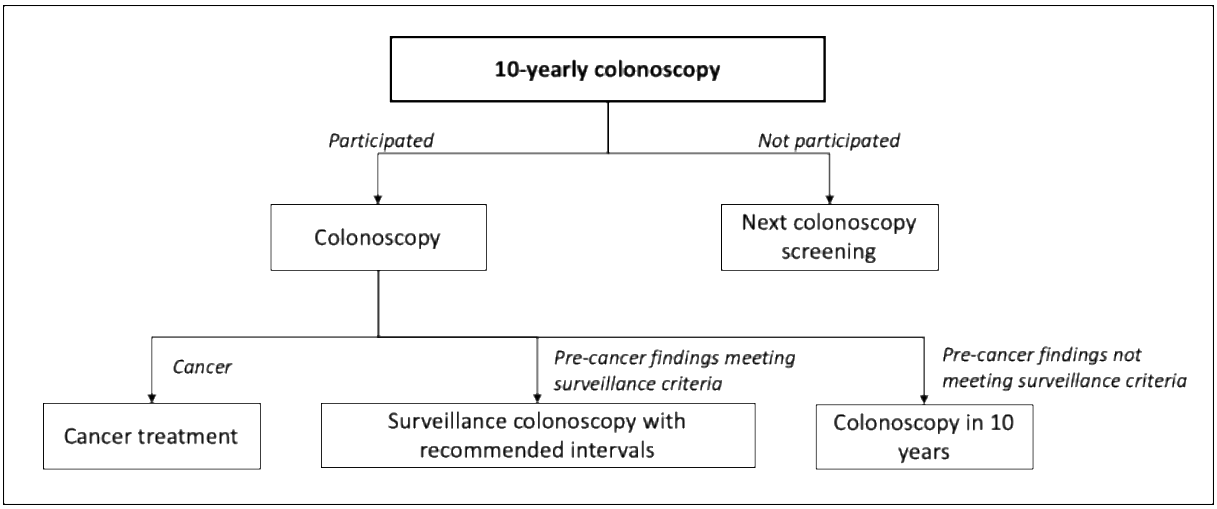


Figure 6. Management flow of screening with 10-yearly colonoscopy
Surveillance colonoscopy management for Germany is detailed in Table 14.

2.3.1.3 Screening test sensitivity when multiple lesions are present

Given that DECAS allows up to 20 precancerous lesions to co-exist in an individual, if multiple lesions of different states (non-advanced or advanced) or types (adenomas or serrated lesions) are present at the time of screening, DECAS needs to determine which test sensitivity for detection

to use. The DECAS logics are as follows: detection rates of advanced lesions take precedence over non-advanced lesions, and detection of serrated lesions has priority over adenomas. Larger or more advanced lesions can naturally be more easily detected by either FIT or colonoscopy. On the other hand, to ensure DECAS considers the fact that both FIT and colonoscopy have worse detection power for serrated lesions (Chang et al., 2017; Zhao et al., 2019), it takes a more conservative approach by applying the probability to detect serrated lesion when both adenomas and serrated lesions co-exist.

2.3.2 Validation of DECAS screening model

Similar to the validation of the DECAS natural history model (Chapter 2.2.6), external and cross validation were performed to assess the prediction credibility of DECAS screening model (Eddy et al., 2012). The aim of these validation exercises was to establish the confidence of DECAS' prediction power for the screening effectiveness using difference screening modalities, including direct visualization and stool-based tests.

2.3.2.1 Validation with UK Flexible Sigmoidoscopy Screening trial

The UK Flexible Sigmoidoscopy Screening (UKFSS) trial is one of the largest randomized control trials evaluating the effectiveness of one-time FS screening for CRC, in which 170,432 participants aged between 55 and 64 years old were recruited across the UK in 1990-1994. It is also the FS screening trial providing the longest available (up to 17 years) follow-up data (Atkin et al., 2010, 2017). In addition, the CISNET models conducted joint external validations using 10-year and 17-year UKFSS follow-up data (Rutter et al., 2016; DeYoreo et al., 2020), which also provided a cross validation opportunity for DECAS.

The parameters used in the UKFSS validation are summarized in **Table 7**. To replicate the UKFSS trial populations, DECAS simulated a no-screening (N=112,936) and a one-time FS screening (N=40,621) population with 51% of female participants. The average age of the screening participants in the UKFSS was 60 years old, therefore, it was assumed the DECAS FS screening cohort received the one-time screening at 60 years old, and they were followed up for 17 years.

Within the reach of FS (complete inspection of sigmoid colon) and colonoscopy, the sensitivities to detect adenomas and CRCs were assumed to be identical with the values used in the CISNET models' joint validation (Rutter et al., 2016; Knudsen et al., 2020), and those for serrated polyps were based on colonoscopy miss rates estimated in a meta-analysis (Zhao et al., 2019). To note, the definition of distal colon in the UKFSS trial included only sigmoid and rectum (Atkin et al., 2017).

Table 7. Parameter assumptions for DECAS validation with UKFSS trial

Variable	Value
One-time FS intervention age	60
Number of people receiving FS[1]	40,621
Female (%)	51%
Adherence of follow-up colonoscopy after positive FS	100%
Adherence of surveillance colonoscopy	80%
Endoscopy performance (FS & colonoscopy)[2]	
Sensitivity	
Non-AA	0.8
AA	0.95
Non-crSP	0.73
crSP	0.76
Cancer	0.95
Specificity for FS	0.92

Note: AA = advanced adenoma; crSP = clinically relevant serrated polyps; FS = flexible sigmoidoscopy.

(1) There were 57,089 people allocated to the FS screening group in the UKFSS trial. However, only 40,621 of them participated and received an FS.

(2) The sensitivity of FS within its reach (up to sigmoid).

Table 8. Assumptions of surveillance colonoscopy intervals after lesion removal used in the DECAS validation with UKFSS trial

Risk	Colonoscopy findings[1]	Recommended surveillance interval
Low	1-2 small adenomas (<10 mm)	Routine screening
Intermediate	3-4 small adenomas (<10 mm) OR ≥1 adenoma 10-19 mm or villous or high-grade neoplasia	3 years
High	≥ 5 small adenomas (<10 mm) OR ≥1 adenoma ≥20 mm	With 1 year[2]

Note:

(1) Assuming the adenomas include serrated lesions in DECAS.

(2) 1 year was used in the validation exercise.

The criteria of referral to a follow-up colonoscopy after positive FS findings and the follow-up schedule for surveillance colonoscopy used in this validation exercise followed the UK recommendations (Atkin et al., 2012) and the assumptions by Rutter et al. (2016) (**Table 8**). The adherence rate to surveillance colonoscopy was assumed to be 80% (Rutter et al., 2016).

One thousand DECAS simulations with the calibrated posterior parameters were performed for both no-screening and screening cohorts, respectively. Primary outputs for comparison were the 17-year relative hazard ratios (HRs) for CRC incidence and mortality rates between the screening group and no-screening group. DECAS predictions will be deemed accurate if the mean HRs are within the 95% CIs of the UKFSS estimates (DeYoreo et al., 2020). DECAS predictions will also be compared with the estimates from the CISNET models.

2.3.2.2 Validation with ESTHER cohort study in Germany

ESTHER study is an ongoing prospective population-based cohort study conducted in Saarland, Germany, in which 9,949 male and female residents of Saarland aged 50–75 years with adequate German language knowledge were recruited in 2000-2002. The health and socioeconomic characteristics of the study cohort was reported to be representative of average German residents (Guo et al., 2021).

In the study, participants were followed up 2, 5, 8 and 17 years after enrolment, and they were asked for the information, among others, whether they have participated in screening colonoscopies. Medical records were also reviewed to identify any incidence and mortality of major chronic diseases, including CRC (Guo et al., 2021). Therefore, ESTHER study contains up to 17-year CRC relevant information and allows DECAS to validate its prediction power for colonoscopy screening on the effectiveness of reduction in CRC incidence and mortality.

The average age at recruitment for ESTHER study was 61 years old, and 55% of them were women. 5,388 (59%) of the study participants received at least one screening colonoscopy (Guo et al., 2021). Applying to the DECAS validation exercise, it was assumed that the screening cohort received a one-off colonoscopy at 61 years old and was followed up for 17 years. 1,000 simulations each for no-screening and screened cohorts were performed in DECAS. To ensure more stable simulation results, ten folds of the ESTHER cohort size was simulated: 38,190 people without screening and 53,880 received one-time screening colonoscopy, both with 55% women.

Colonoscopy was assumed to visualize the entire colon, and sensitivities of colonoscopy to detect different stages of lesion were informed by a meta-analysis (Zhao et al., 2019). The definition of distal colon here included descending colon, sigmoid and rectum. The summary of input parameters is in **Table 9**. The follow-up intervals for surveillance colonoscopy after positive screening colonoscopy were in line with the German S3 guidelines for CRC (**Table 10**) (GGPO, 2019). The surveillance colonoscopy was assumed to stop at age of 75 years (Greuter et al., 2017; Lew et al., 2018).

The primary target for comparison was 17-year relative HRs for CRC incidence and mortality rates between screened and no-screened persons. The accuracy of DECAS' prediction power was evaluated if the mean DECAS predicted HRs were within the 95% CIs of the estimations in the ESTHER study.

Table 9. Parameter assumptions for DECAS validation with ESTHER study

Variable	Value
One-time colonoscopy intervention age	61
Number of people receiving FS	53,880
Female (%)	55%
Adherence of surveillance colonoscopy	63%
Coloscopy performance	
Sensitivity	
Non-AA	0.76
AA	0.91
Non-crSP	0.73
crSP	0.76
Cancer	0.95

Note: AA = advanced adenoma; crSP = clinically relevant serrated polyps.

Table 10. Assumptions of surveillance colonoscopy intervals after lesion removal used in the DECAS validation with ESTHER study

Colonoscopy findings	Recommended surveillance interval
1-2 small tubular adenomas (<10 mm)	5-10 years[1]
3-4 adenomas OR ≥1 adenoma ≥10 mm or villous or high-grade neoplasia	3 years
≥ 5 adenomas	<3 year[2]
Serrated lesions	Same as adenomas

Note:

(1) DECAS used 7.5 years; (2) DECAS used 1.5 years.

2.3.2.3 Validation with Nottingham FOBT trial in the UK

Among three large FOBT screening trials, the study conducted in Nottingham, UK, since 1991 was the largest (Hewitson et al., 2008), in which 152,850 individuals aged 45-74 years were randomly allocated to a control arm or an intervention arm receiving biennial gFOBT (Hemoccult; Rohm Pharma, Weiterstadt, Germany) for CRC screening (Hardcastle et al., 1996; Scholefield et al., 2012). It provided up to 20 years of follow-up data on CRC incidence and mortality reduction, which are suitable to validate for the stool-based test screening algorithm in DECAS.

In the original study, there were 74,998 (52% female) in the control group and 75,253 (52% female) in the screening group, who were offered three to six rounds of biennial gFOBT screening. 44,838 people in the screening group accepted at least one gFOBT screening, and 28,720 (64%) of them completed all rounds of screening. Median age when participant accepted the first gFOBT was in the age interval 60-64 years (Hardcastle et al., 1996).

In DECAS, 1,000 times of simulation for no-screening controls (N=74,998, 52% female) and FOBT screened population (N=44,838, 52% female) were performed seperately. Screening with gFOBT

was simulated starting at age 60, and the screening cohort was followed up for 20 years. If focusing only on the screened population in the Nottingham trial, 64% of them received all 6 cycles of gFOBT. Assuming the proportion of people receiving one to six gFOBTs followed a linear relationship, the proportions of population receiving two to five gFOBTs were extrapolated to be 71%, 78%, 86% and 93%, respectively. During the screening period, once CRC or adenomas were found in the participants during the follow-up colonoscopy due to a positive gFOBT result, screening stopped (Hardcastle et al., 1996). No surveillance colonoscopy was simulated as it was not specifically mentioned in the original trial report (Hardcastle et al., 1996). The summary of input parameters and test characteristics is in **Table 11**.

Similar to the previous two validation exercises, 20-year HRs of CRC incidence and mortality rates between non-screened and screened groups were the primary target, and DECAS prediction should be between the 95% CIs of FOBT trial estimates to be accurate.

Table 11. Parameter assumptions for DECAS validation with Nottingham gFOBT trial

Variable	Value
Biennial gFOBT intervention starting age	60
Number of people receiving gFOBT[1]	44,838
Female (%)	52%
Participation rate of gFOBT[2]	
1 gFOBT	100%
2 gFOBT	93%
3 gFOBT	86%
4 gFOBT	78%
5 gFOBT	71%
6 gFOBT	64%
gFOBT performance	
Sensitivity	
Non-AA	0.05
AA	0.12
Non-crSP	0.05
crSP	0.05
Cancer	0.4
gFOBT specificity	0.98
Coloscopy performance	
Sensitivity	
Non-AA	0.76
AA	0.91
Non-crSP	0.73
crSP	0.76
Cancer	0.95

Note: AA = advanced adenoma; crSP = clinically relevant serrated polyps; gFOBT = guaiac fecal occult blood test.

(1) There were 75,253 people assigned to the gFOBT screening group in the Nottingham trial. However, only 44,838 of them participated and received at least one gFOBT.

(2) 28,720 participants (64% of all screening participants) received all 6 rounds of screening (3-6 rounds in the trial, and it was assumed all were offered 6 rounds in the validation exercise). The participation rates for 2-5 rounds were extrapolated linearly.

2.4 Cost-effectiveness analysis of German CRC screening program

To evaluate the cost-effectiveness of German CRC screening program, the validated DECAS screening model was used. A cohort of 100,000 average-risk individuals without prior screening and CRC diagnosis was repeatedly simulated assuming they do not receive any CRC screening or undergo various screening strategies under different assumptions (described below). They entered the simulation starting at age of 20 and were followed up until age of 90 or death. All screening strategies were simulated 1,000 times using the posterior samples derived from the Bayesian calibration, and the base-case outputs were presented as the averages of the outputs from the 1,000 simulations.

2.4.1 Research questions

The CEA aimed to address the questions concerning German CRC screening program laid out in Chapter 1.5.2. Specifically, there are three research questions:

1) Is the new sex-differentiated screening offer for men to start colonoscopy screening at the age of 50 years more cost-effective compared with the previous offer, where both men and women were entitled to colonoscopy screening from the age of 55 years?
2) How do the current screening strategies offered in Germany compare with single test modalities and strategies starting at age of 45 years?
3) How do the results in questions 2 vary, if participation rates differ as a result of alternative screening invitation approach?

2.4.2 Screening strategies for evaluation

To address research questions 1 and 2, 10 strategies using FIT and colonoscopy, alone and in combination, with different screening frequencies were designed (see **Table 12**). The first three strategies are based on the current and previous screening modality implemented in Germany, i.e., annual FIT for age 50-54 followed by biennial FIT for age 55-75 (FIT1y50+FIT2y55), annual FIT for age 50-54 followed by twice 10-yearly colonoscopy (FIT1y50+COL10y55) and twice 10-yearly colonoscopy for men starting at age 50 together with annual FIT for women age 50-54 followed by twice 10-yearly colonoscopy (mCOL50/fFIT50+COL55). If an individual receives any colonoscopy, the next round of screening (either FIT or colonoscopy) will only resume 10 years later (G-BA, 2018).

Given that majority of the CRC screening program in other countries are based on a single test modality with fixed screening frequency (Schreuders et al., 2015), three strategies with FIT or

colonoscopy alone with fixed screening frequency starting at age 50 years old were designed: biennial FIT (FIT2y50), twice and three-time 10-yearly colonoscopy (COL10y50 and COL10y50-3X). FIT every two years was chosen based on two reasons: it was the most-assessed frequency in almost all FOBT randomized control trials (Hewitson et al., 2008); moreover, a Dutch randomized trial could not ascertain that an annual FIT screening provides more benefits over a biennial one (van Roon et al., 2013). As for the number of colonoscopies to be offered in a lifetime, the two-colonoscopy strategy was based on the current German CRC screening program design (G-BA, 2018). On the other hand, the estimated outcomes in the most recent USPSTF modeling work showed that strategies with three or more colonoscopies appeared to strike a good balance between benefits and burden (Knudsen et al., 2021); thus, a three-time 10-yearly colonoscopy strategy was included.

Another dimension to explore was to evaluate the benefits, harms and burden of CRC screening in Germany if the screening were offered to average-risk people starting at the age of 45, as was recommended by American Cancer Society (Wolf et al., 2018) and USPSTF (US Preventive Services Task Force, 2021). Therefore, two combined strategies (FIT1y45+COL10y50 and a three-time colonoscopy strategy FIT1y45+COL10y50-3X), one biennial FIT (FIT2y45) and three-time 10-yearly colonoscopy (COL10y45-3X) starting at age of 45 were included. All screening strategies involving FIT were assumed to stop at age of 75, as screening for older than 76 years old does not receive strong recommendations in major screening guidelines (Wolf et al., 2018; USPSTF, 2020; Shaukat et al., 2021) despite the open-ended recommendation made by the German guidelines (GGPO, 2019).

2.4.3 Scenarios for screening participation assumptions

To address research question 3, all the 10 strategies were evaluated under four scenarios with various participation rate assumptions. To note, it was assumed that if an individual participates in a screening strategy, one will adhere to all rounds of screening tests. For instance, if one participates in biennial FIT from the age of 50, he or she will receive FIT screening every two years until the age of 75, tested positive or being diagnosed with cancer, without any break in between.

As the base-case scenario, a perfect participation in all screening rounds and perfect adherence to the follow-up colonoscopy after positive FIT (hereafter FIT-positive colonoscopy) and surveillance colonoscopy was assumed (Scenario 1), reflecting the full willingness of eligible population to participate in screening. This also allowed the assessment of different screening strategies' full impact and the comparison among them. To evaluate how different invitation approaches affect the participation rates in an organized CRC screening program, three further scenarios were assumed.

Table 12. Overview of 10 screening strategies to be evaluated

Strategy short name	Screening strategy description
No screening (comparator)	No screening
Combined modality strategies used in German CRC screening program	
FIT1y50+FIT2y55	Annual FIT for age 50-54 years followed by biennial FIT for age 55-75 years
FIT1y50+COL10y55	Annual FIT for age 50-54 years followed by 2 colonoscopies with 10-year interval starting at age 55 years
mCOL50/fFIT50+COL55	Men: from age of 50, 2 colonoscopies with 10-year interval; Women: annual FIT for age 50-54 years followed by 2 colonoscopies with 10-year interval starting at age 55 years
Single modality strategies starting at age of 50	
FIT2y50	Biennial FIT for age 50-75 years
COL10y50	From age of 50, 2 colonoscopies with 10-year interval
COL10y50-3X	From age of 50, 3 colonoscopies with 10-year interval
Combined and single modality strategies starting at age of 45	
FIT1y45+COL10y50	Annual FIT for age 45-49 years followed by 2 colonoscopies with 10-year interval starting at age 50 years
FIT1y45+COL10y50-3X	Annual FIT for age 45-49 years followed by 3 colonoscopies with 10-year interval starting at age 55 years
FIT2y45	Biennial FIT for age 45-75 years
COL10y45-3X	From age of 45, 3 colonoscopies with 10-year interval

Note: COL = colonoscopy; FIT = fecal immunochemical test.

Scenario 2 reflected the current German CRC population-based organized screening program, where an invitation for screening is sent out to eligible participants at ages of 50, 55, 60 and 65 (Gemeinsamer Bundesausschuss (G-BA), 2018), however, FIT kits are required to be picked up from GPs' offices (Gruner et al., 2021). For the strategies starting at age of 45, an additional invitation will be sent out at that age. Given that the organized screening program only started in mid-2019, data on participation rates under the current invitation scheme are currently not available. Nevertheless, two German RCTs (Hoffmeister et al., 2017; Gruner et al., 2020) were done to evaluate the effect of invitation letters on the participation rates. Combining the study results and the participation rates shortly before the organized screening program roll-out can inform the participation rates for Scenario 2.

Based on the information from the Statisische Bundesamt (2021), the participation rates before organized screening were estimated to be 7% for male and 25% for female in annual FIT screening (FOBT participation rate for age 50-54 in 2017), 16% for male and 24% for female in biennial FIT screening (FOBT participation rate for age 55-75 in 2016- 2017), and 17% for male and 19% for

female in 10-yearly colonoscopy screening (screening colonoscopy participation rate for age 55-75 in 2008-2017). According to the two studies, invitation alone has limited influence on FIT participation but can potentially increase the screening colonoscopy participation rate by 1.3 folds (Hoffmeister et al., 2017; Gruner et al., 2020). Therefore, the annual and biennial FIT participation rates were directly taken from the official estimation, and the participation rate for 10-yearly colonoscopy were assumed to be 23% for men and 24% for women in Scenario 2. As for the adherence to FIT-positive colonoscopy, it was assumed to be 64%, as was found in a German study (Gruner et al., 2020).

Although invitation letters alone does not boost the uptake of stool-based tests, the FIT uptake rate has been shown to increase significantly when the test kit is included in the mail-out invitation letter (Klabunde et al., 2015; Goodwin et al., 2019; Gruner et al., 2020, 2021; Toes-Zoutendijk et al., 2020). To reflect such effect, Scenario 3 was assumed to have test kits posted to all eligible individuals for strategies involving FIT, while only having the invitation letters sent out at fixed time-points for strategies involving colonoscopy screening, as in Scenario 2. Thus, the participation rates for both annual and biennial FIT were assumed to increase to 25% for men and 34% for women, as seen in the results in a German randomized study (Gruner et al., 2020). Uptake of FIT-positive and 10-yearly screening colonoscopy were assumed to remain the same as in Scenario 2, and, as a result, the set-up for colonoscopy-only strategies in Scenario 2 and 3 are identical.

Lastly, a higher adherence scenario (Scenario 4) was designed, taking into consideration of other European population-based CRC screening programs with higher participation rates. These programs mostly provide mail-out gFOBT or FIT screening with at least one reminder letter a few weeks after the kits are posted, and the participation rates range from 44% to 75% (Klabunde et al., 2015; Senore et al., 2019; Toes-Zoutendijk et al., 2020). Among the programs with >70% participation rates, e.g., in Netherlands and Basque country in Spain, they send out an advanced notification before mailing the FIT test kits and a reminder 4-6 weeks afterwards (Toes-Zoutendijk et al., 2020). Thus, Scenario 4 was assumed to follow these invitation measures and to have a high uptake for FIT at 71% for male and 75% for female and that for FIT-positive colonoscopy at 83%, as seen in the Dutch population-based screening program (Toes-Zoutendijk et al., 2020). The invitation for screening colonoscopy was assumed to be the same as in Scenario 3 but additionally with a reminder letter, and 42% was assumed for the uptake rates for both men and women, according to the values found in randomized studies in the US (Singal et al., 2016, 2017).

The summary of invitation measure, participation and adherence assumptions for each scenario is in **Table 13**.

Table 13. Summary of four scenarios of invitation methods and screening participation

	Scenarios			
	1. Perfect adherence	**2. Current program**	**3. Mail-out FIT**	**4. High adherence**
Invitation methods				
Advanced notification letter	No	Yes, at (45), 50, 55, 60 and (65) years old[1]	Yes. FIT at the year of screening; COL at (50), 55, 60 and (65) years[1]	Yes. Both FIT and COL at the year of screening
FIT included in the notification letter	No	No	Yes	Yes
Reminder letter	No	No	No	Yes
Participation rate (%)[2]				
Annual FIT	100	Men 7; Women 25	Men 25; Women 34	Men 71; Women 75
Biennial FIT	100	Men 16; Women 24	Men 25; Women 34	Men 71; Women 75
FIT-positive colonoscopy	100	64	64	83
10-yearly screening colonoscopy	100	Men 23; Women 24	Men 23; Women 24	42
Surveillance colonoscopy	100	63	63	63
References		(G-BA, 2018; Hoffmeister et al., 2017, 2019; Gruner et al., 2020; Statistische Bundesamt, 2021)	(Hoffmeister et al., 2017, 2019; Gruner et al., 2020; Toes-Zoutendijk et al., 2020)	(Singal et al., 2016, 2017; Hoffmeister et al., 2019; Toes-Zoutendijk et al., 2020)

Note: COL = colonoscopy; FIT = fecal immunochemical test.

(1) Notification was assumed not to be sent out before the starting age or after completing all screening rounds.

(2) When only one value is shown, it is used for both men and women.

2.4.4 Screening follow-up and surveillance management

For FIT screening participants, if test remains negative, individuals will continue to follow the screening schedule set by the screening strategy. If FIT returns a positive result, the individual will be referred to a FIT-positive colonoscopy. Lesions will be immediately removed if detected by the follow-up colonoscopy, and the individual will receive a further surveillance colonoscopy with an interval defined by the German S3 guidelines for CRC (described below). On the other hand, if the FIT-positive colonoscopy returns negative findings, according to the screening guideline (Gemeinsamer Bundesausschuss (G-BA), 2018), the individual will return to FIT screening in 10 years.

For screening colonoscopy participants, lesions will be removed upon detection, and individuals will then follow the surveillance colonoscopy schedule. Negative findings in the screening colonoscopy will allow participants stay in the screening colonoscopy schedule until the maximum number of screening colonoscopies is reached.

DECAS follows the surveillance colonoscopy intervals recommended by the German S3 guidelines for CRC (**Table 14**) (GGPO, 2019). For 1-2 small tubular adenomas and greater than 5 adenomas, DECAS takes the middle value of the recommended interval, namely, 7.5 years and 1.5 years, respectively. Surveillance colonoscopy was assumed to stop at age of 75 (Greuter et al., 2017; Lew et al., 2018). The adherence rate for surveillance colonoscopy were assumed to be 63%, based on a German cohort study in Saarland (Hoffmeister et al., 2019).

Table 14. Assumptions of surveillance colonoscopy intervals after lesion removal used in the DECAS screening model

Colonoscopy findings	DECAS surveillance interval
1-2 small tubular adenomas (<10 mm)	7.5 years[1]
3-4 adenomas OR ≥1 adenoma ≥10 mm or villous or high-grade neoplasia	3 years
≥ 5 adenomas	1.5 year[2]
Serrated lesions	Same as adenomas[3]

Note:

(1) German S3 guidelines recommends 5-10 years – a mid-value was used.

(2) German S3 guidelines recommends <3 years – a mid-value was used.

(3) When counting the lesions in order to follow the recommended surveillance interval, adenomas and serrated polyps were counted together in DECAS.

2.4.5 Model input parameters

Model input parameters discussed below and the values used for sensitivity analyses are summarized in **Table 15**.

2.4.5.1 Screening test characteristics assumptions

In the literature, colonoscopy performance is usually reported on a per-lesion basis, whereas FIT sensitivity is usually per-person (Knudsen et al., 2020). DECAS operates the lesion detection based on the per-person sensitivity for both screening tests, therefore, the per-lesion detection probability in colonoscopy was treated as per-person in the analysis. When multiple lesions co-exist in an individual, how DECAS determine the test sensitivity was described in Chapter 2.4.1.3.

Since colonoscopy is treated as the gold standard test for detecting and diagnosing polyps and CRC, the test sensitivities can only be inferred by studies evaluating colonoscopy miss rates. The values for colonoscopy sensitivities were based on two meta-analyses, which included miss rates for adenomas, serrated lesions and CRC (Pickhardt et al., 2011; Zhao et al., 2019). Concerning FIT

test characteristics, values from a recent meta-analysis focusing on OC-Sensor (Eiken Chemical, Tokyo, Japan) with 1 stool sample and a cut-off at 100 ng hemoglobin/mL of buffer (which corresponds to 20 μg hemoglobin/g of stool) were used (Imperiale et al., 2019).

2.4.5.2 Colonoscopy complication assumptions

Complications due to colonoscopy considered in the modeling analysis focused on major bleeding and perforation requiring hospitalization. Pox et al. (2012) used the data from German screening colonoscopy program between 2005 and 2008 and found that the complication rate was 5.8 per 10,000 colonoscopies, which was close to a more recent figure at 4 per 10,000 from the program's 2018 annual report (Kretschmann et al., 2020). Stock et al. (2013) took a different approach and analyzed a large German health insurance dataset in 2001-2008, and they estimated the colonoscopy complication rate including major bleeding and perforation at 17 per 10,000 (95% CI 11-24). Zwink et al. (2017) reported a similar rate at 17 per 10,000 from a cohort study evaluating screening colonoscopy in Saarland, Germany in 2010-2013. However, in a recent meta-analysis reviewing international data (Jodal et al., 2019), the complication rate of major bleeding and perforation was estimated to at a lower value between 1 and 3 per 10,000; similarly, the lower range of both complication combined was estimated to be 3.3 per 10,000 (95% CI 2.2 to 4.3) in the systematic review conducted by the USPSTF (Lin et al., 2021). Therefore, in this analysis, complication rate of 17 per 10,000 colonoscopies was assumed to be the base value, and range for sensitivity was set to be between 2 and 24 per 10,000 colonoscopies. No complications resulting in mortality was assumed (Kretschmann et al., 2020).

2.4.5.3 Cost assumptions

For all screening strategies, a screening consultation fee was applied. In Scenario 1, FIT screening costs included the test kit and lab costs, while screening and surveillance colonoscopies accounted for the procedure cost as well as polypectomy and pathology costs if lesions were found. In Scenario 2, postage for the four or five invitation letters was considered. As for scenario 3, the costs involving colonoscopies remained the same, whereas additional posting and pre-paid return costs for mail-out FIT test kit and the result letters (together with invitation to FIT-positive colonoscopy if applicable) to participants and general practitioners were counted. Lastly, postage costs for advanced notification letter prior to mailed-out FIT kit and reminder letters for both FIT and colonoscopy screening participants were added to Scenario 4, on top of those in Scenario 3. The costs for treating colonoscopy complications, including major bleeding and perforation requiring hospitalization, were informed by the German diagnosis-related group (G-DRG) code G48 (InEK, 2021). All costs related to screening were assumed to be one-off costs.

The cancer treatment costs used in this analysis were taken from the annual colon cancer treatment costs by cancer severity and by phase (initial, continuing, and terminal phase) (Cheng et al., 2021), as described in Chapter 2.2.2.1. Low cancer severity was assumed to approximate UICC stage 1 and 2, moderate as stage 3, and advanced stage 4. Initial and terminal phases in this analysis were defined as the first year after lesions entering clinical cancer state and the last year before an individual dying from CRC, respectively. The period in between was defined as continuing phase. The cancer treatment costs in each phase were calculated by multiplying the annual costs and the length (in year) of each phase, and the costs for all applicable phases were summed up. All costs were inflated using the Health Consumer Price Index for Germany to 2021 Euro (Statistisches Bundesamt, 2021).

2.4.5.4 Quality of life assumptions

DECAS reports the quality-adjusted life year (QALYs) as one of its outcomes, which is the product of a quality of life value in one's health state (utility) and the years the person stays in that health status.

Utility used to calculate QALYs was assumed to be different before and after the clinical cancer state, and it slightly varies in the clinical cancer state by cancer stage and phase (the same three phases as described in treatment cost assumptions). The utility values were taken from a health-related quality of life study on a CRC patient cohort in Finland. The study used several questionnaires, including EQ-5D, to survey patients with local or advanced CRC in their primary treatment, rehabilitation, remission or palliative period (Färkkilä et al., 2013). The base-line utility for individuals at any of the states before clinical cancer in DECAS was assumed to be 0.85, a value for the patients in remission period. Utility values for initial and terminal phases of any cancer stages in DECAS were assumed to be 0.76 and 0.64, equating that of primary treatment and palliative periods, respectively. The DECAS CRC stage 1-3 and stage 4 patients in the continuing phase were given the value of 0.84 and 0,82, which are the values of patients with local disease at rehabilitation and those with metastatic disease in the Finnish study (Färkkilä et al., 2013).

DECAS also considers utility loss from screening, including discomfort and complications caused by screening colonoscopy and anxiety during the waiting time for screening test results (including FIT and biopsy after polypectomy). The utility loss was assumed to be one-off which only occurs when specific screening procedure takes place, and the sum of lifetime utility loss is deducted from the lifetime QALYs. The values used were taken from the same values used by the USPSTF modeling study (Knudsen et al., 2021).

Table 15. Summary of model inputs and values for probabilistic sensitivity analysis for the cost-effectiveness analysis of German colorectal cancer screening program

	Input value		PSA		
	Mean	**95% CI**	**Distribution**	**Range**	**Reference**
Screening test performance					
FIT					
Sensitivity					
Non-AA	0.08	0.07-0.09	Uniform	0.07-0.09	
AA	0.26	0.2-0.32	Uniform	0.2-0.32	(Imperiale et al.,
Non-crSP	0.07	0.03-0.15	Uniform	0.03-0.1	2014, 2019; Chang
crSP	0.11	0.04-0.25	Uniform	0.04-0.18	et al., 2017)
Cancer	0.77	0.66-0.85	Uniform	0.66-0.85	
Specificity	0.95	0.92-0.96	Uniform	0.92-0.96	
Colonoscopy					
Sensitivity					
Non-AA	0.76	0.7-0.77	Uniform	0.7-0.77	
AA	0.91	0.84-0.96	Uniform	0.84-0.96	(Zhao et al., 2019;
Non-crSP	0.73	0.6-0.84	Uniform	0.6-0.84	Knudsen et al.,
crSP	0.76	0.63-0.87	Uniform	0.63-0.87	2020)
Cancer	0.95	0.86-1	Uniform	0.86-1	
Specificity	1	--	--	--	
Screening complications					
Major bleeding & perforation from colonoscopy	0.0004	--	Uniform	0.0002-0.0024	(Stock et al., 2013; Lin et al., 2020; Kretschmann et al., 2020)
Utility					
Baseline	0.85	0.83-0.88	Uniform	0.83-0.88	
CRC stage 1-4, initial phase	0.76	0.7-0.82	Uniform	0.7-0.82	
CRC stage 1-3, continuing phase	0.84	0.78-0.88	Uniform	0.78-0.88	(Färkkilä et al., 2013)
CRC stage 4, continuing phase	0.82	0.78-0.86	Uniform	0.78-0.86	
CRC stage 1-4, terminal phase	0.64	0.55-0.75	Uniform	0.55-0.75	
Utility loss (per event)					
Due to colonoscopy itself	0.0005	--	Uniform	0.0004-0.0006	
Due to waiting for FIT results	0.0013	--	Uniform	0.0010-0.0016	(Knudsen et al.,
Due to waiting for polypectomy results	0.0009	--	Uniform	0.0007-0.0011	2020)
Due to colonoscopy complications	0.0055	--	Uniform	0.0044-0.0066	
Costs (2021 Euro)					
Screening related					
Posting notification/reminders	€ 0.80	--	--	--	Assumption
Posting FIT	€ 3.79	--	--	--	Assumption
Screening consultation (one-off)	€ 12.90	--	--	--	
FIT kit	€ 8.34	--	--	--	
FIT process & analysis	€ 6.34	--	--	--	(KVB, 2021)[1]
Colonoscopy	€ 196.35	--	--	--	
Colonoscopy + polypectomy	€ 225.16	--	--	--	
Pathology test	€ 14.57	--	--	--	
Treatment for colonoscopy complications	€ 4973	--	Uniform	4,923-5,098	(InEK, 2021)[2]

(Continue Table 15)

	Input value		PSA		
	Mean	**95% CI**	**Distribution**	**Range**	**Reference**
Treatment for CRC					
Stage 1 & 2					
Initial phase	€ 15,965	13,883 – 18,046	Uniform	13,883 – 18,046	
Continuing phase	-€ 968	-1,215 – -620	Uniform	-1,215 – -620	
Terminal phase	€ 29,826	22,514 – 37,139	Uniform	22,514 – 37,139	
Stage 3					
Initial phase	€ 36,634	33,367 – 39,900	Uniform	33,367 – 39,900	
Continuing phase	€ 1,960	883 – 3,036	Uniform	883 – 3,036	(Cheng et al., 2021)
Terminal phase	€ 23,342	18,968 – 27,715	Uniform	18,968 – 27,715	
Stage4					
Initial phase	€ 61,742	55,969 – 67,514	Uniform	55,969 – 67,514	
Continuing phase	€ 14,099	11,583 – 16,617	Uniform	11,583 – 16,617	
Terminal phase	€ 32,903	27,981 – 37,826	Uniform	27,981 – 37,826	

Note: AA = advanced adenoma; CI = confidence interval; CRC = colorectal cancer; crSP = clinically relevant serrated polyps; FIT = fecal immunochemical test; PSA = probabilistic sensitivity analysis

(1) Einheitlicher Bewertungsmasstab (EBM) codes: 01737, 01738, 01740-10743.

(2) G-DRG code: G48.

2.4.6 Model outcomes

Modeled screening outcomes are categorized into benefits, burden and harms and are generated for each strategy. Benefits include reduction in CRC incidence and mortality rates, total life-years gained (LYG) and quality-adjusted life-years gained (QALYG), all compared to no screening. Total costs and number of screening tests used (including FIT and colonoscopies) represents burden. For Scenario 3 and 4, where mail-out FIT test kits were involved, the number and costs of unused FIT kits were also recorded. Harms are measured by the number of complications from screening. To note, this analysis only counted screening colonoscopies, follow-up colonoscopies after positive-FIT and surveillance colonoscopies, whereas diagnostic colonoscopies to confirm CRC were not considered.

In alignment with the most recent USPSTF modeling evaluation (Knudsen et al., 2021), all outcomes are presented for a cohort of 40-year-olds. Namely, outcomes of the present study were accumulated over the lifetime from the age of 40 and expressed as per 1,000 40-year-old persons. LYs, QALYs and costs were discounted from the age of 40, with a base-case annual rate of 3% as recommended by the General Methods of German Institute for Quality and Efficiency in Health Care (*Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen*, IQWiG) (2020).

2.4.6.1 Cost-effectiveness and burden-benefits analysis

To illustrate the relationship between benefits and burdens among various strategies, two measurements were used in the present analysis. The first one is the conventional cost-effectiveness, measured by costs and benefits (LYs and QALYs). The other one considers number of colonoscopies needed as burdens and number of CRC death averted as benefits (hereafter, "burden-benefits analysis"). Given population-based CRC screening requires a significant resource commitment, especially colonoscopies, understanding what could be the most efficient strategy in terms of the number-needed-to-colonoscope (NNC) to avert one CRC death is crucial. This concept has been used as one of the assessment outcomes for efficient CRC screening strategies by Lew et al. (2018) and similarly in the most recent USPSTF modeling report (Knudsen et al., 2021).

For both cost-effectiveness and burden-benefits analysis of all screening alternatives, efficiency frontiers are first plotted by placing the outcomes of all alternatives on a two-dimensional plane, where x-axis represents costs or burdens and y-axis benefits (**Figure 7**) (Institute for Quality and Efficiency in Health Care (IQWiG), 2020). The alternatives which are with similar benefits but with lower costs (located at the upper left position of the reference) are deemed as "dominant" (A, B, C and D in **Figure 7**). By connecting all dominant alternatives, an efficiency frontier can be drawn (the line connecting A, B, C and D in **Figure 7**). Alternatives located on the frontier can be viewed as efficient, while everything located lower right to the frontier is regarded as dominated or less efficient (Institute for Quality and Efficiency in Health Care (IQWiG), 2020; Knudsen et al., 2020).

Among the dominant alternatives on the efficiency frontier, incremental cost-effectiveness ratios (ICERs) and incremental NNC (INNC) ratios can be calculated. ICER is calculated using the additional costs divided by additional LYG or QALYG between the more beneficial alternative and the next most beneficial alternative (Drummond et al., 2015). INNC ratios take the additional colonoscopies needed as numerator and additional number CRC death averted as denominator (Lew et al., 2018). ICER and INNC help illustrate how much resource is devoted to have one LY/QALY or CRC death saved (i.e., how much buck for a bang). In relation to the efficiency frontier, the ICER and INNC are the inverse of the slope on the efficiency frontier (**Figure 7**).

2.4.6.2 Perspective

The perspective of the present analysis was taken from a pure SHI perspective, where only reimbursement costs and transferrable costs (not applicable in the present analysis) were considered, according to the IQWiG methods for health economics evaluation (Institute for Quality and Efficiency in Health Care (IQWiG), 2020).

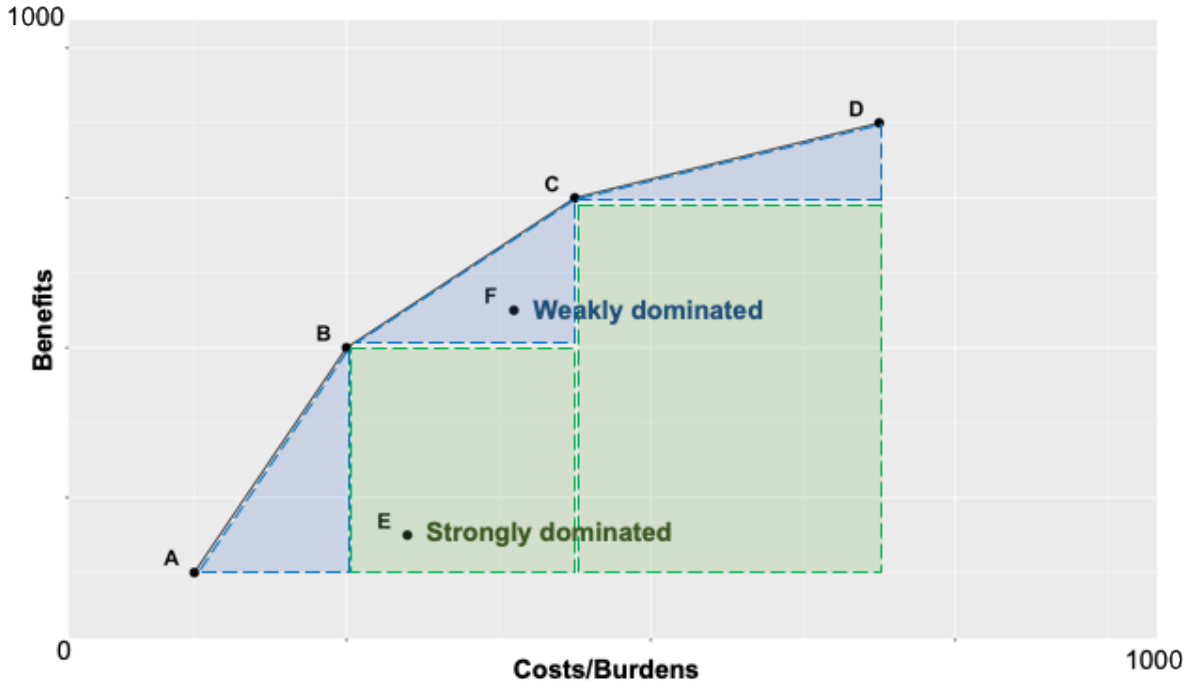


Figure 7. Illustration of efficiency frontier.
A is the reference strategy, which results in lowest benefits but also requires the least resources. A, B, C and D are located on the efficiency frontier and deemed efficient. Any alternatives located in the green rectangles are strongly dominated (e.g., E is strongly dominated by B as E provides less benefits but costs more), and those located in blue triangles are weekly dominated (e.g., F is weekly dominated by B and C – B and C do not provide more benefits and cost less at the same time compared to F). **Reference:** (Institute for Quality and Efficiency in Health Care (IQWiG), 2020; Knudsen et al., 2020)

2.4.7 Base-case and sensitivity analyses

Given the nature that DECAS takes the 1,000 sets of posterior parameters from the Bayesian calibration to simulate each strategy across scenarios, the outputs already resemble probabilistic sensitivity analyses (PSA) for that each DECAS natural history model parameter is drawn from a distribution. Therefore, to complete the PSA, ranges were specified for the rest of the model inputs (including test characteristics, complication rates, treatment costs and utility values), and random numbers within each range were drawn 1,000 times following a uniform distribution. Screening costs were the only inputs that were not altered. The ranges used for PSA were summarized in **Table 15**.

Base-case simulated outputs for each strategy were presented with as means of the 1,000 simulated outputs and 95% credible intervals (CrI), and the means were used for cost-

effectiveness analysis and burden-benefit analysis to derive the frontiers. Base-case LYs, QALYs, and costs were discounted at 3% annual rates from age 40 years (Basu and Ganiats, 2016).

To further evaluate the impact of parameter uncertainties on the cost-effectiveness and burden-benefit ratio, all the 1,000 simulated outputs were used to plot cost-effectiveness acceptance curves (CEAC), which depicts the probabilities of a certain alternative being the most cost-effective at various WTP thresholds. To derive CEACs, net health benefit (NHB) method was adopted to transform the outcomes in the unit of health benefit (i.e., LYs or QALYs) and to facilitate comparison (Stinnett and Mullahy, 1998; Institute for Quality and Efficiency in Health Care (IQWiG), 2020). NHB approach started with specifying various WTP values per LYG or QALYG. In the present thesis, WTP €0 to €100,000 per LYG or QALYG were applied. Then, incremental costs of each strategy were divided by the WTP thresholds, followed by subtracting the derived values from the LYG or QALYG of each strategy. By doing so, one can compare NHBs across different strategies given different WTP thresholds, in which the one with the highest NHB can be deemed as the most cost-effective (Drummond et al., 2015).

In addition, all scenarios and strategies were independently assessed with different annual discount rates of 0% and 5%, as recommended by IQWiG (2020).

3 RESULTS

3.1 Net colon cancer treatment costs

(Part of Chapter 3.1 has been published (Cheng et al., 2021)*)*

In the net colon cancer treatment cost analysis, there were 21,853, 20,035 and 18,036 matched controls for initial, continuing and terminal phases, respectively. The characteristics of the patient and control populations (including age, sex, cancer severity, and number of comorbidities) are presented in **Table 16**.

Table 16. Characteristics of the cost analysis study population

	Patients (*N*=4,438)			**Controls (*N*=44,378)**		
	Initial (*N*=3,552)	**Continuing (*N*=2,850)**	**Terminal (*N*=1,747)**	**Initial (*N*=21,853)**	**Continuing (*N*=20,035)**	**Terminal (*N*=18,036)**
Age[1] (Median, IQR)	74 (67-80)	73 (66-79)	79 (72-85)	74 (67-80)	74 (67-80)	82 (74-88)
Female (*N*, %)[2]	1,695 (48%)	1,338 (47%)	852 (49%)	10,491 (48%)	9,619 (48%)	9,155 (51%)
Cancer Severity (*N*, %)						
Low	2,289 (64%)	1,956 (69%)	791 (45%)	-	-	-
Moderate	740 (21%)	613 (22%)	259 (15%)	-	-	-
Advanced	523 (15%)	281 (10%)	697 (40%)	-	-	-
Number of comorbidities (*N*, %)						
No comorbidity	380 (11%)	334 (12%)	88 (5%)	1,434 (7%)	1378 (7%)	0 (0%)
1-2 comorbidities	1,775 (50%)	1,443 (50%)	751 (43%)	10,124 (46%)	9,489 (47%)	8,894 (49%)
3-4 comorbidities	1,397 (39%)	1,073 (38%)	366 (52%)	10,295 (47%)	9,168 (46%)	9,142 (51%)

Note: IQR = interquartile range

(1) Age of diagnosis for initial and continuing phase, and age of death for terminal phase.

(2) Percentages might not sum up to 100% due to rounding.

Overall, the mean annualized net treatment costs were the highest in the terminal phase (€30,093; 95% CI €26,190 – €33,996), followed by the initial phase (€27,010; 95% CI €25,200 – €28,821) and continuing phase (€1,147; 95% CI €702 – €1,593) (**Table 17**).

Subgroup analyses by cancer severity are also presented in **Table 17**. Low cancer severity showed the same pattern as the overall net costs, where terminal phase bore the highest costs (€29,826;

95% CI €22,514 – €37,139). Interestingly, the continuing phase net costs in low cancer severity group were estimated to be negative (-€968; 95% CI -€1,215 to -€620), namely, the treatment costs were lower compared with the control group. On the other hand, for moderate and advanced cancer severity, highest costs occurred in the initial phase (moderate €36,634; 95% CI €33,367 – €39,900; advanced €61,742; 95% CI €55,969 – €67,514). It is worth of noting that the initial and continuing phase net costs for people with advanced colon cancers were significantly higher than those with low or moderate disease: 2-4 times higher in the initial phase costs, and the continuing phase net costs in the advanced cancer group was almost equal to the initial phase net costs in the low severity group.

Table 17. Results of the annualized net colon cancer treatment costs by phase and subgroup analysis by cancer severity (in 2021 Euro, costs discounted with 3% annual rate)

	Overall patient group	By cancer severity		
		Low	Moderate	Advanced
Annualized net treatment costs (mean, 95% CI)				
Initial	€ 27,010 (25,200 – 28,821)	€ 15,965 (13,883 – 18,046)	€ 36,634 (33,367 – 39,900)	€ 61,742 (55,969 – 67,514)
Continuing	€ 1,147 (702 – 1,593)	- € 968 (-1,215 – -620)	€ 1,960 (883 – 3,036)	€ 14,099 (11,583 – 16,617)
Terminal	€ 30,093 (26,190 – 33,996)	€ 29,826 (22,514 – 37,139)	€ 23,342 (18,968 – 27,715)	€ 32,903 (27,981 – 37,826)

Note: CI = confidence interval

3.2 DECAS natural history model calibration and validation results

(Part of Chapter 3.2 has been submitted and currently under review (Cheng et al., 2022)*)*

3.2.1 Calibration results

Following the calibration algorithm described in Chapter 2.2.4, after the pilot run, the APMC algorithm took 11 cycles to reach the threshold $p_{acc_{min}} \leq 0.05$ and stopped. Eventually, the final 1,000 parameter sets from the $\alpha = 0.1$ portion of 10,000 simulations in the last cycle were obtained as the posterior parameter samples. The calibration took approximately 10 days.

In general, the APMC algorithm converged the crucial parameters determining the transition rates well (e.g., baseline risks, progression to advanced stage pre-cancer lesions, and the base risk for progressing to cancer), as can be seen when comparing parameter distributions between the first cycle of APMC and the final cycle (**Figure 8**). **Table 18** shows the summary statistics of the posterior samples of each parameter. DECAS outputs derived from the final posterior parameter

samples were also plotted against the calibration targets together with the 95% credible intervals (CrI) (**Figure 2 to 4**). Generally, the ranges of 95% CrIs of the DECAS simulated outputs captured all calibration targets, however, the trend of prevalence along the age did not match perfectly in the prevalence of non-AA, female non-crSP, and cr-SP prevalence for age >70 years.

Table 18. Summary of the posterior estimates of DECAS parameters

Model parameters	Notation	Prior distribution	Posterior estimates	
			Mean	95% CrI
Precancerous lesion initiation (L0 to A1 or S1)				
Adenoma				
Baseline log-risk, mean	aind	~ U(-9, -4.6)	-7.240	(-8.899, -4.970)
Baseline log-risk, standard deviation	ind_sd	~ U(0.8, 5.4)	2.723	(1.693, 4.096)
Sex effect	ar_sex	~ U(-0.65, 0)	-0.373	(-0.628, -0.078)
Age effect, 20 ≤ age < 50 years	ar_2050	~ U(-0.06, 0.1)	0.023	(-0.035, 0.072)
Age effect, 50 ≤ age < 70 years	ar_5070	~ U(-0.1, 0.15)	0.035	(-0.083, 0.137)
Age effect, age ≥ 70 years	ar_70	~ U(-0.1, 0.2)	0.036	(-0.091, 0.181)
Serrated polyp				
Baseline log-risk, mean	sind	~ U(-9.8, -5.4)	-8.648	(-9.744, -6.696)
Baseline log-risk, standard deviation	ind_sd	Same as in adenoma		
Sex effect	sr_sex	~ U(-0.65, 0.25)	-0.265	(-0.606, 0.142)
Age effect, 20 ≤ age < 50 years	sr_2050	~ U(-0.12, 0.06)	0.018	(-0.033, 0.055)
Age effect, 50 ≤ age < 70 years	sr_5070	~ U(-0.12, 0.15)	-0.005	(-0.111, 0.118)
Age effect, age ≥ 70 years	sr_70	~ U(-0.12, 0.2)	0.026	(-0.112, 0.185)
Progression to the advanced stage of precancerous lesion (A1 to A2 or S1 to S2)				
Hazard of non-AA progressing to AA	r_aa	~ U(0.002, 0.3)	0.004	(0.002, 0.012)
Hazard of non-crSP progressing to crSP	r_crsp	~ U(0.002, 0.6)	0.014	(0.005, 0.040)
Progression to pre-clinical cancer (A2 or S2 to PC)				
Adenoma				
Base risk of colonic lesion progressing to pre-clinical cancer, male at age 20 years	r_aca	~ U(0.002, 0.3)	0.005	(0.002, 0.014)
Location effect, rectum	rr_ar	~ U(2, 30)	5.871	(2.192, 23.252)
Age effect, 50 ≤ age < 70 years	rr_a50	~ U(1, 5)	1.923	(1.040, 4.384)
Age effect, age ≥ 70 years	rr_a70	~ U(1.2, 10)	4.004	(1.439, 9.395)
Serrated polyp				
Base risk of colonic lesion progressing to pre-clinical cancer, male at age 20 years	r_sca	~ U(0.002, 0.6)	0.004	(0.002, 0.008)
Location effect, rectum	rr_sr	~ U(4, 50)	18.897	(6.133, 47.341)
Age effect, 50 ≤ age < 70 years	rr_s50	~ U(1, 5)	1.651	(1.022, 3.624)
Age effect, age ≥ 70 years	rr_s70	~ U(1.2, 10)	3.761	(1.404, 9.123)

Note: CrI = credible interval; ~ U(a, b) denotes the uniform distribution bounded by (a, b); L0: no lesions; A1: non-advanced adenoma; S1: non-clinically relevant serrated polyp; A2: advanced adenoma; S2: clinically relevant serrated polyp; PC: pre-clinical cancer

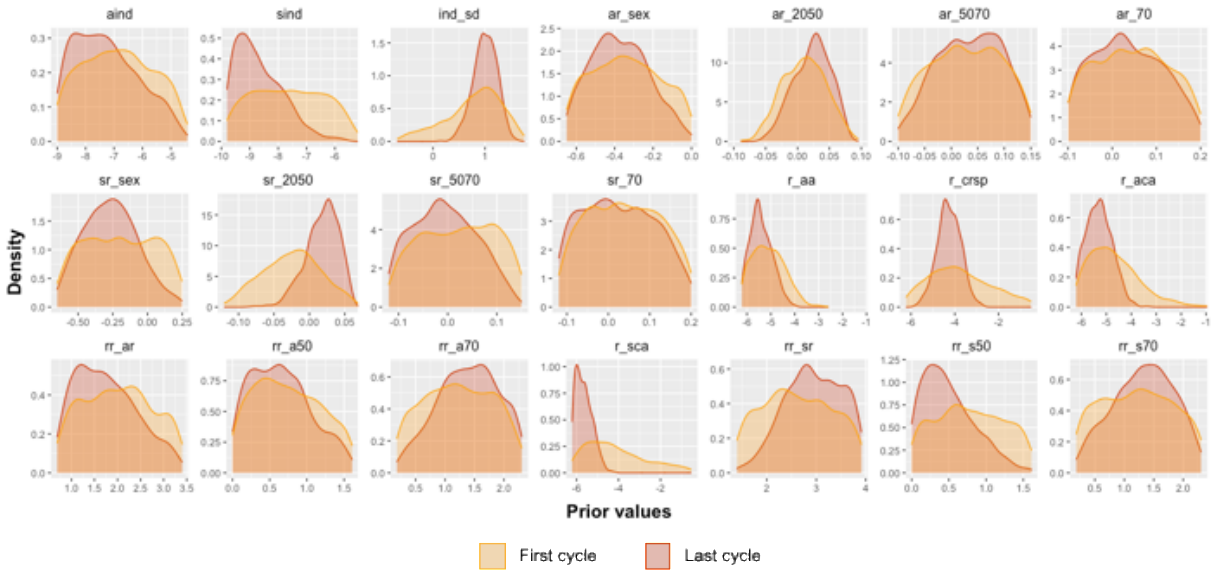


Figure 8. Density plots of posterior parameter sample distributions from the first cycle and final cycle of APMC calibration algorithm
Notations are listed in Table 18.

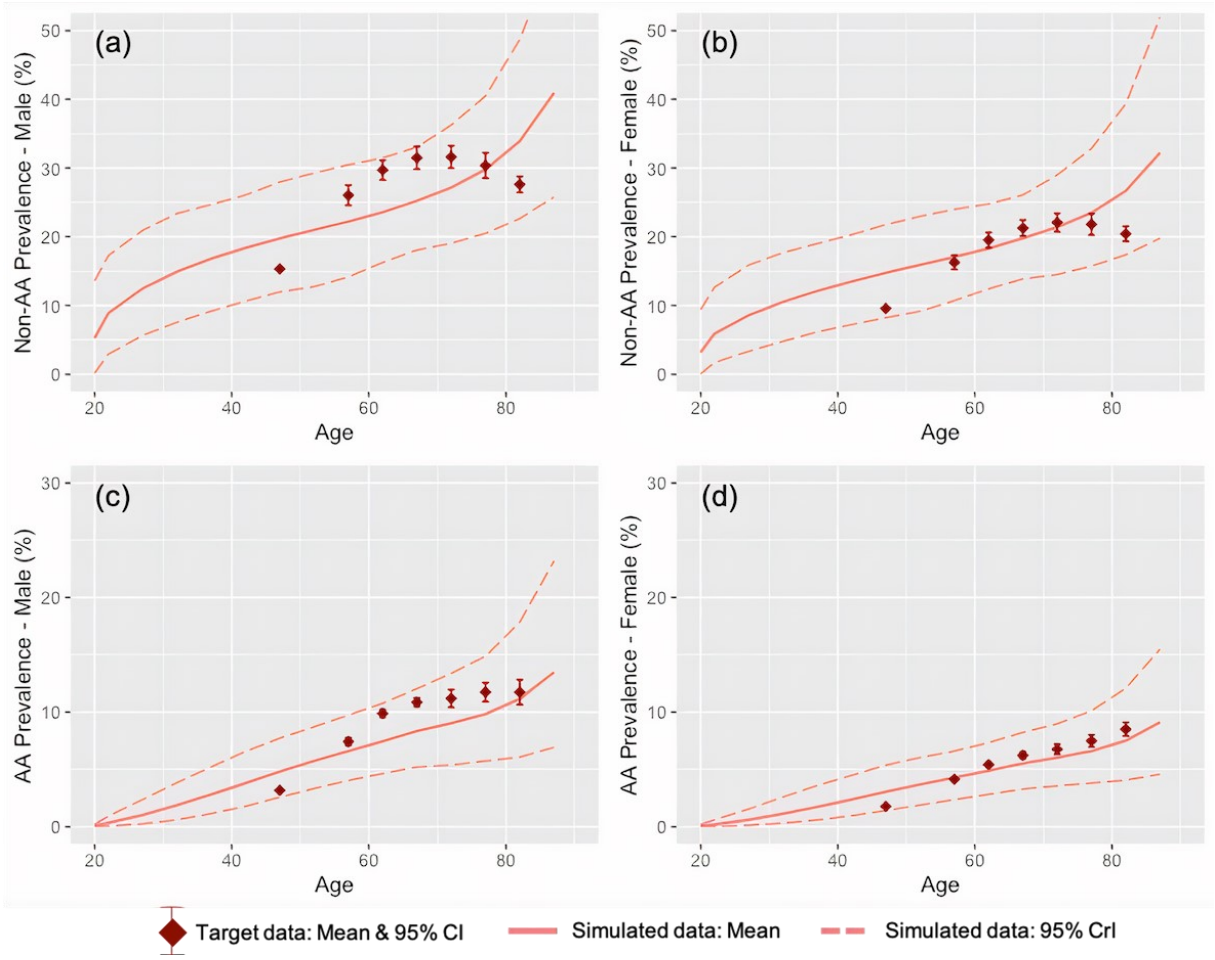


Figure 9. DECAS predicted prevalence for (a) non-advanced adenoma in men, (b) non-advanced adenoma in women, (c) advanced adenoma in men, and (d) advanced adenoma in women
AA = advanced adenoma; CI = confidence interval; CrI = credible interval.

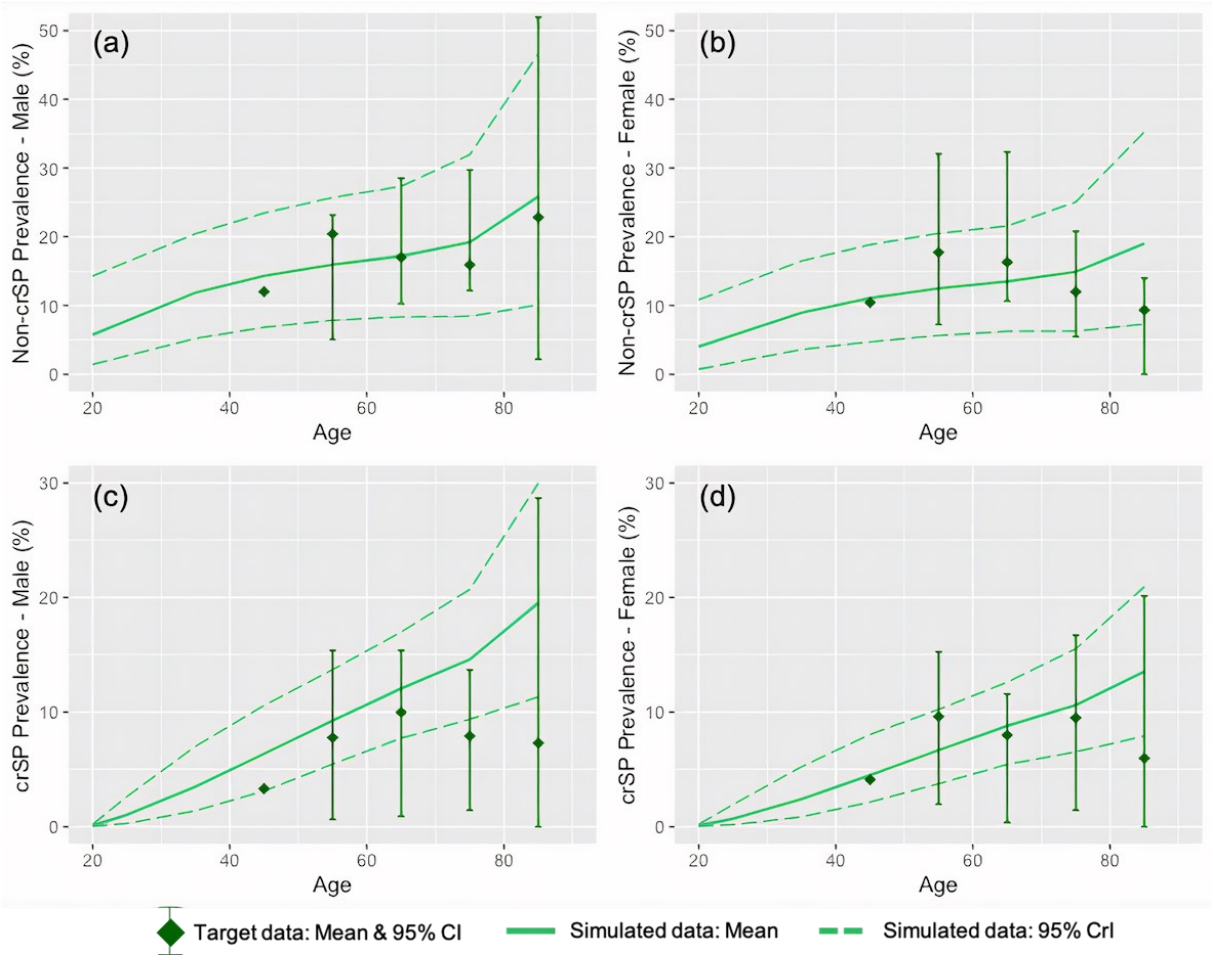


Figure 10. DECAS predicted prevalence for (a) non-crSP in men, (b) non-crSP adenoma in women, (c) crSP in men, and (d) crSP in women
CI = confidence interval; CrI = credible interval; crSP = clinically relevant serrated polyps.

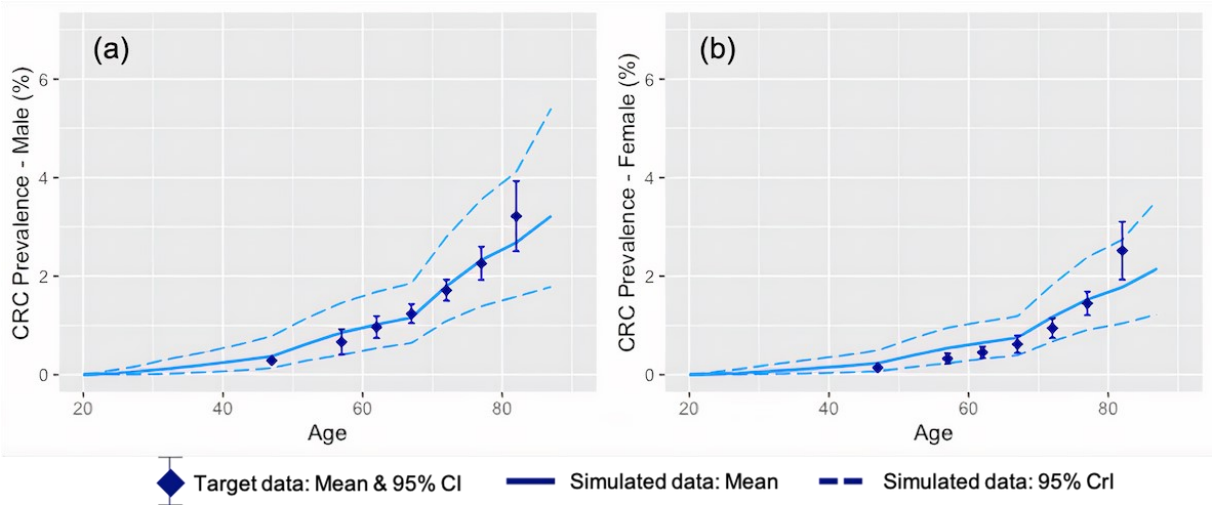


Figure 11. DECAS predicted CRC prevalence (a) in men and (b) in women
CI = confidence interval; CrI = credible interval.

3.2.2 Validation results

DECAS obtained face validity of the model structure, parameters and data sources by the evaluation from a gastroenterologist experienced in colonoscopy and an epidemiologist specialized in CRC screening. Internal validation of DECAS was executed by an experienced modeler within the Division of Health Economics at the DKFZ, who did not have direct involvement in this project. Systematic approach was taken to perform step-wise parameters alteration and observe the corresponding output change to reassure that DECAS produced reasonable outputs consistently. As for the external validation, DECAS simulated the age- and sex-specific CRC incidence to compare with ZfKD data (Robert Koch Institute, 2019) by using a population size of 30,000 and the final posterior parameter samples. The graphic examination demonstrated the predictive power of DECAS to capture the CRC incidence from another independent source of data (**Figure 12**).

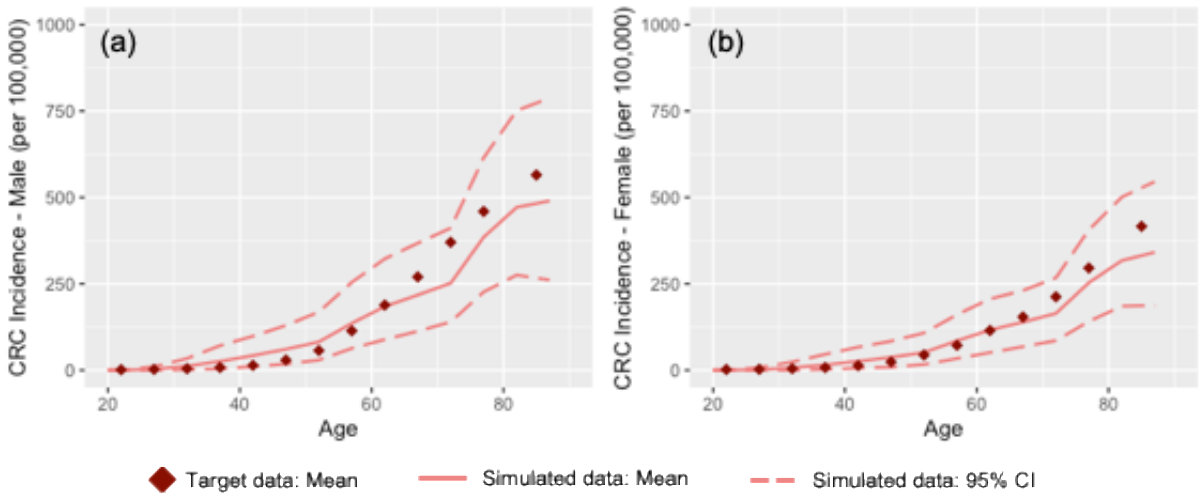


Figure 12. DECAS predicted CRC incidence (a) in men and (b) in women
CI = confidence interval.

3.2.3 Dwell time and sojourn time

The average adenoma and serrated lesion dwell time in DECAS were 38.6 years (interquartile range (IQR), 28.4-49.7) and 37.2 years (IQR, 26.6-48.5), respectively. The mean sojourn time, as expected from the pre-defined Weibull distribution, was 4.7 years (IQR, 4-5.4).

3.3 DECAS screening model validation results

(Part of Chapter 3.3 has been submitted and currently under review (Cheng et al., 2022)*)*

3.3.1 Validation results with UK Flexible Sigmoidoscopy Screening trial

DECAS predicted the primary outputs, the 17-year HRs for CRC incidence and mortality rates between the screening group and control group, accurately (**Table 19 and Figure 13**), with the mean HR and 95% CIs almost completely overlap with the estimates from the 17-year results from the UKFSS trial. The predictions for HRs of proximal and distal CRC incidence were precise as well (**Table 19**).

Comparing with the CISNET models' validation results against UKFSS trial, DECAS and SimCRC predictions were the more accurate and stable than CRC-SPIN 2.0 and MISCAN, if considering overall accuracy of HRs in overall CRC incidence, proximal CRC incidence, distal CRC incidence, and CRC mortality. In terms of HR of proximal cancer incidence, SimCRC appeared to have an even more precise prediction than DECAS. Interestingly, the deviation in HR predictions of proximal and distal cancer incidence from the UKFSS data seen in DECAS and SimCRC went into different directions – DECAS slightly overestimated whereas SimCRC tended to underestimate.

Although absolute rates were not defined as outputs for the accuracy evaluation, results from DECAS were still recorded to provide more insights to understand the difference between models. When it comes to the 17-year cumulative rates, none of the four models consistently replicate the results seen in the UKFSS trials. The DECAS prediction of the overall CRC incidence rates in the screening and control groups was the closest to the trial data, while MISCAN had the most accurate absolute mortality rate prediction. Of note, the credible intervals for the DECAS incidence predictions, both for the relative and absolute outputs, were wider than the CIs from other models (**TABLE 19**).

Table 19. Comparison of hazard ratios between screening and no-screening and 17-year rates of CRC incidence and mortality estimated by UKFSS trial, CISNET models, and DECAS

Output	Source	Hazard ratio (Mean, 95% CI/CrI)	17-year rate per 100,000 person-years	
			Control (Mean, 95% CI/CrI)	Screening (Mean, 95% CI/CrI)
CRC incidence (overall)	**UKFSS**	**0.65 (0.59-0.71)**	**184 (178-191)**	**120 (112-128)**
	CRC-SPIN 2.0	0.56 (0.52-0.77)	200 (193-207)	114 (105-123)
	SimCRC	0.66 (0.61-0.72)	212 (204-219)	143 (122-153)
	MISCAN	0.61 (0.57-0.66)	231 (223-238)	144 (134-154)
	DECAS	**0.64 (0.52-0.77)**	**183 (106-283)**	**116 (64-188)**
CRC incidence (proximal)[1]	**UKFSS**	**0.95 (0.83-1.09)**	**71 (67-75)**	**66 (60-73)**
	CRC-SPIN 2.0	0.81 (0.73-0.90)	101 (96-105)	82 (46-148)
	SimCRC	0.99 (0.88-1.12)	68 (64-72)	67 (61-74)
	MISCAN	0.67 (0.60-0.74)	113 (108-118)	75 (68-82)
	DECAS	**0.85 (0.70-1.02)**	**103 (55-170)**	**87 (46-148)**
CRC incidence (distal)[1]	**UKFSS**	**0.44 (0.38-0.50)**	**112 (107-117)**	**50 (45-56)**
	CRC-SPIN 2.0	0.33 (0.28-0.38)	99 (94-104)	32 (28-37)
	SimCRC	0.53 (0.48-0.58)	144 (138-150)	76 (69-82)
	MISCAN	0.59 (0.53-0.65)	117 (112-123)	69 (62-75)
	DECAS	**0.36 (0.27-0.47)**	**80 (42-123)**	**29 (14-53)**
CRC mortality	**UKFSS**	**0.59 (0.49-0.70)**	**56 (53-59)**	**33 (29-38)**
	CRC-SPIN 2.0	0.47 (0.40-0.54)	70 (47-138)	33 (29-38)
	SimCRC	0.60 (0.52-0.69)	73 (69-77)	44 (39-49)
	MISCAN	0.49 (0.41-0.57)	54 (50-57)	26 (22-30)
	DECAS	**0.59 (0.49-0.70)**	**86 (47-138)**	**51 (27-83)**

Notes: CI = confidence interval; CrI = credible interval; CRC = colorectal cancer.

(1) Colon segmentation in the UKFSS trial: proximal includes colon segments proximal to end of descending colon, and distal includes sigmoid and rectum.

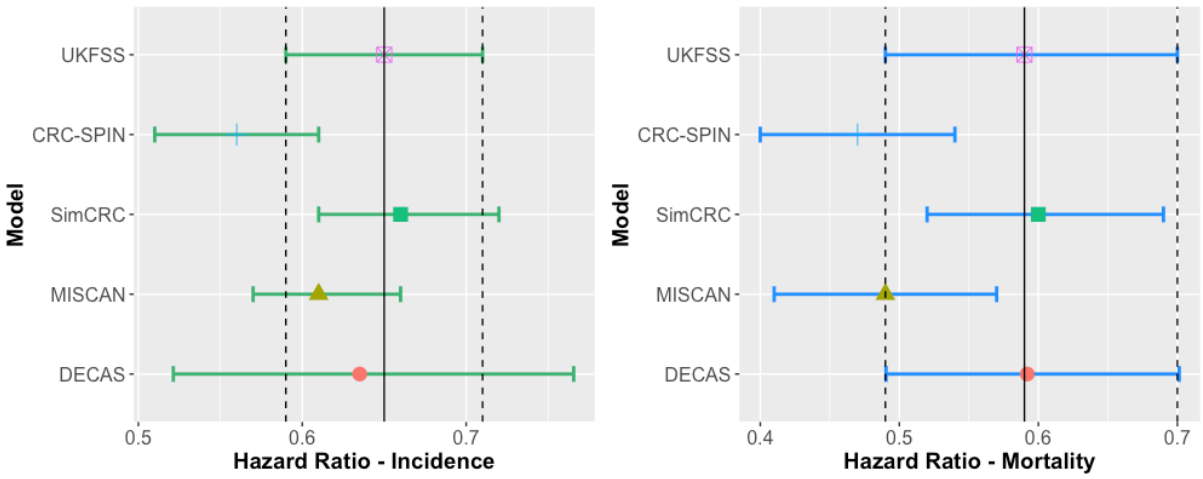


Figure 13. Hazard ratios of CRC incidence and mortality between screening and no-screening groups – comparison between the estimation of the UKFSS Trial, CISNET models, and DECAS. Point estimates (mean) and 95% confidence interval or credible interval of hazard ratios from the UKFSS Trial, DECAS, CRC-SPIN 2.0, SimCRC, and MISCAN are displayed. The vertical solid line and dashed lines signify the point estimate and 95% confidence interval of the UKFSS trial, respectively.

3.3.2 Validation results with ESTHER cohort study

DECAS predicted the overall CRC incidence and reduction well, reaching the pre-defined accuracy as DECAS prediction fell within the 95% error bounds from ESTHER study. The overall 17-year HR of CRC incidence was slightly overestimated with an absolute difference in mean HR of 0.07 (ESTHER HR 0.44, 95% CI 0.33-0.57; DECAS HR 0.37, 95% CI 0.31-0.44), while the mean mortality HR was slightly underestimated by 0.04 compared with ESTHER estimation (ESTHER HR 0.34, 95% CI 0.21-0.53; DECAS HR 0.38, 95% CI 0.32-0.45). However, the ESTHER study estimated higher incidence and mortality reduction in the distal cancers (mean incidence HR 0.36 and mortality HR 0.33) than in proximal cancers (mean incidence HR 0.69 and mortality HR 0.62), whereas the DECAS predicted similar reduction effects both in distal and proximal colons (all mean HRs around 0.38) (**Table 20 & Figure 14**).

As for the absolute rates, DECAS predicted lower 17-year incidence rates both in the screening and non-screening groups compared with ESTHER cohorts. On the other hand, DECAS predicted the 17-year mortality rates close to the estimation in the ESHTER study (**TABLE 20**).

Table 20. Comparison of hazard ratios between screening and no-screening and 17-year rates of CRC incidence and mortality estimated by ESTHER study and DECAS

Output	Source	Hazard ratio (Mean, 95% CI/CrI)	17-year rate per 100,000 person-years	
			Control (Mean, 95% CI/CrI)	Screening (Mean, 95% CI/CrI)
CRC incidence (overall)	ESTHER	0.44 (0.33-0.57)	248	122
	DECAS	0.37 (0.31-0.44)	189 (112-292)	71 (41-114)
CRC incidence (proximal)[1]	ESTHER	0.69 (0.42-1.13)	53	43
	DECAS	0.38 (0.30-0.45)	89 (49-145)	33 (18-55)
CRC incidence (distal)[1]	ESTHER	0.36 (0.25-0.51)	159	64
	DECAS	0.37 (0.30-0.45)	100 (55-152)	37 (19-63)
CRC mortality (overall)	ESTHER	0.34 (0.21-0.53)	95	37
	DECAS	0.38 (0.32-0.45)	88 (48-141)	34 (18-57)
CRC mortality (proximal)[1]	ESTHER	0.62 (0.26-1.45)	18	14
	DECAS	0.38 (0.30-0.48)	41 (20-69)	16 (8-27)
CRC mortality (distal)[1]	ESTHER	0.33 (0.19-0.59)	62	23
	DECAS	0.38 (0.30-0.47)	48 (24-76)	18 (8-32)

Notes: CI = confidence interval; CrI = credible interval; CRC = colorectal cancer.

(1) Colon segmentation: proximal includes colon segments proximal to splenic flexure, and distal includes descending colon, sigmoid and rectum.

Figure 14. Hazard ratios of CRC incidence and mortality between screening and no-screening groups – comparison between the estimation of the ESTHER study and DECAS predictions.
Point estimates and 95% confidence interval or credible interval of hazard ratios from the ESTHER study and DECAS are displayed. The vertical solid line and dashed lines signify the point estimate and 95% confidence interval of the ESTHER study, respectively.

3.3.3 Validation results with Nottingham FOBT trial

Overall, the validation with Nottingham FOBT trial also rendered accurate results by pre-defined end point, but DECAS generally predicted with slight overestimation on the intervention effects. DECAS predicted a slightly lower 20-year incidence HR by 0.04 (Nottingham incidence HR 0.94, 95% CI 0.85-1.05 vs DECAS incidence HR 0.90, 95% CI 0.81-1.00), as well as a lower 20-year mortality HR by 0.09 (Nottingham mortality HR 0.82, 95% CI 0.7-0.95 vs DECAS mortality HR 0.73, 95% CI 0.64-0.82) (**Table 21 & Figure 15**). In terms of 20-year absolute cumulative rates, the Nottingham trial only reported the mortality rates. Generally, DECAS predicted lower 20-year cumulative CRC mortality for both screening and control groups.

Table 21. Comparison of hazard ratios between screening and no-screening and 20-year rates of CRC incidence and mortality estimated by Nottingham FOBT trial and DECAS

Output	Source	Hazard ratio (Mean, 95% CI/CrI)	20-year rate per 1,000 person-years	
			Control (Mean, 95% CI/CrI)	Screening (Mean, 95% CI/CrI)
CRC incidence (overall)	Nottingham	0.94 (0.85-1.05)	--	--
	DECAS	0.90 (0.81-1.00)	2 (1.22-3.02)	1.79 (1.10-2.67)
CRC mortality (overall)	Nottingham	0.82 (0.70-0.95)	1	0.91
	DECAS	0.73 (0.64-0.82)	0.92 (0.52-1.44)	0.67 (0.38-1.03)

Notes: CI = confidence interval; CrI = credible interval; CRC = colorectal cancer; -- = not reported.

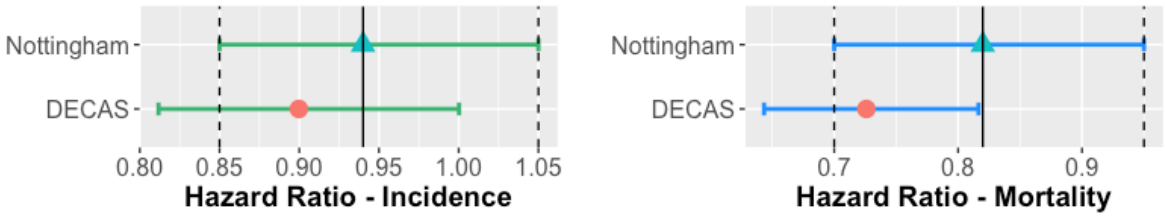


Figure 15. Hazard ratios of CRC incidence and mortality between screening and no-screening groups – comparison between the estimation of the Nottingham FOBT trial and DECAS predictions.
Point estimates and 95% confidence interval or credible interval of hazard ratios from the Nottingham FOBT trial and DECAS are displayed. The vertical solid line and dashed lines signify the point estimate and 95% confidence interval of the Nottingham FOBT trial, respectively.

3.4 Evaluation of German CRC Screening Program

3.4.1 Overall results

3.4.1.1 Benefits

The base-case results of benefits are shown in **Table 22**. When assuming 100% adherence to screening (Scenario 1), all screening strategies demonstrated significant protective effects in reducing new CRC cases and deaths. The mean HRs for CRC incidence comparing any screening strategies with no-screening ranged between 0.25 and 0.66, and those for CRC mortality varied between 0.2 and 0.48, with all 95% CrIs well under 1. Screening strategies involving colonoscopy generally demonstrated more protective effects than those with pure FIT screening. In addition, three-time colonoscopies appeared to be more effective to prevent new CRC cases and deaths in comparison with two colonoscopies. Screening with FIT was shown to provide more protection in preventing CRC mortality than CRC incidence, while strategies involving colonoscopies demonstrated comparable protective effects against new CRC cases and deaths.

In Scenario 2 and 3, where the participation rates in both FIT and colonoscopy screening were under 35% and those for FIT-positive colonoscopy as well as surveillance colonoscopy around 65%, the protective effects against CRC new cases and deaths apparently decreased. In pure FIT screening strategies, the reduction in CRC incidence even became insignificant (95% CrIs of HRs crossed 1). The mean HRs for CRC incidence and mortality of the screening strategies in these two scenarios were 0.81-0.97 and 0.7-0.84, respectively.

In the high adherence scenario (Scenario 4), participation in FIT screening (>70%) was assumed to be higher than colonoscopy screening (42%). The protective effect of FIT screening to prevent

new CRC cases remained inferior (mean HRs 0.79-0.83) to colonoscopy screening. In contrast, the protective effects in preventing CRC mortality appeared to be comparable across strategies (mean HRs 0.54-0.65). The two combined strategies starting at age 45 years old, FIT1y45+COL10y50 and FIT1y45+COL10y50-3X, stood out in this scenario, with the mean incidence HRs of 0.66 and 0.64 and the mean mortality HRs of 0.55 and 0.54, respectively.

When looking at the discounted LYs, generally the lower the HR of CRC mortality was, the higher the total discounted LYs were. For instance, COL10y45 in Scenario 1 had the lowest mean CRC mortality HR of 0.2 and the highest mean total LYs of 22,479 life-years, while FIT2y50 had the highest mean CRC mortality HR of 0.48 and the lowest mean total LYs of 22,438 life-years. Same was observed across Scenario 2 to 4: pure FIT strategies were weaker in the protective effects against CRC mortality, and they also resulted in fewer total discounted LYs.

As for discounted QALYs in each screening strategy, overall, the strategies with higher total LYs also ended up with higher QALYs. However, some exceptions were noticed. When participation rate was 100% (Scenario 1), the discounted total LYs of COL10y50 and COL10y50-3X were lower than those of the combined strategies (FIT1y45+COL10y50 and FIT1y45+COL10y50-3X). Yet, the discounted total QALYs of the former two strategies resulted in higher numbers (19,189 and 19,189 QALYs, respectively) than those of the latter two combined strategies (19,186 and 19,185, respectively). Similarly, the exception was observed in Scenario 3 among FIT1y45+COL10y50, FIT1y45+COL10y50-3X and COL10y45, where COL10y45 had lower total discounted LYs (22,419 LYs) but higher total discounted QALYs (19140 QALYs) than the other two.

3.4.1.2 Burdens and harms

Table 23 summarizes the base-case burdens and harms of all screening strategies from four scenarios, calculated for the lifetime of 1,000 persons who are alive at 40 years old. In general, strategies with shorter screening intervals, earlier starting age, and more colonoscopy entitled were associated with higher resource use (higher burdens). In the perfect screening scenario (Scenario 1), participants in the three pure FIT screening strategies in average received approximately 8-10 rounds of screening, resulting in 7,918-9,782 FIT usage. The utilization of colonoscopies ranged from 809 to 3,240 colonoscopies per 1,000 people among all strategies. Using strategy FIT2y50 as benchmark, two-colonoscopy and three-colonoscopy strategies required approximately three and four times the amount of colonoscopies, respectively. The number of complications from colonoscopies was positively correlated with the number of colonoscopies used, ranging from 1.05 to 4.21 cases per 1,000 people.

When the participation rates were no longer perfect, the number of FIT and colonoscopy used declined. In contrast to the FIT usage in Scenario 1, the number of FIT used in Scenario 2 decreased by 74-83%, Scenario 3 by 64-69%, and Scenario 4 by 20-25%. As for the colonoscopy usage, comparing to Scenario 1, the usage in Scenario 2 decreased by 77-88%, Scenario 3 by 76-83%, and Scenario 4 by 42-61%. The complication rates across Scenario 2 to 4 decreased in the same magnitude as seen in number of colonoscopies used.

In terms of costs, all screening strategies in all four scenarios resulted in lower discounted lifetime costs compared with no-screening strategy (in other words, cost-saving). Among the screening strategies, in Scenarios 1, FIT1y50+FIT2y55 had the highest average lifetime discounted costs of €830,886, and COL10y50 brought about the lowest costs of €690,682. In general, pure FIT strategies had higher costs compared with strategies involving colonoscopies, across all scenarios. This was especially true for the pure FIT screening strategies in Scenario 3, where FITs were posted to all eligible individuals, no matter they eventually participated or not. Similar pattern was observed in Scenario 4, but at a smaller scale.

In Scenario 3 and 4, since not all posted FIT kits were used and returned, the number for wasted FITs and the wasted costs were recorded. As expected, the lower the FIT participation rates were, the more wasted FITs and higher wasted costs there were. In Scenario 3, the three pure FIT screening strategies resulted in 8,006-9,967 wasted FIT test kits, which were more than three times of the number of used ones. In the combined strategies, the wasted FIT kits were fewer due to a maximum of five FIT testing rounds in total, but the wasted FIT kits also mounted to three times more of the used ones. In Scenario 4, given the participation rate was double the rate of that in Scenario 3, the number of wasted FITs significantly reduced – the wasted ones were about half the number of the used ones.

Pertaining the wasted costs due to unused FIT kits and postage, the total discounted wasted costs in Scenario 3 were the highest in FIT2y45, which summed to €93,323 (95% CI €92,058-€94,191), and the lowest in mCOL50/fFIT50+COL55, with a sum of €18,263 (95% CI €18,059-€18,437). In Scenario 4, same as the number of wasted FITs, the wasted costs also significantly decreased. FIT2y45 and mCOL50/fFIT50+COL55 remained the strategies with highest and lowest wasted FIT costs, with a sum of €35,724 (95% CI €25,164-36,218) and €6,917 (95% CI €6,799-€7,029).

Table 22. Modeled benefits of strategies per 1,000 people for a cohort of 40 years old

Strategy[1,2]	CRC incidence			CRC mortality			dLY	dQALY
	Case	Rate	Hazard ratio	Case	Rate	Hazard ratio		
No Screening	58 (37-86)	1.64 (1.03-2.46)	--	27 (17-41)	0.75 (0.47-1.16)	--	22,381 (22,255-22,466)	19,107 (18,563-19,652)
Scenario 1 (perfect adherence)								
FIT1y50+FIT2y55	38 (22-58)	1.07 (0.64-1.65)	0.66 (0.55-0.81)	13 (7-20)	0.35 (0.20-0.54)	0.46 (0.38-0.56)	22,441 (22,348-22,498)	19,146 (18,613-19,681)
FIT1y50+COL10y55	24 (13-39)	0.67 (0.37-1.11)	0.41 (0.33-0.49)	8 (5-14)	0.23 (0.13-0.40)	0.31 (0.24-0.38)	22,450 (22,361-22,507)	19,164 (18,633-19,693)
mCOL50/fFIT50+COL55	21 (11-34)	0.58 (0.31-0.98)	0.35 (0.28-0.43)	8 (4-13)	0.21 (0.11-0.36)	0.27 (0.21-0.34)	22,458 (22,374-22,517)	19,178 (18,640-19,719)
FIT2y50	40 (24-60)	1.11 (0.67-1.70)	0.68 (0.58-0.83)	13 (8-20)	0.36 (0.21-0.56)	0.48 (0.40-0.58)	22,438 (22,344-22,498)	19,143 (18,608-19,679)
COL10y50	19 (10-32)	0.53 (0.27-0.89)	0.32 (0.25-0.41)	7 (4-12)	0.19 (0.10-0.33)	0.26 (0.20-0.33)	22,465 (22,382-22,517)	19,189 (18,655-19,720)
COL10y50-3X	17 (9-30)	0.48 (0.24-0.83)	0.29 (0.22-0.37)	7 (3-11)	0.18 (0.09-0.32)	0.24 (0.18-0.31)	22,465 (22,379-22,520)	19,189 (18,648-19,721)
FIT1y45+COL10y50	20 (10-32)	0.55 (0.29-0.91)	0.33 (0.25-0.42)	7 (4-11)	0.19 (0.10-0.32)	0.25 (0.19-0.31)	22,469 (22,399-22,520)	19,186 (18,648-19,713)
FIT1y45+COL10y50-3X	19 (10-32)	0.52 (0.27-0.89)	0.31 (0.24-0.40)	6 (3-11)	0.18 (0.09-0.31)	0.23 (0.17-0.30)	22,469 (22,391-22,519)	19,185 (18,645-19,718)
FIT2y45	36 (21-55)	1.00 (0.59-1.56)	0.61 (0.50-0.77)	11 (7-18)	0.31 (0.18-0.49)	0.42 (0.34-0.52)	22,452 (22,371-22,507)	19,156 (18,617-19,688)
COL10y45-3X	15 (8-27)	0.42 (0.21-0.74)	0.25 (0.18-0.34)	6 (3-10)	0.15 (0.07-0.26)	0.20 (0.14-0.27)	22,479 (22,406-22,524)	19,205 (18,666-19,736)
Scenario 2 (current program)								
FIT1y50+FIT2y55	56 (35-82)	1.57 (0.98-2.35)	0.96 (0.92-1.00)	22 (14-34)	0.62 (0.38-0.96)	0.83 (0.78-0.88)	22,405 (22,294-22,480)	19,122 (18,580-19,666)
FIT1y50+COL10y55	49 (31-73)	1.39 (0.86-2.10)	0.85 (0.81-0.88)	20 (12-31)	0.55 (0.34-0.86)	0.74 (0.69-0.78)	22,413 (22,307-22,485)	19,132 (18,594-19,671)
mCOL50/fFIT50+COL55	49 (31-73)	1.38 (0.86-2.09)	0.84 (0.81-0.88)	20 (12-31)	0.55 (0.33-0.85)	0.73 (0.69-0.78)	22,415 (22,304-22,487)	19,135 (18,593-19,680)
FIT2y50	56 (36-83)	1.59 (1.00-2.38)	0.97 (0.94-1.01)	23 (14-35)	0.63 (0.39-0.98)	0.84 (0.79-0.89)	22,403 (22,289-22,474)	19,120 (18,576-19,659)
COL10y50	49 (31-73)	1.39 (0.87-2.08)	0.85 (0.81-0.88)	20 (12-31)	0.56 (0.33-0.86)	0.74 (0.70-0.78)	22,415 (22,308-22,486)	19,135 (18,593-19,672)
COL10y50-3X	48 (30-72)	1.36 (0.85-2.07)	0.83 (0.80-0.86)	20 (12-31)	0.55 (0.34-0.86)	0.73 (0.69-0.77)	22,416 (22,310-22,486)	19,136 (18,595-19,678)
FIT1y45+COL10y50	48 (30-72)	1.36 (0.84-2.06)	0.83 (0.80-0.87)	20 (12-30)	0.55 (0.33-0.85)	0.72 (0.68-0.77)	22,416 (22,305-22,485)	19,136 (18,593-19,675)
FIT1y45+COL10y50-3X	48 (30-71)	1.34 (0.83-2.04)	0.82 (0.79-0.86)	19 (12-30)	0.54 (0.32-0.83)	0.72 (0.67-0.76)	22,417 (22,309-22,486)	19,137 (18,597-19,679)
FIT2y45	56 (35-82)	1.57 (0.99-2.36)	0.96 (0.93-1.00)	22 (14-34)	0.62 (0.38-0.95)	0.83 (0.79-0.88)	22,405 (22,291-22,477)	19,122 (18,582-19,664)
COL10y45-3X	48 (30-72)	1.35 (0.84-2.05)	0.83 (0.79-0.86)	20 (12-30)	0.54 (0.33-0.84)	0.72 (0.68-0.77)	22,419 (22,307-22,487)	19,140 (18,595-19,680)

(Continue Table 22)

Strategy[1,2]	CRC incidence			CRC mortality			dLY	dQALY
	Case	Rate	Hazard ratio	Case	Rate	Hazard ratio		
Scenario 3 (mail-out FIT)								
FIT1y50+FIT2y55	54 (34-80)	1.53 (0.95-2.30)	0.94 (0.89-1.00)	22 (13-33)	0.60 (0.37-0.91)	0.80 (0.75-0.85)	22,407 (22,295-22,481)	19,122 (18,585-19,660)
FIT1y50+COL10y55	49 (31-72)	1.37 (0.84-2.07)	0.84 (0.80-0.87)	20 (12-30)	0.54 (0.33-0.85)	0.72 (0.68-0.77)	22,414 (22,311-22,487)	19,132 (18,592-19,668)
mCOL50/fFIT50+COL55	49 (30-73)	1.37 (0.85-2.07)	0.84 (0.80-0.88)	20 (12-31)	0.55 (0.33-0.85)	0.73 (0.69-0.77)	22,414 (22,307-22,483)	19,134 (18,597-19,668)
FIT2y50	55 (35-82)	1.56 (0.98-2.33)	0.95 (0.92-1.01)	22 (14-34)	0.61 (0.37-0.94)	0.82 (0.77-0.87)	22,406 (22,297-22,481)	19,122 (18,577-19,661)
COL10y50	49 (31-73)	1.39 (0.87-2.08)	0.85 (0.81-0.88)	20 (12-31)	0.56 (0.33-0.86)	0.74 (0.70-0.78)	22,415 (22,308-22,486)	19,135 (18,593-19,672)
COL10y50-3X	48 (30-72)	1.36 (0.85-2.07)	0.83 (0.80-0.86)	20 (12-31)	0.55 (0.34-0.86)	0.73 (0.69-0.77)	22,416 (22,310-22,486)	19,136 (18,595-19,678)
FIT1y45+COL10y50	48 (30-71)	1.34 (0.83-2.04)	0.82 (0.78-0.86)	19 (12-30)	0.53 (0.32-0.83)	0.71 (0.67-0.76)	22,419 (22,314-22,486)	19,138 (18,594-19,674)
FIT1y45+COL10y50-3X	47 (29-70)	1.32 (0.82-1.98)	0.81 (0.77-0.85)	19 (12-29)	0.53 (0.32-0.81)	0.70 (0.66-0.75)	22,419 (22,312-22,487)	19,138 (18,600-19,675)
FIT2y45	55 (35-81)	1.54 (0.96-2.31)	0.94 (0.90-0.99)	22 (14-33)	0.61 (0.37-0.93)	0.81 (0.76-0.86)	22,408 (22,298-22,486)	19,124 (18,579-19,668)
COL10y45-3X	48 (30-72)	1.35 (0.84-2.05)	0.83 (0.79-0.86)	20 (12-30)	0.54 (0.33-0.84)	0.72 (0.68-0.77)	22,419 (22,307-22,487)	19,140 (18,595-19,680)
Scenario 4 (high adherence)								
FIT1y50+FIT2y55	46 (19-69)	1.30 (0.80-1.98)	0.80 (0.72-0.91)	17 (10-26)	0.46 (0.28-0.72)	0.62 (0.55-0.70)	22,426 (22,330-22,494)	19,135 (18,594-19,672)
FIT1y50+COL10y55	40 (24-61)	1.13 (0.68-1.73)	0.69 (0.63-0.75)	16 (9-24)	0.43 (0.25-0.68)	0.57 (0.52-0.63)	22,429 (22,332-22,492)	19,144 (18,605-19,684)
mCOL50/fFIT50+COL55	41 (25-62)	1.16 (0.71-1.79)	0.71 (0.66-0.76)	16 (10-25)	0.45 (0.27-0.71)	0.60(0.56-0.65)	22,427 (22,325-22,490)	19,146 (18,604-19,685)
FIT2y50	48 (30-72)	1.36 (0.84-2.05)	0.83 (0.76-0.94)	18 (11-27)	0.49 (0.20-0.75)	0.65 (0.59-0.72)	22,421 (22,319-22,489)	19,131 (18,590-19,671)
COL10y50	42 (26-64)	1.19 (0.73-1.81)	0.73 (0.68-0.77)	17 (10-26)	0.47 (0.28-0.74)	0.63 (0.58-0.68)	22,426 (22,317-22,490)	19,148 (18,604-19,685)
COL10y50-3X	41 (25-62)	1.15 (0.70-1.76)	0.70 (0.66-0.75)	17 (10-26)	0.46 (0.28-0.73)	0.61 (0.57-0.65)	22,428 (22,328-22,493)	19,149 (18,614-19,688)
FIT1y45+COL10y50	39 (24-58)	1.08 (0.66-1.66)	0.66 (0.60-0.72)	15 (9-23)	0.41 (0.25-0.65)	0.55 (0.49-0.61)	22,439 (22,349-22,501)	19,156 (18,612-19,692)
FIT1y45+COL10y50-3X	37 (23-57)	1.05 (0.63-1.62)	0.64 (0.58-0.70)	15 (9-23)	0.40 (0.24-0.63)	0.54 (0.48-0.59)	22,439 (22,343-22,500)	19,156 (18,618-19,686)
FIT2y45	46 (29-68)	1.29 (0.80-1.96)	0.79 (0.72-0.90)	17 (10-26)	0.46 (0.28-0.71)	0.62 (0.56-0.69)	22,429 (22,332-22,492)	19,138 (18,601-19,668)
COL10y45-3X	40 (25-61)	1.13 (0.69-1.72)	0.69 (0.65-0.74)	16 (10-26)	0.45 (0.27-0.71)	0.60 (0.56-0.64)	22,433 (22,338-22,499)	19,155 (18,617-19,690)

Note: COL = colonoscopy; dLY = discounted life-years; dQALY = discounted quality-adjusted life-years; FIT = fecal immunochemical test.

(1) The life-years and quality-adjusted life-years were discounted with 3% annual rates.

(2) Results are presented as mean and 95% credible interval.

Table 23. Modeled burdens, harms and wasted resources of strategies per 1,000 people for a cohort of 40 years old

	Burdens and harms				Wasted resources	
Strategy[1,2]	**FIT**	**COL**	**COL complications**	**dCosts**	**Wasted FIT**	**Wasted dCosts**
No Screening				1,043,243 (620,397-1,716,334)		
Scenario 1 (perfect adherence)						
FIT1y50+FIT2y55	8,965 (8,277-9,688)	895 (671-1,150)	1.16 (0.22-2.31)	830,886 (545,872-1,247,989)	-	-
FIT1y50+COL10y55	3,737 (3,534-3,934)	2,250 (1,991-2,550)	2.93 (0.56-5.39)	774,071 (547,243-1,137,025)	-	-
mCOL50/fFIT50+COL55	1,967 (1,863-2,076)	2,354 (2,070-2,687)	3.06 (0.59-5.63)	725,658 (513,556-1,056,507)	-	-
FIT2y50	7,918 (7,366-8,495)	809 (609-1,043)	1.05 (0.21-2.07)	828,888 (548,539-1,259,355)	-	-
COL10y50	-	2,393 (2,084-2,745)	3.12 (0.61-5.80)	690,682 (499,794-1,006,566)	-	-
COL10y50-3X	-	2,957 (2,587-3,378)	3.85 (0.78-7.02)	720,950 (535,354-1,034,443)	-	-
FIT1y45+COL10y50	3,856 (3,645-4,065)	2,483 (2,168-2,828)	3.24 (0.62-5.99)	763,187 (572,747-1,078,880)	-	-
FIT1y45+COL10y50-3X	3,856 (3,644-4,064)	3,007 (2,644-3,411)	3.91 (0.76-7.11)	794,684 (605,654-1,111,955)	-	-
FIT2y45	9,782 (9,076-10,531)	991 (737-1,280)	1.29 (0.25-2.55)	815,225 (557,806-1,201,719)	-	-
COL10y45-3X	-	3,240 (2,825-3,700)	4.21 (0.83-7.76)	752,938 (590,021-1,027,066)	-	-
Scenario 2 (current program)						
FIT1y50+FIT2y55	2,041 (1,930-2,154)	111 (85-141)	0.15 (0.02-0.32)	910,245 (557,020-1,458,465)	-	-
FIT1y50+COL10y55	647 (616-679)	509 (468-558)	0.66 (0.12-1.26)	864,312 (536,252-1,372,459)	-	-
mCOL50/fFIT50+COL55	516 (493-540)	525 (481-575)	0.69 (0.13-1.29)	861,306 (537,401-1,372,708)	-	-
FIT2y50	1,750 (1,650-1,855)	93 (72-118)	0.12 (0.01-0.27)	912,768 (547,808-1,468,899)	-	-
COL10y50	-	517 (470-572)	0.67 (0.12-1.28)	859,372 (531,736-1,366,035)	-	-
COL10y50-3X	-	671 (613-734)	0.87 (0.15-1.65)	863,201 (539,901-1,376,369)	-	-
FIT1y45+COL10y50	664 (634-693)	549 (497-609)	0.71 (0.14-1.35)	862,509 (547,884-1,360,539)	-	-
FIT1y45+COL10y50-3X	664 (633-694)	701 (640-768)	0.91 (0.17-1.69)	867,460 (544,469-1,362,242)	-	-
FIT2y45	2,173 (2,042-2,308)	114 (86-144)	0.15 (0.02-0.31)	912,261 (563,210-1,450,670)	-	-
COL10y45-3X	-	722 (658-793)	0.94 (0.17-1.76)	871,808 (554,011-1,367,235)	-	-

(Continue Table 23)

Strategy[1,2]	Burdens and harms				Wasted resources	
	FIT	COL	COL complications	dCosts	Wasted FIT	Wasted dCosts
Scenario 3 (mail-out FIT)						
FIT1y50+FIT2y55	3,174 (3,002-3,351)	174 (133-222)	0.23 (0.03-0.47)	1,024,527 (676,168-1,564,616)	9,935 (9,781-10,036)	89,283 (88,005-90,092)
FIT1y50+COL10y55	1,167 (1,115-1,223)	535 (489-588)	0.69 (0.13-1.30)	913,442 (584,360-1,416,362)	3,373 (3,343-3,396)	37,711 (37,376-37,964)
mCOL50/fFIT50+COL55	702 (669-734)	534 (487-588)	0.69 (0.12-1.32)	886,942 (568,116-1,392,467)	1,634 (1,616-1,649)	18,263 (18,059-18,437)
FIT2y50	2,572 (2,424-2,725)	138 (106-173)	0.18 (0.02-0.37)	1,000,351 (647,455-1,538,955)	8,006 (7,878-8,094)	69,380 (68,356-70,099)
COL10y50	-	517 (470-572)	0.67 (0.12-1.28)	859,372 (531,736-1,366,035)	-	-
COL10y50-3X	-	671 (613-734)	0.87 (0.15-1.65)	863,201 (539,901-1,376,369)	-	-
FIT1y45+COL10y50	1,202 (1,146-1,258)	577 (520-641)	0.75 (0.14-1.40)	918,877 (603,329-1,410,067)	3,432 (3,408-3,450)	44,471 (44,155-44,699)
FIT1y45+COL10y50-3X	1,201 (1,148-1,260)	728 (664-802)	0.95 (0.18-1.79)	923,446 (608,362-1,413,087)	3,432 (3,408-3,450)	44,473 (44,165-44,700)
FIT2y45	3,194 (3,010-3,390)	168 (128-213)	0.22 (0.04-0.45)	1,031,382 (686,998-1,575,156)	9,967 (9,812-10,072)	93,323 (92,058-94,191)
COL10y45-3X	-	722 (658-793)	0.94 (0.17-1.76)	871,808 (554,011-1,367,235)	-	-
Scenario 4 (high adherence)						
FIT1y50+FIT2y55	7,142 (6,685-7,624)	516 (394-649)	0.67 (0.11-1.34)	988,951 (676,713-1,457,477)	3,805 (3,746-3,854)	34,192 (33,697-34,600)
FIT1y50+COL10y55	2,797 (2,658-2,937)	1,012 (915-1,123)	1.31 (0.24-2.41)	877,553 (587,727-1,339,486)	1,292 (1,274-1,307)	14,438 (14,245-14,611)
mCOL50/fFIT50+COL55	1,509 (1,436-1,587)	979 (891-1,082)	1.27 (0.26-2.38)	849,919 (559,194-1,305,894)	619 (608-629)	6,917 (6,799-7,029)
FIT2y50	6,073 (5,697-6,470)	441 (339-554)	0.58 (0.10-1.16)	970,040 (642,441-1,457,373)	3,063 (3,005-3,110)	26,548 (26,076-26,926)
COL10y50	-	924 (841-1,021)	1.20 (0.23-2.24)	819,095 (532,720-1,281,521)	-	-
COL10y50-3X	-	1,197 (1,098-1,307)	1.56 (0.29-2.87)	826,566 (535,228-1,286,515)	-	-
FIT1y45+COL10y50	2,884 (2,739-3,028)	1,090 (976-1,222)	1.41 (0.27-2.64)	879,800 (611,284-1,308,609)	1,314 (1,297-1,328)	17,026 (16,807-17,209)
FIT1y45+COL10y50-3X	2,884 (2,739-3,027)	1,351 (1,229-1,496)	1.76 (0.33-3.29)	890,098 (629,108-1,303,325)	1,314 (1,299-1,328)	17,027 (16,834-17,208)
FIT2y45	7,531 (7,060-8,032)	539 (410-679)	0.70 (0.12-1.40)	992,965 (689,958-1,459,604)	3,815 (3,750-3,872)	35,724 (25,164-36,218)
COL10y45-3X	-	1,290 (1,178-1,415)	1.67 (0.31-3.10)	841,950 (568,756-1,289,481)	-	-

Note: COL = colonoscopy; dCosts = discounted costs; FIT = fecal immunochemical test.

(1) The costs were discounted with 3% annual rates and in 2021 Euro.

(2) Results are presented as mean and 95% credible interval.

3.4.2 Sex-differentiated starting age for screening colonoscopy

To answer Question 1 raised in Chapter 2.4.1, comparison between the strategies FIT1y50+COL10y55 and mCOL50/fFIT50+COL55 was made. The latter, sex-differentiated-starting-age strategy, demonstrated a superior overall reduction in CRC incidence and mortality (mean incidence HR 0.41 vs. 0.3 and mean mortality HR 0.31 vs. 0.27) in the perfect adherence scenario (Scenario 1). Diving into the sex-specific HRs, one can see that if men were entitled to colonoscopy 5 years earlier, it brought about 10% and 6% further reduction in CRC incidence risk and CRC mortality risk, respectively.

However, such advantage in protective effects for the overall cohort and men in the rest three scenarios was no longer seen. The two strategies in Scenario 2 and 3 offered very similar benefits. In Scenario 4, where FIT participation was >70% while colonoscopy participation only 42%, the non-sex-differentiated strategy provided more protective effects: mean HRs 0.69 and 0.57 for CRC incidence and mortality, respectively, in comparison with the counterpart mean HRs 0.71 and 0.60 in the sex-differentiated strategy. The results of the benefit comparison are presented in **Table 24**.

Table 25 summarizes the ICERs and INNCs between FIT1y50+COL10y55 and mCOL50/fFIT50+COL55 and treats the former as reference. Across the four scenarios, sex-differentiated strategy was more cost-saving than the undifferentiated one. The mean LYG across Scenario 1 to 3 and mean QALYG across all four scenarios showed positive results, hence, the majority of ICERs, both in terms of LYG and QALYG, signified that mCOL50/fFIT50+COL55 was the dominant strategy. Only in Scenario 4, it led to an ICER of €15,404 per LYG. However, when considering the 95% CrIs around the costs and benefits, only the ICER calculated with LYG in Scenario 1 and those calculated with QALYG in Scenario 1 and 3 showed absolute dominant results.

Regarding INNC, the sex-differentiated strategy resulted in more colonoscopies used in Scenario 1 and 2 (additional 104 and 16, respectively) as well as more CRC deaths prevented (additional 0.88 and 0.07 per 1,000, respectively). A reverse picture was observed in Scenario 3 and 4, where the sex-differentiated strategy resulted in fewer colonoscopies (1 and 33 less, respectively) used but also fewer CRC death prevented (0.19 and 0.76 per 1,000 less, respectively). These results led to the mean INNCs ranging between 7 and 235 more colonoscopies needed to prevent one additional CRC death.

Table 24. Comparison of CRC incidence and mortality reduction effects in screening colonoscopy strategies with and without differentiated starting age for different sex

	Overall		Men		Women	
Strategy[1]	**Incidence HR**	**Mortality HR**	**Incidence HR**	**Mortality HR**	**Incidence HR**	**Mortality HR**
Scenario 1 (perfect adherence)						
FIT1y50+COL10y55	0.41 (0.33-0.49)	0.31 (0.24-0.38)	0.42 (0.34-0.51)	0.32 (0.25-0.40)	0.39 (0.31-0.47)	0.29 (0.23-0.36)
mCOL50/fFIT50+COL55	0.35 (0.28-0.43)	0.27 (0.21-0.34)	0.32 (0.25-0.41)	0.26 (0.20-0.33)	0.39 (0.30-0.48)	0.29 (0.22-0.36)
Scenario 2 (current program)						
FIT1y50+COL10y55	0.85 (0.81-0.88)	0.74 (0.69-0.78)	0.86 (0.82-0.90)	0.75 (0.69-0.81)	0.83 (0.78-0.88)	0.72 (0.65-0.78)
mCOL50/fFIT50+COL55	0.84 (0.81-0.88)	0.73 (0.69-0.78)	0.85 (0.81-0.90)	0.74 (0.69-0.81)	0.83 (0.78-0.88)	0.72 (0.66-0.79)
Scenario 3 (mail-out FIT)						
FIT1y50+COL10y55	0.84 (0.80-0.87)	0.72 (0.68-0.77)	0.84 (0.80-0.89)	0.73 (0.67-0.79)	0.83 (0.78-0.88)	0.71 (0.65-0.78)
mCOL50/fFIT50+COL55	0.84 (0.80-0.88)	0.73 (0.69-0.77)	0.85 (0.81-0.89)	0.74 (0.69-0.80)	0.83 (0.77-0.88)	0.71 (0.65-0.77)
Scenario 4 (high adherence)						
FIT1y50+COL10y55	0.69 (0.63-0.75)	0.57 (0.52-0.63)	0.70 (0.64-0.76)	0.58 (0.52-0.64)	0.68 (0.62-0.75)	0.57 (0.50-0.64)
mCOL50/fFIT50+COL55	0.71 (0.66-0.76)	0.60 (0.56-0.65)	0.73 (0.68-0.78)	0.63 (0.57-0.69)	0.68 (0.62-0.76)	0.57 (0.50-0.63)

Note: HR = hazard ratio.

(1) Results are presented as mean and 95% credible interval.

Table 25. Incremental cost-effectiveness ratio and incremental number-needed-to colonoscope between screening colonoscopy strategies with and without differentiated starting age for different sex

Strategy[1,2,3,4]	ΔCosts	By LY		By QALY		ΔColonoscopy	ΔCRC deaths prevented	INNC
		LYG	ICER	QALYG	ICER			
Scenario 1 (perfect adherence)								
FIT1y50+COL10y55	--	--	--	--	--	--	--	--
mCOL50/fFIT50+COL55	CS (CS-CS)	8 (10-13)	Dominant (D-D)	14 (8-27)	Dominant (D-D)	104 (79-137)	0.88 (0.61-1.42)	119 (96-130)
Scenario 2 (current program)								
FIT1y50+COL10y55	--	--	--	--	--	--	--	--
mCOL50/fFIT50+COL55	CS (248-1,149)	2 (-4-1)	Dominant (D'ed-181)	3 (-2-9)	Dominant (D'ed-28)	16 (13-18)	0.07 (0.13-0.18)	235 (96-99)
Scenario 3 (mail-out FIT)								
FIT1y50+COL10y55	--	--	--	--	--	--	--	--
mCOL50/fFIT50+COL55	CS (CS-CS)	0 (-3- -4)	Dominant (4,955-5,943)	2 (0-5)	Dominant (D-D)	-1 (-2-0)	-0.19 (-0.15- 0.05)	7 (1-D)
Scenario 4 (high adherence)								
FIT1y50+COL10y55	--	--	--	--	--	--	--	--
mCOL50/fFIT50+COL55	CS (CS-CS)	-2 (-7- -1)	15,404 (4,146-23,785)	2 (-1-0)	Dominant (27,567-D)	-33 (-42- -23)	-0.76 (-0.61- -1.17)	44 (35-38)

Note: -- = reference; Δ = difference; CS = cost-saving; D = dominant; D'ed = dominated; ICER = incremental cost-effectiveness ratio; INNC = incremental number-needed-to-colonoscope; LY = life-years; QALY= Quality-adjusted life-years.

(1) The costs, life-years and quality-adjusted life-years are expressed as per 1,000 40-year-olds and were discounted with 3% annual rates. Costs were in 2021 Euro.

(2) FIT150+COL10y55 was used as the reference.

(3) The results are presented as: difference between means of the two strategies and difference between the upper value and lower values of 95% credible intervals of the two strategies.

(4) "Dominant" denotes a strategy being more effective and cost-saving; "Dominated" denotes a strategy being less effective and costing more.

3.4.3 Efficient strategies under different assumptions of participation rates

To answer questions 2 and 3 raised in Chapter 2.4.1, all strategies were evaluated based on the "efficiency", which was defined by cost-effectiveness using costs and LYs/QALYs and by NNC using number of colonoscopies used and number of CRC death prevented (burden-benefit analysis). The efficient frontiers derived from the cost-effective and burden-benefit analysis results for all four strategies are presented in **Figure 16** (costs and LYs), **Figure 17** (costs and QALYs), and **Figure 18** (number of colonoscopies and CRC deaths prevented). The cost-effectiveness compared to no-screening and ICERs among the efficient strategies are shown in **Table 26 and Table 27**, and NNCs and INNCs are illustrated in **Table 28**.

3.4.3.1 Efficiency frontiers based on cost-effectiveness

Overall, all strategies across all scenarios dominated no-screening in terms of cost-effectiveness, namely, they cost less and provided more LYs and QALYs compared to no-screening strategy. When focusing on LYs, COL10y50, COL10y45-3X, and FIT1y45+COL10y50 were the efficient ones across all four scenarios (expect for FIT1y45+COL10y50 in Scenario 1). To note, COL10y50-3X was only weakly dominated across all four scenarios, and in the current program scenario, the sex-differentiated strategy and FIT1y45+COL10y50-3X were also only slightly dominated. In general, the pure colonoscopy screening strategies and the combined strategies appeared to be more efficient than pure FIT ones, either starting at 45 or 50 years old. Strategies with FITs concentrated at the lower right corner and were strongly dominated by strategies involving colonoscopies. If €50,000/LYG was set as the WTP threshold, all efficient strategies across all scenarios could be deemed cost-effective when compared with their references, except FIT1y45+COL10y50 vs. COL10y45-3X in scenario 3, which the ICER was €68,273/LYG.

If QALYs were measured, the picture appeared to be very similar to the efficient frontiers seen in LYs: COL10y50 and COL10y45-3X remained the efficient strategies, as well as FIT1y45+COL10y50 in Scenario 4. Given €50,000/QALYG as the WTP threshold, ICERs among the efficient strategies across all scenarios were all well under the threshold, with the highest being FIT1y45+COL10y50 vs. COL10y45-3X in scenario 4, in which the ICER was €29,774/QALYG.

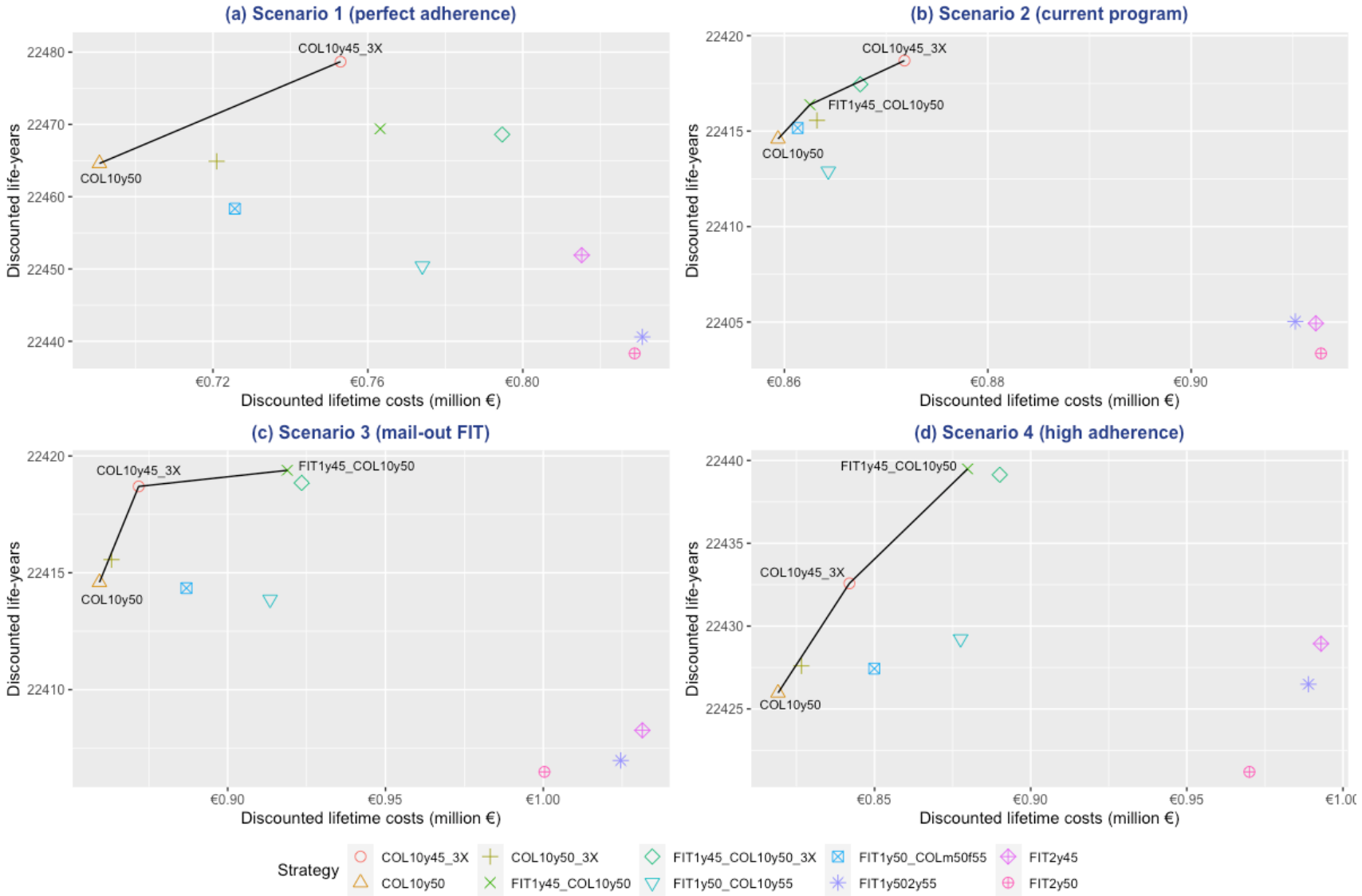


Figure 16. Efficiency frontier based on cost-effectiveness (life-years as benefits, 3% discount rate).
Texts shown in the graphs mark the dominant strategies in each scenario, which will be used to calculate incremental cost-effectiveness ratios (ICER).

Table 26. Cost-effectiveness and incremental cost-effectiveness ratios (ICERs) between efficient strategies (life-years as benefits)

Strategy[1,2,3,4]	Compared to no screening			ICER		
	ΔCosts	LYG	CER	ΔCosts	LYG	ICER
Scenario 1 (perfect adherence)						
COL10y50	Cost-saving (CS-CS)	83 (51-128)	Dominant (D-D)	--	--	--
COL10y45-3X	Cost-saving (CS-CS)	97 (58-151)	Dominant (D-D)	62,257 (20,501-90,227)	14 (7-23)	4,424 (2,858-3,852)
Scenario 2 (current program)						
COL10y50	Cost-saving (CS-CS)	33 (20-53)	Dominant (D-D)	--	--	--
FIT1y45+COL10y50	Cost-saving (CS-CS)	35 (19-50)	Dominant (D-D)	3,137 (CS-16,148)	2 (-3-0)	1,750 (D'ed-27,795)
COL10y45-3X	Cost-saving (CS-CS)	37 (22-53)	Dominant (D-D)	9,299 (6,127-6,696)	2 (2-3)	4,022 (2,153-3,291)
Scenario 3 (mail-out FIT)						
COL10y50	Cost-saving (CS-CS)	33 (20-53)	Dominant (D-D)	--	--	--
COL10y45-3X	Cost-saving (CS-CS)	37 (22-53)	Dominant (D-D)	12,436 (1,200-22,275)	4 (0-2)	3,030 (D'ed-653)
FIT1y45+COL10y50	Cost-saving (CS-CS)	38 (20-59)	Dominant (D-D)	47,070 (42,831-49,318)	1 (-1-7)	68,273 (D'ed-7,136)
Scenario 4 (high adherence)						
COL10y50	Cost-saving (CS-CS)	45 (24-63)	Dominant (D-D)	--	--	--
COL10y45-3X	Cost-saving (CS-CS)	51 (33-83)	Dominant (D-D)	22,855 (7,961-36,036)	7 (9-20)	3,454 (839-1,768)
FIT1y45+COL10y50	Cost-saving (CS-CS)	58 (35-94)	Dominant (D-D)	37,850 (19,128-42,528)	7 (1-11)	5,479 (3,979-14,587)

Note: -- = reference; Δ = difference; CER = cost-effectiveness ratio; CS = cost-saving; D = dominant; D'ed = dominated; ICER = incremental cost-effectiveness ratio; LYG = life-years gained.

(1) The costs and life-years are expressed as per 1,000 40-year-olds were discounted with 3% annual rates. Costs were in 2021 Euro.

(2) CERs were calculated by comparing each strategy with no-screening strategy. ICERs were calculated by comparing each strategy with the next most effective one (the rows next to each other).

(3) The results are presented as: difference between means of the two strategies and difference between the upper value and lower values of 95% credible intervals of the two strategies.

(4) "Dominant" denotes a strategy being more effective and cost-saving; "Dominated" denotes a strategy being less effective and costing more.

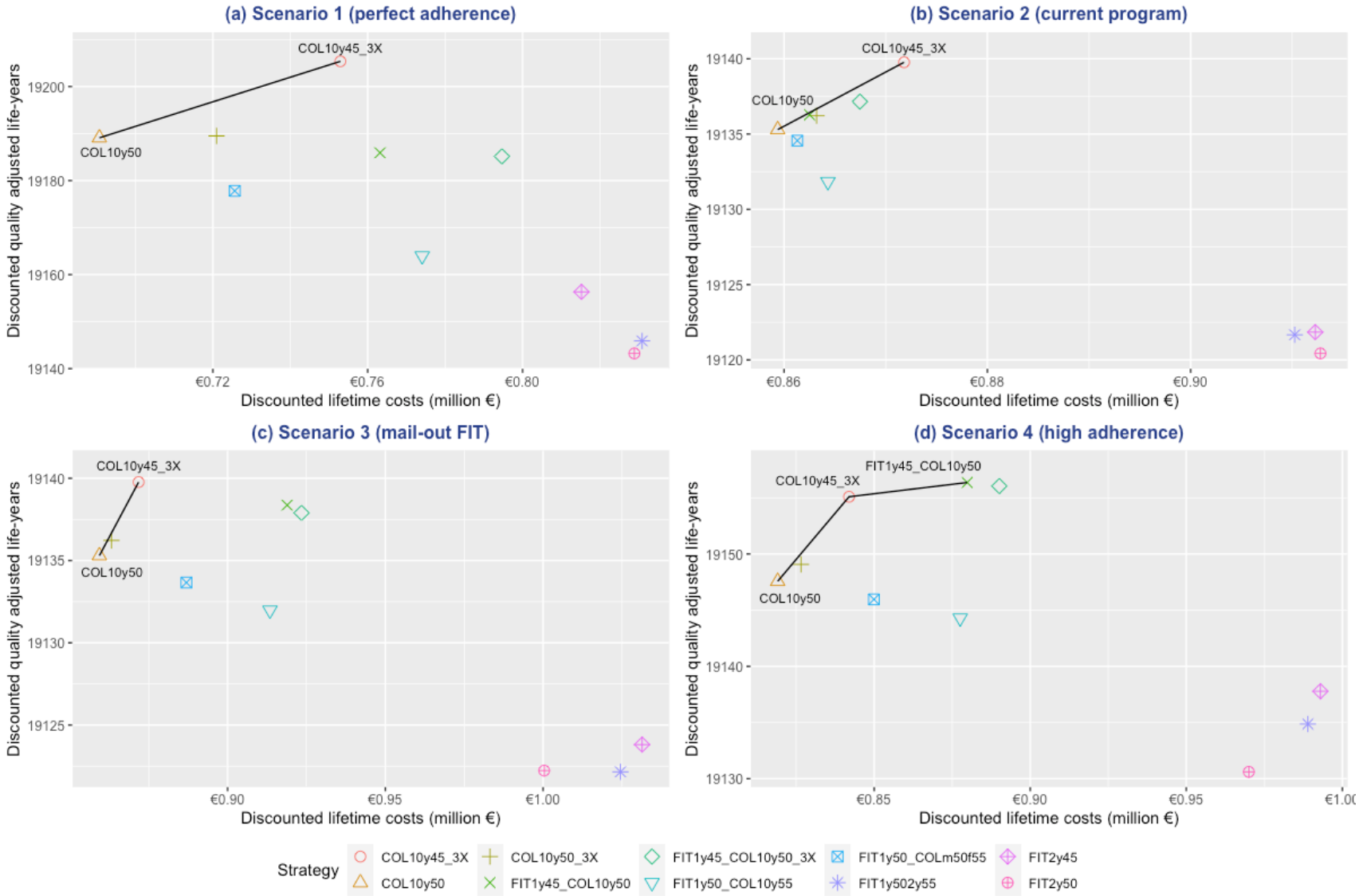


Figure 17. Efficiency frontier based on cost-effectiveness (quality adjusted life-years as benefits, 3% discount rate).
Texts shown in the graphs mark the dominant strategies in each scenario, which will be used to calculate incremental cost-effectiveness ratios (ICER).

Table 27. Cost-effectiveness and incremental cost-effectiveness ratios (ICERs) between efficient strategies (quality-adjusted life-years as benefits)

	Compared to no screening			Per QALY		
Strategy[1,2,3,4]	**ΔCosts**	**QALYG**	**CER**	**ΔCosts**	**QALYG**	**ICER**
Scenario 1 (perfect adherence)						
COL10y50	Cost-saving (CS-CS)	82 (68-92)	Dominant (D-D)	--	--	--
COL10y45-3X	Cost-saving (CS-CS)	98 (85-103)	Dominant (D-D)	62,257 (20,501-90,227)	16 (11-17)	3,826 (1,234-8,304)
Scenario 2 (current program)						
COL10y50	Cost-saving (CS-CS)	28 (21-30)	Dominant (D-D)	--	--	--
COL10y45-3X	Cost-saving (CS-CS)	33 (28-33)	Dominant (D-D)	12,436 (1,200-22,275)	4 (3-7)	2,780 (161-8,259)
Scenario 3 (mail-out FIT)						
COL10y50	Cost-saving (CS-CS)	28 (21-30)	Dominant (D-D)	--	--	--
COL10y45-3X	Cost-saving (CS-CS)	33 (28-33)	Dominant (D-D)	12,436 (1,200-22,275)	4 (3-7)	2,780 (161-8,259)
Scenario 4 (high adherence)						
COL10y50	Cost-saving (CS-CS)	40 (33-42)	Dominant (D-D)	--	--	--
COL10y45-3X	Cost-saving (CS-CS)	48 (39-54)	Dominant (D-D)	22,855 (7,961-36,036)	8 (5-13)	3,038 (1,460-2,874)
FIT1y45+COL10y50	Cost-saving (CS-CS)	49 (41-49)	Dominant (D-D)	37,850 (19,128-42,528)	1 (-5-2)	29,774 (D'ed-9,869)

Note: -- = reference; Δ = difference; CER = cost-effectiveness ratio; CS = cost-saving; D = dominant; D'ed = dominated; ICER = incremental cost-effectiveness ratio; QALYG= quality-adjusted life-years gained.

(1) The costs and quality-adjusted life-years are expressed as per 1,000 40-year-olds were discounted with 3% annual rates. Costs were in 2021 Euro.

(2) CERs were calculated by comparing each strategy with no-screening strategy. ICERs were calculated by comparing each strategy with the next most effective one (the rows next to each other).

(3) The results are presented as: difference between means of the two strategies and difference between the upper value and lower values of 95% credible intervals of the two strategies.

(4) "Dominant" denotes a strategy being more effective and cost-saving; "Dominated" denotes a strategy being less effective and costing more.

3.4.3.2 Efficiency frontiers based on burden-benefit analysis

When number of colonoscopies and CRC deaths were used to determine efficient strategies (burden-benefit analysis), it appeared different to the picture seen on the cost-effective planes. Using this measurement, some pure FIT screening strategies became efficient. In perfect adherence scenario (Scenario 1), FIT2y50, FIT2y45, COL10y50, FIT1y45+COL10y50, and COL10y45-3X were efficient, and their NNCs ranged between additional 57-151 colonoscopies to prevent one extra CRC death compared to no-screening. As for the INNCs among the efficient strategies, the steepest increase was seen in COL10y45-3X vs. FIT1y45+COL10y50, with an INNC of 605 more colonoscopies to prevent one additional CRC death. To note, all non-efficient strategies in scenario 1 were located in the triangle spaces below the efficiency frontier, meaning they were only weakly dominated.

The efficient strategies in Scenario 2 to 4 were very consistent, with FIT2y50, FIT1y50+2y55, FIT1y45+COL10y50, and FIT1y45+COL10y50-3X always on the list. The NNCs of the efficient strategies ranged from 22 to 92 extra colonoscopies to prevent one additional CRC death in Scenario 2 and from 47 to 108 extra colonoscopies in Scenario 4. The INNCs in the three scenarios were very close to those in Scenario 1. Pure colonoscopies were predominantly strongly dominated in Scenario 2 to 4.

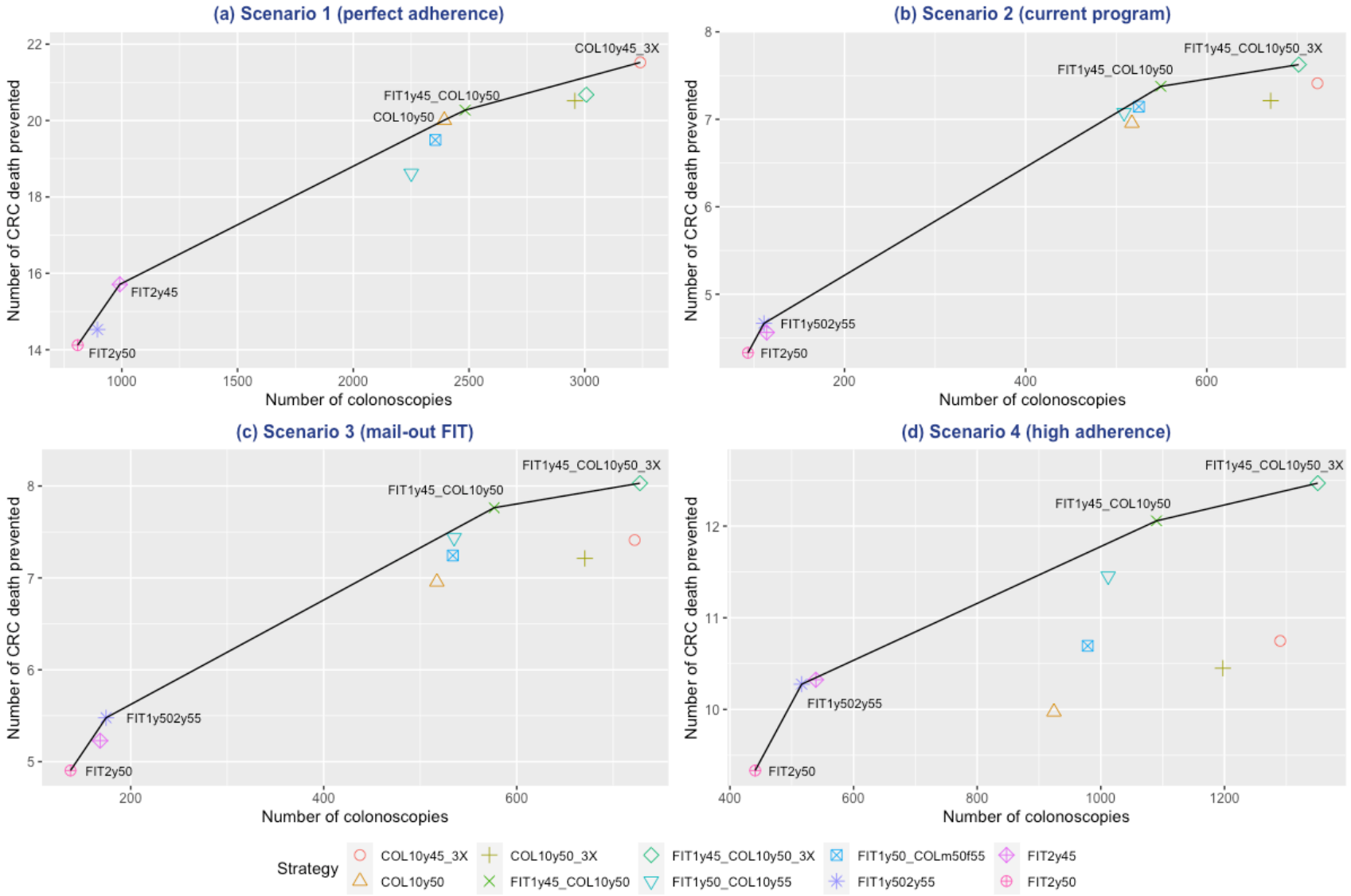


Figure 18. Efficiency frontier based on number-needed-to-colonoscope to prevent one CRC death.
Texts shown in the graphs mark the dominant strategies in each scenario, which will be used to calculate incremental number-needed-to-colonoscope (INNC).

Table 28. Number-needed-to-colonoscope (NNC) and incremental number-needed-to-colonoscope (INNC) between efficient strategies

Strategy[1,2]	NNC[3]			INNC[3]		
	ΔColonoscopy	ΔCRC deaths prevented	NNC	ΔColonoscopy	ΔCRC deaths prevented	INNC
Scenario 1 (perfect adherence)						
FIT2y50	809 (609-1,043)	14.12 (9.34-21.49)	57 (49-65)	--	--	--
FIT2y45	991 (737-1,280)	15.71 (10.41-23.77)	63 (54-71)	182 (128-237)	1.59 (1.07-2.28)	114 (104-120)
COL10y50	2,393 (2,084-2,745)	20.01 (13.29-29.43)	120 (93-157)	1,584 (1,476-1,702)	5.89 (3.95-7.94)	269 (214-374)
FIT1y45+COL10y50	2,483 (2,168-2,828)	20.27 (13.29-30.01)	122 (94-162)	1,492 (1,431-1,548)	4.56 (2.98-6.24)	327 (248-480)
COL10y45-3X	3,240 (2,825-3,700)	21.52 (14.25-31.94)	151 (116-198)	757 (657-872)	1.25 (0.86-1.93)	605 (452-765)
Scenario 2 (current program)						
FIT2y50	93 (72-118)	4.33 (2.91-6.39)	22 (18-25)	--	--	--
FIT1y50+FIT2y55	111 (85-141)	4.67 (3.22-7.27)	24 (19-26)	18 (13-23)	0.34 (0.31-0.88)	52 (26-41)
FIT1y45+COL10y50	549 (497-609)	7.38 (5.02-11.11)	74 (55-99)	438 (413-468)	2.71 (1.90-3.84)	162 (122-229)
FIT1y45+COL10y50-3X	701 (640-768)	7.63 (5.21-11.55)	92 (67-123)	152 (143-160)	0.25 (0.19-0.44)	612 (363-751)
Scenario 3 (mail-out FIT)						
FIT2y50	138 (106-173)	4.90 (3.38-7.84)	28 (22-31)	--	--	--
FIT1y50+FIT2y55	174 (133-222)	5.48 (3.69-8.59)	32 (26-36)	37 (27-49)	0.57 (0.31-0.75)	64 (65-88)
FIT1y45+COL10y50	577 (520-641)	7.76 (5.34-11.72)	74 (55-97)	402 (387-419)	2.29 (1.65-3.13)	176 (134-235)
FIT1y45+COL10y50-3X	728 (664-802)	8.03 (5.45-12.44)	91 (64-122)	151 (144-161)	0.27 (0.11-0.73)	565 (221-1298)
Scenario 4 (high adherence)						
FIT2y50	441 (339-554)	9.33 (6.21-14.55)	47 (38-55)	--	--	--
FIT1y50+FIT2y55	516 (394-649)	10.28 (6.87-15.74)	50 (41-57)	75 (55-94)	0.94 (0.66-1.19)	80 (83-79)
FIT1y45+COL10y50	1,090 (976-1,222)	12.06 (8.07-18.11)	90 (67-121)	574 (574-582)	1.78 (1.20-2.37)	322 (242-484)
FIT1y45+COL10y50-3X	1,351 (1,229-1,496)	12.47 (8.35-18.86)	108 (79-147)	261 (253-274)	0.41 (0.28-0.75)	634 (365-907)

Note: -- = reference; Δ = difference; CS = cost-saving; D'ed = dominated; INNC = incremental number-needed-to-colonoscope; NNC = number-needed-to-colonoscope.

(1) The results are presented as per 1,000 40-year-olds and as difference between means of the two strategies and difference between the upper value and lower values of 95% credible intervals of the two strategies.

(2) "Dominant" denotes a strategy being more effective and cost-saving; "Dominated" denotes a strategy being less effective and costing more.

(3) NNC was calculated by comparing each strategy with no-screening strategy. INNC was calculated by comparing each strategy with the next most effective strategy (the rows next to each other).

3.4.4 Results of sensitivity analyses

3.4.4.1 Probabilistic sensitivity analyses and cost-effectiveness acceptance curves

The PSA results with 3% discounting annual rate (base case) using the NHB methods given different WTP thresholds were plotted as CEAC in **Figure 19 and 20**. Across the four scenarios, the PSA results were generally consistent regardless of using LYs or QALYs as the benefit measure. COL10y45-3X was consistently the most cost-effective strategy from WTP €5,000 to €100,000 per LYG or QALYG, even with a >50% probability from WTP >€15,000 being the most cost-effective in Scenario 1 using QALYs. In the scenario with mail-out FIT (Scenario 3), COL10y45-3X was the leading strategy under €20,000 per LYG, but with increasing WTP thresholds, the gaps between COL10y45-3X and the other two combined strategies starting at age 45 narrowed. On the other hand, if measured by costs per QALYG, COL10y45-3X remained the most cost-effective across WTP €5,000 to €100,000 per QALYG. In high adherence scenario (Scenario 4), the combined strategies starting at age 45 years (FIT1y45+COL10y50 and FIT1y45+COL10y50-3X) were the most cost-effective, with a probability around 25% being the most cost-effective from WTP €5,000 to €100,000 per LYG or QALYG. Generally, colonoscopy-related strategies were more likely to be cost-effective if starting from age of 45 years. Whether two or three colonoscopies were offered did not seem to affect the cost-effectiveness significantly, both in terms of LYs and QALYs. On the other hand, pure FIT strategies were the least cost-effective ones across all scenarios.

3.4.4.2 Sensitivity analyses with different discount rates

The results of sensitivity analyses using 0% and 5% discounting annual rates are presented in Chapter 9.3 **Appendix C**. In general, all screening strategies, even under alternative discount rate assumptions, dominated no screening. The only non-dominant exceptions were found in the strategy FIT1y50+FIT2y55 and FIT2y45 in Scenario 3 and 4 under the assumption of 5% discount rate, yet they were both with a competitive ICER under €2,000 per LYG or QALYG compared to no screening (**Appendix C, Supplementary Table 7, ICERs not shown**).

The sensitivity test results with both 0% and 5% discount rates for questions 1 remained unchanged: the strategy with sex-differentiated starting age for colonoscopy (mCOL50/fFIT50+COL55) was the dominant strategy compared to the non-sex-differentiated strategy (FIT1y50+COL10y55) in Scenario 1 to 3, and with a competitive ICER in Scenario 4 (**Appendix C, Supplementary Table 8 and 9**).

Pertaining the efficient strategies, 5% discount rates yielded the same efficient strategies as seen in the base case with 3% discount rate across scenarios (**Appendix C, Supplementary Figure 3, 4, 7, 8 and Table 12,13**). In addition, one can notice that the cost difference between the pure FIT screening strategies and colonoscopy or combined strategies became smaller, yet the narrowed cost difference was still not able to make the pure FIT screening strategies fall on the efficiency frontier. On the other hand, the results from the tests with 0% discount rate revealed a somewhat different picture compared to the base case with 3% discount rate (**Appendix C, Supplementary Figure 1, 2, 5, 6 and Table 10,11**). In Scenario 2, COL10y50 was no longer on the efficient frontier, but the three strategies involving colonoscopies starting at age 45 were efficient. In Scenario 3, COL10y50 was replaced by COL10y50-3X on the efficient frontier, and the slope between COL10y45-3X and FIT1y45+COL10y50 became flatter, which resulted in a higher ICER of €308,421 per LYG. In Scenario 4, FIT1y45+COL10y50 and FIT1y45+COL10y50-3X clearly dominated the rest of the strategies.

CEACs of alternative discount rates are illustrated in **Appendix C, Supplementary Figure 5-8**. In perfect adherence scenario, COL10y45-3X was clearly still the most cost-effective strategies regardless of the discount rates across the WTP thresholds. In Scenario 2, COL10y45-3X remained the most cost-effective in both 0% and 5% discount rate cases, but the advantage was less obvious in the 0% discount rate case. In Scenario 3, FIT1y45+COL10y50 and FIT1y45+COL10y50-3X overtook COL10y45-3X as the most cost-effective strategies above WTP €10,000 per LYG or QALYG in 0% discount rate case, while COL10y45-3X was persistently the most efficient one in 5% case as in the base case. The only difference in the sensitivity tests for Scenario 4 was found to be in the CEAC measured by costs per QALYG in the 5% case, where COL10y45-3X was the most cost-effective, unlike in the rest of cases. Overall, the sensitivity analyses confirmed that colonoscopy or FIT-colonoscopy combined strategies starting from age 45 were likely to be the most cost-effective ones even under alternative discount rate assumptions.

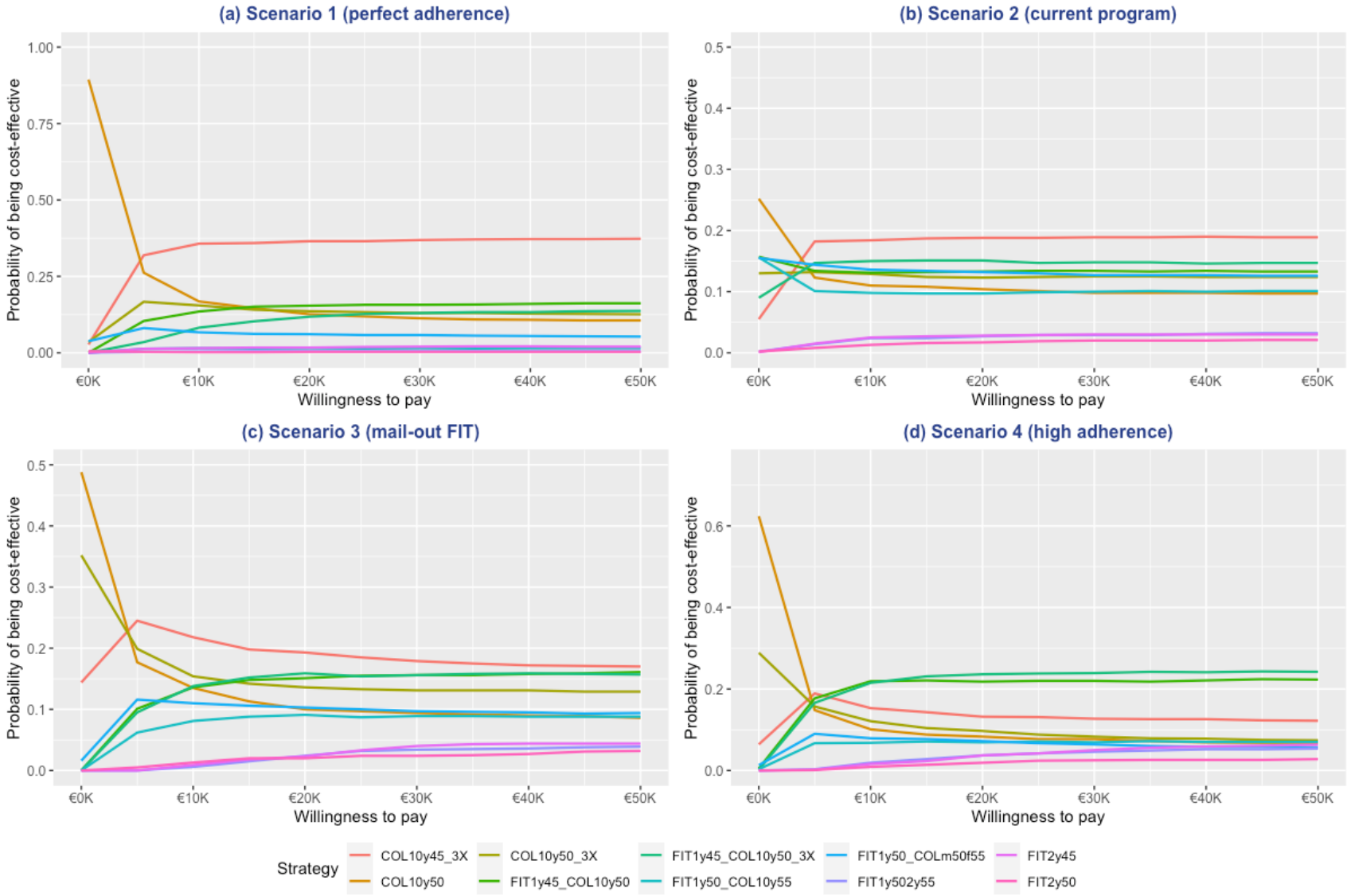


Figure 19. Cost-effectiveness acceptance curve (life-years as benefits, 3% discount rate).
Each line denotes how likely each strategy could be the most cost-effective strategy under the given willingness-to-pay threshold.

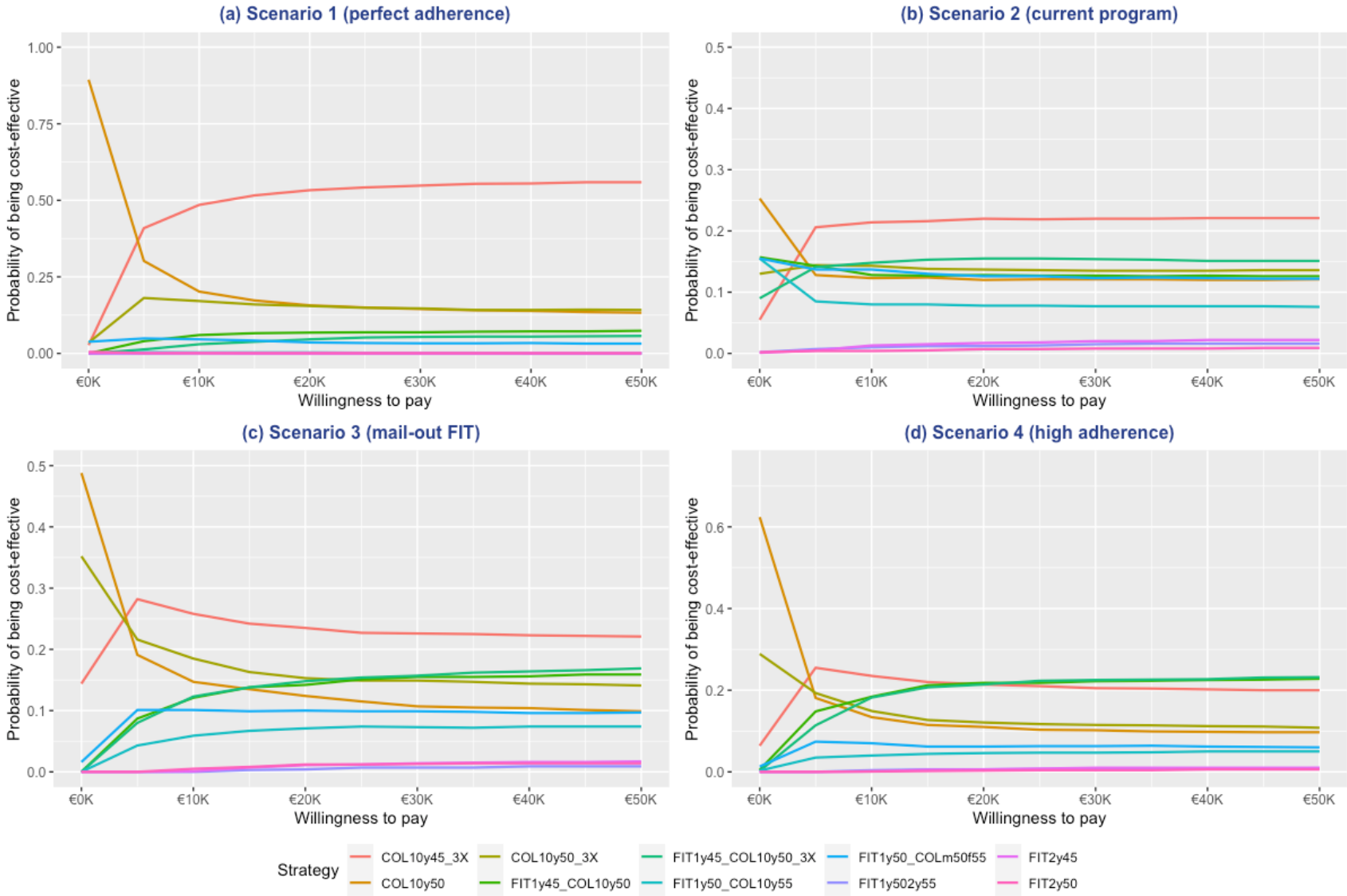


Figure 20. Cost-effectiveness acceptance curve (quality adjusted life-years as benefits, 3% discount rate).
Each line denotes how likely each strategy could be the most cost-effective strategy under the given willingness-to-pay threshold.

4 DISCUSSION

In this chapter, the discussion will start with the observations over the distinct points related to DECAS development, its calibration and validation results, and interpretation of the cost-effectiveness analysis. This chapter will wrap up with the strengths and limitations of the present thesis and future research directions.

4.1 DECAS natural history and screening model development

(Part of Chapter 4.1 has been submitted and currently under review (Cheng et al., 2022)*)*

As the center piece of the present doctoral thesis, DECAS, a DES model for CRC natural history, was developed. It considers both adenoma-carcinoma and serrated neoplasia pathways, and is calibrated by using an approximate Bayesian computation algorithm, APMC, which realizes the estimation of 21 input parameters and their uncertainties. The calibration made use of years of nationwide data from the German screening colonoscopy program containing information of millions of participants. DECAS demonstrated the ability to reproduce real-world data in Germany through the validation against the Garman cancer registry data.

4.1.1 CRC natural history modeling compared with existing models

4.1.1.1 With CISNET models

The three CISNET models (Kuntz et al., 2011; van Ballegooijen et al., 2011) are the most widely used models to inform CRC screening policies (Shaukat et al., 2021; US Preventive Services Task Force, 2021; Wolf et al., 2018). Although they do not consider the serrated neoplasia pathway, they have provided extensive transparency in their model structures and assumptions (Knudsen et al., 2020). It will provide an informative overview by comparing DECAS with the three CISNET models (**Table 29**).

DECAS employs very similar assumptions in the precancerous lesion initiation risk as with CRC-SPIN1.0/2.0 and MISCAN while being the closest to MISCAN in terms of lesion growth modeling. Pertaining to the transitioning to pre-clinical cancer, all four models make slightly different assumptions regarding varying risks due to age, sex, adenoma size and location: DECAS considers random risk for each lesion and systematically only by lesion type, age and location effects in this state transition. As for the step of transitioning to clinical cancers, all four models make comparable assumptions except if transition times vary across locations (Knudsen et al., 2020).

Table 29. Natural history model structure comparison between DECAS and CISNET models

Component	DECAS	SimCRC	CRC-SPIN	MISCAN
Precancerous lesion initiation				
Mechanism	Poisson process	Logistic function	Poisson process	Poisson process
Risk varies:				
Randomly across individuals	Yes	Yes	Yes	Yes
Systematically by age and sex	Yes	Yes	Yes	Yes
Precancerous lesion progression				
Mechanism	Time to next lesion state	Time in each size category	Growth curve	Time in each size category
Risk varies:				
Randomly across individuals	Yes	Yes	Yes	Yes
Systematically by location	No	Yes	Yes	No
Transition times correlated across size categories	Yes	No	Yes	Yes
Transition to pre-clinical CRC				
Mechanism	Poisson process	Logistic function	Adenoma size at transition	Overall transition probability
Risk varies:				
Randomly across lesions	Yes	Yes	Yes	No
Systematically by				
Sex	No	Yes	Yes	No
Age	Yes	Yes	Yes	Yes
lesion size	NA	No	Yes	Yes
Location	Yes	Yes	Yes	No
Transition times correlated across pre-clinical stages	No	No	NA	Yes
Transition to clinical CRC				
Mechanism	Time to transition	Time to transition	Time to transition	Time to transition
Transition times vary:				
Randomly across pre-clinical CRCs	Yes	Yes	Yes	Yes
Systematically by				
Sex	No	No	No	Yes
Location	No	Yes	Yes	Yes
Correlated with duration of pre-clinical CRC	No	No	No	Yes

Note: CRC = colorectal cancer; NA = not applicable.

Reference: (Knudsen et al., 2020)

Three CISNET models are anchored with calibration data of adenoma prevalence from autopsy studies and CRC incidence data from the US Surveillance, Epidemiology, and End-Results (SEER) data in the pre-screening era (Rutter et al., 2016). Between the two anchoring states, CISNET models have jointly reported the dwell time and sojourn time, which are two of the most crucial parameters in CRC modeling (Silva-Illanes and Espinoza, 2018) and can be seen as summary outputs of the four model components altogether (**Table 29**). The length of the two time parameters may have directly influence on the effectiveness of screening tests to detect precancerous lesions or pre-clinical cancers as the time serves the window of opportunity for early detection and lesion removal (Rutter et al., 2016).

In comparison to CISNET models, the dwell time in DECAS is longer, and it is likely that some people might harbor lesions for half of their lifetime without the lesions transitioning to cancers (**Table 30**). Serrated lesions have a slightly shorter dwell time compared to adenomas in DECAS. One potential explanation for longer dwell time is that in DECAS calibration, the CRC prevalence from the German screening colonoscopy registry were treated as the proxy for the prevalence of pre-clinical cancers. This treatment will inevitably increase the dwell time approximation in DECAS slightly as it was unknown how long those pre-clinical cancers found in screening participants have already existed, and, strictly speaking, this unknown period of time should be regarded as sojourn time and taken out from the dwell time. Nevertheless, the CRC prevalence discovered from screening registry were the best available data in Germany to serve as the proxy for pre-clinical cancer prevalence, and Brenner et al. also used the data in a similar way to estimate the transition probabilities from advanced adenoma to pre-clinical cancers (Brenner et al., 2013). Additionally, the slight increase in dwell time estimation did not appear to influence the forecast of CRC incidence reduction as seen in the shorter dwell time (see Chapter 4.1.3 for more details), therefore, it still justifies the use of screening-detected CRC prevalence as calibration targets for pre-clinical cancers.

In regard with sojourn time, as DECAS took it from a predefined Weibull distribution with a mean of 4.7 years, which was a direct input from another German study (Brenner et al., 2011), it does not show too much variations in the DECAS simulated results. Among the CISNET models, MISCAN used the same source of sojourn time (Brenner et al., 2011), while the sojourn time assumptions in CRC-SPIN2.0 and SimCRC were somewhat shorter, which had means of 3.6 and 4.0 years, respectively (Knudsen et al., 2020). The authors of CISNET models reached a consensus that the reasonable sojourn time should be close to 4 years (Rutter et al., 2016). Hence, the sojourn time assumption in DECAS is comparable with CISNET models.

Table 30. Dwell time and sojourn time comparison between DECAS and CISNET models

	DECAS[1]		SimCRC	CRC-SPIN2.0	MISCAN
	Adenoma	Serrated lesion			
Dwell time (precancerous lesion occurrence until transition to pre-clinical cancers)					
Mean (IQR)	38.6 (28.4-49.7)	37.2 (26.6-48.5)	21.9 (12-29)	25.4 (16-33)	12.5 (4-18)
Sojourn time (pre-clinical cancer occurrence until transition to clinical cancers)					
Mean (IQR)	4.7 (4-5.4)		4.0 (2-5)	3.6 (2-5)	4.7 (1-7)

Note: IQR = interquartile range.

(1) DECAS recorded the dwell time separately for adenomas and serrated polyps as it is random which type of lesion occurs first in an individual.

Reference: (Knudsen et al., 2020)

4.1.1.2 With models considering two pathways

DECAS is most similar to the ASCCA model (Greuter et al., 2014) and its variant, Policy1-Bowel (Lew et al., 2017), in terms of model structure – all incorporated adenoma-carcinoma and serrated neoplasia pathways for CRC tumorigenesis. To date, DECAS, the ASCCA model and Policy1-Bowel are believed to be the only three CRC models intended for economic evaluation that explicitly address both CRC carcinogenesis pathways (Silva-Illanes and Espinoza, 2018). The other existing models assume CRCs arise only from the adenoma-carcinoma pathway (Greuter et al., 2014; Silva-Illanes and Espinoza, 2018), although some models incorporate the design of "de novo cancers" that arise directly from normal epithelium or representing alternative pathways (Whyte et al., 2011; Prakash et al., 2017; Silva-Illanes and Espinoza, 2018). The serrated neoplasia pathway should arguably be an indispensable part of CRC models, especially when used for projecting the effectiveness of screening. The reason being that fecal immunochemical testing appears to have a lower sensitivity for serrated polyps (Chang et al., 2017; Zorzi et al., 2017), and the colonoscopy miss rate for serrated polyps also appears to be higher (Zhao et al., 2019), which is likely due to the flat or sessile morphology, similar color to the epithelium and camouflage by a mucus cap (Rex et al., 2012). Without the inclusion of the serrated pathway and adjusting the sensitivities of the screening intervention, the modeled estimation of screening effectiveness might be over-optimistic (Greuter et al., 2014).

In addition to the inclusion of serrated pathway, ASCCA models included lesion characteristics such as morphology, dysplasia and villosity (Greuter et al., 2014), which provided more details in modeling the disease progression. DECAS achieved in accounting for similar lesion features by incorporating them into the concept of AA and crSP, and the approach allowed a relative model parsimony and mitigated the influence of the "curse-of-dimensionality" (Nott et al., 2019) by reducing the amount of parameters needed.

DECAS offers two advantages over the ASCCA models. First, DECAS takes the DES approach to model the disease progression in continuous time, while the ASCCA models (as well as most of the existing CRC microsimulation models (Loeve et al., 1999; Kuntz et al., 2011; Prakash et al., 2017)) simulate in discrete time. One of the key advantages of using DES model is to incorporate time-varying event risks along the disease course in a straightforward manner (Karnon and Haji Ali Afzali, 2014). For example, in ASCCA models, the probabilities of precancerous lesion occurrence are discrete values by age baskets of 5- to 20-year intervals. In theory, that creates the need for "tunnel states": multiple sub-states divided from a health state, each representing the health state in a pre-defined time after the start of the state (Karnon and Haji Ali Afzali, 2014). With finer stratification and more tunnel states, it might lead to so-called "state explosion." (Siebert et al., 2012) On the other hand, DES models simulate time-varying risks by mathematical functions (in

the case of DECAS, **Equation 1**), which simplifies the model structure. Continuous-time modeling also offers another advantage: it only updates when events happen without periodical check, when most of the time nothing is to be updated (Caro and Möller, 2016).

Second, DECAS is calibrated with a Bayesian inference method, which is to intentionally address parameter uncertainty. On the other hand, the ASCCA models (Greuter et al., 2014; Lew et al., 2017) and the majority of the existing models (Loeve et al., 1999; Kuntz et al., 2011; Prakash et al., 2017) take the optimization approach for calibration, specifically the Nelder-Mead algorithm. Commonly used parameter optimization algorithms for the calibration of disease models applied in economic evaluation (Vanni et al., 2011), including the Nelder-Mead algorithm, output only one single set of "optimal" parameters, and it may lead to loss of a great amount of valuable information from the data, which would have been helpful to understand the parameter uncertainty (Daly et al., 2017). Although some models (e.g., ASCCA) mitigate this drawback by keeping multiple sets of best-fit parameters to capture the uncertainty using confidence intervals (Greuter et al., 2014; Lew et al., 2017), it is of concern that the frequentist methods might have unsatisfactory performance in the face of large parameter spaces and highly nonlinear models (Daly et al., 2017). Bayesian calibration methods offer remedy to the challenges faced by frequentist approach.

4.1.1.3 With models calibrated with Bayesian methods

Bayesian inference methods, including ABC, naturally encapsulate the uncertainty of parameter space as well as inter-dependencies in the joint posterior distribution of the parameters, and they are more robust to address the parameter uncertainty (Daly et al., 2017). Therefore, for a complex model with high parameter dimensions as DECAS is, the implementation of calibration with a Bayesian method is arguably preferable. Rutter et al. (2009) and Whyte et al. (2011) successfully applied a Bayesian inference method, MCMC with Metropolis-Hasting sampling, to calibrate a microsimulation and a cohort CRC model, respectively. However, MCMC requires repeated evaluation of the likelihood function, which describes the relationship between model parameters and observed data, to accept parameter sets. Rutter et al. (2009) admitted that the exact functional form of their likelihood was essentially unknown. To tackle the issue, they proposed using an approximate Metropolis-Hasting algorithm embedded within the microsimulation model to estimate the likelihood (Rutter et al., 2009).

In complex models where the exact likelihood function is intractable, a more natural and sensible approach is to use the likelihood-free ABC methods, which have been shown to approximate the "true" posterior distribution in a satisfactory manner (Beaumont, 2019). In CRC screening modeling, Rutter et al. (2018) proposed an ABC method, Incremental Mixture ABC (IMABC), which

can be regarded to be in the same family as APMC that uses adaptive importance sampling and a clear stopping rule, and they successfully updated the CRC-SPIN 2.0. However, IMABC requires several steps of fine-tuning with various calibration population sizes for different components of the model sequentially (Rutter et al., 2018), which does not seem to be intuitively generalizable. DECAS sought to use a more universal ABC algorithm and, consequently, demonstrated a successful example of applying an established ABC method, APMC (Lenormand et al., 2013), without major modification of the original algorithm. With the APMC calibration, DECAS captures the valuable information of parameter uncertainty surrounding the CRC-related data used for calibration. The captured uncertainty during calibration will be carried to the forecasted results of CRC screening effectiveness and cost-effectiveness in the form of probabilistic sensitivity analyses, which play a crucial role in the interpretation of model simulated results.

4.1.2 Challenges to calibrate DECAS natural history model

There are some challenges faced during the calibration process. The first one is pertaining to parameter estimation. Bayesian calibration methods aim to capture the parameter uncertainty and describe them with distributions. However, as shown in **Figure 8**, some posterior parameter distributions concentrated a relatively high probability mass close to the prior boundaries. This indicated that better fits were generally achieved when the parameter in question was close to the boundary. The current results were derived after several iterations in adjusting the prior ranges, yet, the prior ranges were not further extended to completely avoid boundary-hitting based on two reasons: (1) some parameters should be strictly positive (e.g., risk of progressing to advanced lesions or cancers) and the lower range already corresponds to very low risks, and (2) some were already given a very generous range compared with the data informing our priors (e.g., the age factors and location factors).

The second challenge is related to structural uncertainty (Briggs et al., 2012). Although DECAS predictions captured the general trend of calibration targets well in a way that the simulated data and their CrIs covered the targets (except in the older age >75 years in certain type of lesions), the targets did not appear to randomly spread around the predicted means, namely, residual autocorrelation might be present. This might indicate the need for a more precise algorithm to better estimate such parameters, and that the data provided only limited information to the parameters given the current model form and parameter constraints. For instance, a more granular piece-wise age effect or alternative change-points in the initiation and cancer progression might mitigate the problem. However, one must balance between the prediction accuracy and the demanded resources (both data and computational resources). Given that the

main trends were captured with good overall accuracy, the current model formulation and settings were kept.

Thirdly, DECAS struggled to capture the trend of non-AA, female non-crSP, and crSP prevalence that well compared with the other target data, especially after the age of 70 to 75 years, where all three real-world prevalence have a downward trend. Although the age effects in lesion occurrence function allowed negative values to flatten the prevalence trend, the trend in older age among these parameters still showed mismatch patterns. The reasons are likely to be multifaceted. After all, the final calibration targets, CRC prevalence, are strictly increasing with age, and it has a steeper slope after age of 70 years. When the APMC calibration algorithm attempts to consider a global fit, it is more likely to accept higher simulated older-age precancerous lesion prevalence in order to yield higher older-age CRC prevalence, which contributes to the overall increasing trend in precancerous lesion prevalence and their mismatch in older age. Attempt to mitigate this gap was taken by assigning age factors in the transition from advanced precancerous lesions (AA and crSP) to CRC, which reflected the fact that a higher grade of dysplasia is more likely to happen in older age (Brenner et al., 2007; Lash et al., 2010; Bouwens et al., 2014). Despite the efforts, the trend mismatch still existed.

Another explanation could lie in the data used for calibration. For non-AA prevalence, several other studies evaluating prevalence using screening colonoscopy data revealed an increasing trend of prevalence until age of 80 (Diamond et al., 2011; Ferlitsch et al., 2011), which is 10 years more of the increasing trend compared to the German data used in this thesis. Interestingly, one German study in 2014/15 surveying the colonoscopy usage for any reason in the last 10 years discovered a >60% usage rate for people >70 years old (Starker et al., 2017). In contrast, the 10-year screening colonoscopy participation rate of people older than 75 years old is only slightly over 10% (Gesundheitsberichterstattung des Bundes (GBE), 2021). Therefore, the downward trend in non-AA prevalence in older age might be associated with healthy user bias in the screening data, in which the high usage of colonoscopy in older age outside of screening colonoscopy creates a screening population with lower risk of having non-AA lesions.

Data issues could as well be one of the explanations for trend mismatch in non-crSP and cr-SP prevalence. In literature, there is no obvious increasing or decreasing trend associated with older age in non-crSP or crSP prevalence (Ijspeert et al., 2017; Schramm et al., 2018) , unlike in advanced adenomas and CRC prevalence. However, the calibrated targets for serrated lesions used in this thesis showed a downturn in older age prevalence. This could be the case due to broad variability in the data given data scarcity (Ijspeert et al., 2017; Schramm et al., 2018). CRC-SPIN 2.0 (Rutter et al., 2018) and ASCCA (Greuter et al., 2014) also did not match all the data points in their

calibration. This challenge related to the potential data bias can only be solved by more future epidemiological research to shed light on the precancerous lesion prevalence.

4.1.3 Predicted CRC screening effects by DECAS screening model

DECAS performed three validation exercises, and all of them showed that DECAS can predict the CRC screening effectiveness well, either with stool tests and endoscopy screening.

4.1.3.1 Validation with UKFSS trial and cross-validation with CISNET models

Among the three validation exercises, validation against UKFSS trial carries the most weights to give credits to the screening effect prediction capability of DECAS. The UKFSS trial serves as a unique study well suited for CRC screening model validation, as the trial is designed with a one-time FS screening intervention in the era when CRC screening was yet in place, which avoids the "contamination" in the control group (i.e., control group is unlikely to receive FS screening during the follow-up) (Rutter et al., 2016). This allows the models to evaluate the clinical incidence and mortality reduction over a certain period of time, capturing the difference between the post-screening and background incidence and mortality rates (van Ballegooijen et al., 2011). The incidence reduction is deemed as an important summary indicator for CRC screening models, and, unlike the unknown real-world dwell time and sojourn time, the incidence reduction can be assessed by benchmarking against trial or cohort study data (e.g., the UKFSS trial).

Overall, DECAS predicted the 17-year CRC incidence and mortality reduction accurately compared with the estimates from the UKFSS trial. DECAS also predicted the overall 17-year CRC incidence cumulative rate precisely comparing with the observation in the UKFSS trial. However, DECAS output more proximal cancers than distal cancers, which was the opposite in the UKFSS trial. In the UKFSS trial, distal cancers are defined as cancers located in sigmoid and rectum (Atkin et al., 2017). DECAS assigns the cancer locations when the precancerous lesions first occur in the model using the location distribution observed in screening colonoscopy, where 40% of adenomas and merely 4% of serrated polyps (and the cancers deriving from each pathway) are located in sigmoid and rectum. It is likely the very low percentage of distal cancers stemming from serrated pathway in DECAS contributes to the phenomenon – after all, DECAS assume 15% of cancers originate from serrated polyps. In terms of 17-year CRC mortality rate, DECAS estimated a higher rate for both the screening and control arm. This is likely due to the background CRC mortality rate variation between Germany and UK, in which the CRC mortality rate in the UK was lower than Germany in 1990s-2000s (Ait Ouakrim et al., 2015).

In comparison with the CISNET models, DECAS' accurate performance in predicting the 17-year FS screening effect is closest to that of SimCRC. In the joint validation with 10-year results of the UKFSS trial, the CISNET model authors pointed out the overall length of dwell and sojourn time would affect the predicted screening effect, as this is the window period when the lesions can be detected and removed (Rutter et al., 2016). If the duration is shorter, the lesions might come back in individuals and grow to cancers during the post-screening follow-up. It potentially explained why MISCAN back then (with mean overall dwell and sojourn time 10.6 years) did not predict FS screening effect accurately, especially underestimating the incidence reduction (Kuntz et al., 2011; Rutter et al., 2016). In the 17-year UKFSS validation, despite still having the shortest dwell and sojourn time among the updated CISNET models, MISCAN was recalibrated to have a longer time of 17.2 years (IQR 4-18 years). The updated MISCAN improved the prediction and predicted stronger CRC incidence and mortality reduction than UKFSS trial (DeYoreo et al., 2020; Knudsen et al., 2020). This suggests the minimum dwell and sojourn time to prevent underestimation of screening effect could fall between 10.6 and 17.2 years. In this sense, DECAS equips the minimum length of overall dwell and sojourn time to predict screening effects accurately for a follow-up time up to 17 years. However, it is still unknown if the longer dwell time in DECAS would result in different screening effect forecast in longer term, which will require longer-term real-world data to assess.

4.1.3.2 Validation of FOBT and colonoscopy screening effect prediction

In the validation of stool-test screening effect, DECAS slightly overestimated the gFOBT effect in reducing CRC incidence and mortality, but generally with good precision – the 95% CIs from the trial covered the DECAS predictions. This gives the confidence for DECAS to predict the screening effect of FIT, which follows the same screening algorithm as gFOBT.

As for the validation of colonoscopy screening effect, DECAS's predictions were overall accurate, but slightly higher in CRC incidence reduction and slightly lower in CRC mortality reduction than the estimation in ESTHER study. Also, in ESTHER study, the reduction in proximal cancer incidence and mortality was not significant, whereas DECAS predicted the reduction in proximal cancers as equally effective as in distal cancers. Compared with another modeling study focusing on colonoscopy screening in Germany (Heisser et al., 2021a), which predicted 60-65% in incidence reduction and 75-80% in mortality reduction, the results from DECAS are comparable in the incidence reduction effect but slightly more conservative in the mortality reduction effect.

The colonoscopy protective effect against distal cancer has been well-documented, but that against proximal cancers is still uncertain (Brenner et al., 2014b). The discrepancy in the colonoscopy screening effect in proximal cancers observed between the ESTHER study and

DECAS prediction might stem from two factors. On one hand, proximal cancers develop more often from serrated lesions, which are more challenging for endoscopists to detect and remove (Zhao et al., 2019). This was reflected in the serrated lesion location distribution (more in proximal colon) and colonoscopy missing rates (slightly higher than adenomas) in DECAS. However, the authors of the meta-analysis that DECAS cited for the colonoscopy missing rates suggested that the serrated lesion missing rate might be underestimated given the data paucity (Zhao et al., 2019), i.e., the colonoscopy sensitivity to detect non-crSP and crSP could be even lower. Therefore, with a potentially overestimated colonoscopy sensitivity for serrated lesions, DECAS-predicted effect to reduce proximal cancer incidence and mortality might as well be overestimated.

On the other hand, the proximal cancer cases in the ESTHER study were small, which led to a high variability when assessing the proximal cancer incidence and mortality reduction, and the non-significance might be a result of the high variability (Guo et al., 2021). There are two other cohort studies analyzing the colonoscopy protective effect against proximal cancers, both with a follow-up time >15 years. One study in France only analyzed the incidence and found no significant reduction (HR 0.87, 95% CI 0.64-1.18) (Morois et al., 2014), and the other US study only looked at the mortality and showed significant reduction (HR 0.47, 95% CI 0.29-0.76) (Nishihara et al., 2013). Recently, a large observational study from Poland followed up with a cohort undergoing one-time colonoscopy for 17 years discovered that the screening colonoscopy significantly reduced proximal cancer incidence (HR 0.27, 95% CI 0.17-0.36) and mortality (HR 0.5, 95% CI 0.19-0.81). However, the reduction was only consistent in those receiving high-quality colonoscopies, which were defined as those with cecal intubation, adequate bowel preparation, and endoscopist ADR of 20% or greater. Nevertheless, the magnitude of protective effect (in terms of reduction in both incidence and mortality) against proximal cancers was still lower than distal cancers (Pilonis et al., 2020).

Regarding the absolute CRC incidence rate, unlike in the validation with UKFSS trial, DECAS predicted more distal cancers than proximal cancers, in alignment with the estimation from the ESTHER study. This is mainly because the definition of distal colon in ESTHER study included descending colon. The different definition of distal colon made DECAS allocated 50% of adenomas and 18% of the serrated polyps to these segments, contrasting the 40% adenomas and 4% serrated polyps in the UKFSS trial validation.

4.1.3.3 Lessons learnt from validation exercises

In summary, in the model cross-validation with CISNET models using the same benchmarks from the UKFSS trial, there were some minor differences among the four models. Nevertheless, DECAS

demonstrated the capability to accurately predict the FS screening effect on overall CRC incidence and mortality reduction, comparable to CISNET models. Same prediction accuracy from DECAS was also observed in the predicted screening effect of colonoscopy and gFOBT, with the caveat that DECAS had a slight tendency of overestimating screening effects on CRC incidence for colonoscopy and CRC mortality for gFOBT. Overall, this supports the use of DECAS to forecast the effect of German CRC screening program, in which colonoscopy and FIT are used.

In addition, taking together the observations in the validation with the UKFSS trial and the ESTHER study, it appears that DECAS predicts a universal colonoscopy protective effect against both distal and proximal cancers, while the current clinical evidence suggests the protective effect is likely to be weaker against proximal cancers. Therefore, caution needs to be taken when using DECAS to dive into the screening effects for cancer sub-groups in different colon segments, especially in predicting the effect against proximal cancers.

4.2 Cost-effective analysis of German CRC screening program

4.2.1 Colon cancer treatment costs in Germany

(Part of Chapter 4.2.1 has been published (Cheng et al., 2021)*)*

The mean net colon cancer treatment costs by phase showed a "U-shaped" pattern, consistent with the cancer treatment cost estimations in the literature (Yabroff et al., 2008; Haug et al., 2014; Laudicella et al., 2016). Among the literature, one study used comparable methods to analyze another German SHI dataset to estimate the mean net colorectal cancer treatment costs. Their cost estimates for the initial, continuing and terminal phases were €33,535, €2,967 and €66,683 (inflated to 2021 Euro) (Haug et al., 2014), which were higher than the results in this thesis. However, the former study did not match for comorbidity in the control group. In contrast, the cost analysis in the present thesis matched for comorbidity, and it consequently resulted in higher treatment costs in the control group (Sarfati et al., 2016), which, in turn, narrowed the cost difference between patient and control groups and led to lower net treatment costs.

Diving into the sub-group analysis of net colon cancer treatment costs by cancer severity, which are used to inform the costs in the cost-effectiveness analysis for the German CRC screening program, the results are consistent with the literature that late stage cancers result in higher net treatment costs (Yabroff et al., 2008; Knudsen et al., 2012). Interestingly, the continuing phase net costs in patients with low cancer severity are negative, which means the treatment costs for the matched control are even higher. This might again be the results of matching comorbidities for the control in the present thesis.

4.2.2 Modeled benefits, burden and harms compared to literature

In the literature, there are two recent modeling studies evaluated various CRC screening strategies with similar efficiency measurements which can be used to compare with the results in the present thesis. One is the most recent modeling analysis by the three CISNET models to inform the USPSTF CRC screening recommendations (Knudsen et al., 2020, 2021) (hereafter, the USPSTF modeling study). They evaluated over 200 CRC screening strategies, including the majority of the strategies evaluated in the present thesis. The efficiency measurement used in the USPSTF modeling study is number of extra colonoscopies needed to save one additional LY, although not exactly the same as the measurement used in the present study, it still provides valuable information for comparison. The other one is an Australian modeling study by Lew et al. (2018) using Policy1-Bowel, one of the other models considering serrated polyp pathway (hereafter, Policy1-Bowel study). They also evaluated biennial FIT and 10-yearly colonoscopy, both starting from age 50, and used the same efficiency assessment as in the present thesis, which serves a valuable source for comparison.

4.2.2.1 Benefits under perfect adherence assumption

In summary, all strategies yielded positive benefits compared to no-screening: increased life expectancy (mean 57-97 LYG per 1,000 40-year-olds), increased quality-adjusted life expectancy (mean 36-98 QALYG per 1,000 40-year-olds), and reduction in lifetime CRC cases (mean 18-43 cases prevented per 1,000 40-year-olds) and deaths (mean 14-22 deaths prevented per 1,000 40-year-olds).

Compared the outcomes from DECAS with the USPSTF modeling study (Knudsen et al., 2021), the benefits estimated by DECAS are generally lower. For examples, the LYG in the comparable strategies from the USPSTF analysis (including FIT, FIT-colonoscopy and colonoscopy screening from the age of 45 or 50) ranged 256-361 per 1,000 40-year-olds and the CRC deaths averted ranged 21-28 deaths per 1,000 40-year-olds. The more conservative estimation from DECAS in mortality reduction has been observed in the cross-validation exercise against UKFSS trials (Chapter 3.3.1), which estimated only 17-year effects. Therefore, it is no surprise that the gap between the benefit estimates from CINSET models and DECAS widens when the forecasting period extends to lifetime. However, if considering the 95% CrIs from DECAS, deaths prevented are comparable to the USPSTF analysis, but values of LYG were still around half of the results from the USPSTF analysis. The pure colonoscopy or FIT-colonoscopy combined strategies in the USPSTF study prevented more CRC deaths compared to FIT alone in the same screening period, which is consistent with the findings from DECAS.

In comparison with Policy1-Bowel study (Lew et al., 2018), the mortality reduction predicted in the Australian study for biennial FIT and 10-yearly colonoscopy starting from age 50 years old was 74% and 78% respectively, in comparison with the DECAS prediction of 52% and 74%. The two results in mortality reduction by 10-yearly colonoscopy are comparable. In contrast, the higher mortality reduction predicted by Policy-1 Bowel is likely to derive from the higher FIT sensitivity assumption to detect non-advance adenoma as well as the fact that DECAS only allows next FIT screening following a negative FIT-positive colonoscopy in 10 years.

Additionally, Chen et al. (2019) used a Markov model to estimate the protective effects of screening colonoscopy in the German context. Their model is driven by transition probabilities between adenoma and CRC states estimated from the data of German screening colonoscopy program, the same source as DECAS. Under 100% adherence, Chen et al. predicted 82% and 94% mortality reduction from twice and three times 10-yearly colonoscopy starting at age 50, respectively. In contrast, the mortality reduction rates were 74% and 76% in DECAS estimation. There are no known validation studies done with the Markov model, hence, it is difficult to verdict the prediction power of the Markov model. Given that DECAS tends to underestimate the mortality reduction by screening colonoscopy as seen in the validation exercise with the ESTHER study, the gap between the predictions from the two models could be narrower in reality. Another explanation to account for the difference could be that the Markov model did not apply different colonoscopy sensitivities for different type of lesions. Instead, they used the proportion of detected cases from the screening registry and applied the proportion directly to the cohort model (Chen et al., 2019). The cohort approach and lesion-based approach might result in different detection rates as well as the protective effects by screening.

4.2.2.2 Benefits under imperfect adherence assumptions

The Australian (Lew et al., 2018) and German (Chen et al., 2019) modeling studies also estimated the benefits under different participation rates. Policy1-Bowel (Lew et al., 2018) assumed 2 other scenarios: one with 29% for FIT and 15% for colonoscopy, and the other 57% for FIT and 35% for colonoscopy. The assumptions are close to the Scenario 3 and 4 in the present CEA. The mortality reductions for biennial FIT starting from age 50 in the two scenarios were 36% and 51% in Policy1-Bowel and 18% and 35% in DECAS. On the other hand, the mortality reductions for 10-yearly colonoscopy starting from age 50 in the two scenarios were 16% and 34% in Policy1-Bowel and 27% and 39% in DECAS. Similar to the perfect adherence scenarios, DECAS predicted lower FIT mortality reduction effect but higher colonoscopy mortality reduction effect than Policy1-Bowel.

Despite that the Markov model by Chen et al. (2019) predicts much more mortality reduction effects for twice or three-time 10-yearly colonoscopy screening under 100% participation rates, their estimation with 25% participation rate is closer to those in Scenario 3 of the present thesis. Under 25% participation rate, Chen et al. predicted about 20% and 22% of CRC mortality reduction resulted from twice and three-time 10-yearly colonoscopy from the age of 50, respectively, in comparison with 26% and 27% reduction in DECAS prediction using a similar participation rate.

4.2.2.3 Burdens and harms under perfect adherence assumption

In terms of lifetime number of colonoscopies used, DECAS predicted similar results as in the USPSTF modeling analysis (Knudsen et al., 2021). The number of colonoscopies used per 1,000 40-year-olds in the 10-yearly colonoscopy strategy starting at age of 45 was 2,825-3,700 in DECAS and 3,679-3,782 in USPSTF analysis. In biennial FIT strategy starting at age of 45, the number was 737-1,280 in DECAS and 1,147-1,361 in USPSTF analysis. On the other hand, DECAS estimated a slightly lower number of colonoscopies used per 1,000 40-year-olds when compared with the Policy1-Bowel study (Lew et al., 2018): for 10-yearly colonoscopy starting at age of 50, the number was 2,084-2,745 in DECAS and 3,001 in Policy1-Bowel; for biennial FIT from age of 50, the number was 609-1,043 in DECAS and 1,105 in Policy1-Bowel. The causes for the differences are difficult to disentangle, as the colonoscopy usage is the jointly determined by test sensitivity, screening and surveillance colonoscopy management flow in the models.

All strategies were predicted by DECAS to be cost-saving compared to no-screening. Interestingly, all screening strategies resulted in higher lifetime costs than no screening in the Australian study estimated by Policy1-Bowel (Lew et al., 2018). One potential reason for the high costs in Policy-1 Bowel model could be that it assumes three times higher costs of the treatment for colonoscopy complications and five times higher colonoscopy complication rates, in conjunction with slightly higher estimated number of colonoscopies used. Moreover, the total CRC treatment costs in CRC stage 3-4 in DECAS could be higher than in Policy1-Bowel, as the treatment costs in DECAS depend on overall survival length, in which the costs in continuing phase could accumulate. With higher CRC treatment costs, the higher number of CRC cases result in higher lifetime costs in no-screening strategy. On the other hand, the results from DECAS are consistent with other estimation done in the US context: compared with no-screening, screening with biennial FIT and 10-yearly colonoscopy is cost-saving (Ran et al., 2019).

Regarding the number of complications from colonoscopy, given that colonoscopy complication rates used in DECAS are lower than those used in the USPSTF modeling analysis (Knudsen et al., 2021), the number of complications also reflected the difference. For instance, there were 1-8

complications per 1,000 in the DECAS prediction compared to 15-16 complications per 1,000 in USPSTF analysis for 10-yearly colonoscopy screening starting at age 45.

4.2.3 Sex-differentiated starting age for screening colonoscopy

The results comparing FIT1y50+COL10y55 and mCOL50/fFIT50+COL55 demonstrated that the sex-differentiated screening strategy could be a better option in terms of cost-effective, even at a lower participation rate setting.

There has been extensive evidence suggesting men have a higher risk of developing CRC and advanced adenoma (Nguyen et al., 2009), which could be explained by the biological differentiation between the sexes, e.g., level of testosterone (Amos-Landgraf et al., 2014). Brenner et al. (2010) and Ferlitsch et al. (2011) analysed the data collected from screening colonoscopy participants in Germany and Austria and concluded that the prevalence of AA and CRC could reach the same level in men 5-10 years earlier than women. Specifically, the Austrian study included screening participants between the age of 45 and 55, and the prevalence of AA in men aged 45 to 49 years was comparable to that in women aged 55 to 59 years (Ferlitsch et al., 2011). Similarly, in a pilot study in Germany extending screening colonoscopy offers to both men and women from age of 50 years, the prevalence of CRC and AA together were found to be 8.6% in men and 4.5% in women (Brenner et al., 2017). These epidemiological evidence supports sex-specific recommendations on the starting age of screening colonoscopy, which already took place in Germany in 2019.

Under perfect adherence assumption, our model predicted similar mortality reduction effects in both sexes when sex differentiated strategy was applied. The results are similar to the findings from Heisser et al. (Heisser et al., 2021b), who used a Markov model to evaluate the screening strategies in Germany, and the mortality reduction in men receiving colonoscopy at 50 and 60 years old was almost en par to women receiving annual FIT from 50 to 54 years and colonoscopies at 55 and 65 years (81% and 82%, respectively).

In regard with the other cost-effectiveness evidence on sex-differentiated strategy, there is no literature available for comparison with the present thesis. The closest simulated results for sex-specific strategies are provided by Wong et al. (2016) in the context of Hong Kong. In their Markov simulation study, sex-specific strategies were found to be the most cost-effective, however, they recommended purely 5-year FS for women and purely 10-yearly colonoscopy for men, both starting at age of 50 years. Despite not evaluating the same sex-differentiated strategies, the evidence from Wong et al. (2016) and the present thesis suggests that providing alternative

screening test other than colonoscopy or a later colonoscopy starting age to women given the lower risk could be more cost-effective than colonoscopy to both sexes at the same age.

4.2.4 Efficient strategies

4.2.4.1 Under perfect adherence assumption

When the adherence to screening and surveillance is perfect, the three pure colonoscopy screening strategies (COL10y50, COL10y50-3X, and COL10y45-3X) are the efficient strategies or only slightly dominated in terms of cost-effectiveness. However, if NNC to prevent one additional CRC death is used to evaluate efficient strategies, all 10 strategies are either efficient or only weakly dominated.

Comparing the results from the present thesis with the USPSTF modeling study (Knudsen et al., 2020, 2021), the results are mostly comparable within strategies using the same test modality. The three pure colonoscopy strategies evaluated in the present thesis are also shown as efficient strategies in the USPSTF modeling analysis, as well as the FIT-colonoscopy combined strategies. However, the USPSTF modeling study did not directly compare across the strategies using different screening modalities, therefore, there is no direct information on the efficiency between colonoscopy and FIT strategies to allow the comparison with DECAS. The only certain information is that colonoscopy strategies require more colonoscopies but reduce more CRC deaths than FIT strategies in USPSTF study, which is consistent with the findings from DECAS.

In the Australian modeling study (Lew et al., 2018), when evaluated with LYG and costs, Polocy1-Bowel deemed annual and biennial FIT starting from age 50 years as the efficient strategies. This is foreseeable given the higher estimated benefits from FIT strategies along with the higher lifetime costs in colonoscopy strategies in Policy1-Bowel, which makes FIT strategies more cost-effective than colonoscopy strategies – a reverse picture compared to the results shown in DECAS forecast. When using NNC as the efficiency measurement, both 10-yearly colonoscopy and biennial FIT were deemed as efficient in Policy1-Bowel, which is consistent with the results in the present thesis.

In the other and the only CEA so far in the German context by Ladabaum et al. (2014), the authors also plotted the strategies on a cost-effectiveness plane so that the efficient frontier can be drawn. The four strategies on the efficiency frontiers were 10-yearly colonoscopy at age 60 and 70, annual FIT plus 10-yearly colonoscopy at age 55 and 65 or 60 and 70, and annual FIT starting at 50 years old followed by biennial FIT. The findings agreed with those from DECAS that strategies involving colonoscopies tend to be more efficient even though the starting age in Ladabaum et

al.'s study was a bit later. On the other hand, they also predicted the pure FIT strategy to be efficient, which is opposite to what DECAS found. The discrepancy might derive from their lower CRC death rate (hence higher QALYs given no disutility was assumed) and lower treatment costs in the FIT strategies than in colonoscopy strategies.

4.2.4.2 Under imperfect participation rates

When using LYG and costs as measurement to assess the efficient strategies under imperfect scenarios, all 3 scenarios with various suboptimal participation rates for FIT and colonoscopy yielded 4 consistent efficient or near-efficient strategies, including the 3 pure colonoscopy strategies and FIT1y45+COL10y50. When assessing with QALY, FIT1y45+COL10y50 was not consistently on the efficient frontier, only in Scenario 4, where FIT participation rate was almost double the colonoscopy participation rate. The reason is that waiting for the FIT-positive colonoscopy given a positive FIT result was assumed to cause disutility. Overall, purely FIT strategies were strongly dominated by pure colonoscopy or combined strategies both in terms of LYG or QALYG with costs.

Lew et al. (2018) designed similar scenarios with different participation rates in the Policy1-Bowel study. However, in the two imperfect adherence scenarios in the Australian study, only annual and biennial FIT strategies were regarded as efficient, which deviated from the outcomes in the present thesis. The reasons are likely to be the same as for the perfect adherence scenario. To note, in the low adherence scenario in the Policy1-Bowel study, 10-yearly colonoscopy and a combined strategy involving colonoscopy were very close to the efficient frontier. It is similar to that in the DECAS analysis, where in Scenario 2, other colonoscopy and combined strategies were near the efficiency frontier.

Opposite to the reverse picture on the efficiency regarding the pure FIT and colonoscopy strategies using LY and costs, when using the NNC as the basis, both DECAS and Policy1-Bowel models agree that FIT strategies are efficient. However, Policy1-Bowel also found 10-yearly colonoscopy on the efficiency frontier, whereas DECAS deemed it to be dominated by FIT or combined strategies.

4.2.4.3 Screening strategies starting from 45 years old: the most cost-effective

Although looking at the efficiency frontiers produced by the mean values of each strategy, many of the colonoscopy or combined strategies were deemed efficient in terms of cost-effectiveness, regardless of the starting age. However, when considering the CEAC across different WTP, it is

obvious that the colonoscopy or combined strategies with a starting age from 45 years are consistently being the most cost-effective options across all four scenarios.

The cost-effectiveness findings in the present thesis align with the major updated CRC screening recommendations in the recent years, which recommend CRC screening to start from the age of 45 (Wolf et al., 2018; Shaukat et al., 2021; US Preventive Services Task Force, 2021). In the modelling study informing the 2021 USPSTF CRC screening recommendations (Knudsen et al., 2021), 41 out of the 49 efficient strategies started at age of 45 years. Taking 10-yearly colonoscopy for example, the USPSTF modelling study estimated that if moving the starting age from the age of 50 to 45, it would provide 27 extra LYs, reduce additional 3 CRC cases and 1 CRC death, but add 2 more colonoscopy complications and 784 colonoscopies per 1,000 40-year-olds. In comparison, if compare COL10y50-3X and COL10y45-3X in the DECAS forecast, there could be in average 14 more LYs, 2 fewer CRC cases, 1 fewer CRC death, 0.5 more colonoscopy complication, and only 283 more colonoscopies per 1,000 40-year-olds. Both models showed that the extension of CRC screening 5 years earlier results in more benefits with reasonable extra burden and harms.

Another well-validated Markov model by Ladabaum et al. (2019) was also used to evaluate the impact of extending the CRC screening to the age of 45. In this Markov study, the ICERs of 10-yearly colonoscopy 45-75 years old and annual FIT 45-49 years old followed by 10-yearly colonoscopy 50-75 years compared to 10-yearly colonoscopy 50-75 years were $33,900 and $2,500 per QALYG, respectively (in 2018 USD). The ICER for COL10y45-3X vs. COL10y50-3X in DECAS estimation was €2,017 per QALYG (in 2021 Euro), but FIT1y45+COL10y50-3X was dominated by COL10y50-3X, given the disutility assumption for FIT. If calculated by LYG, the ICER between FIT1y45+COL10y50-3X and COL10y50-3X would be €19,860 per LYG (2021 Euro). Overall, both models again predicted the strategies starting from age of 45 are cost-effective than those starting from the age of 50 if assuming the WTP threshold at $50,000 per LYG or QALYG.

4.2.5 Influence of invitation methods and corresponding participation rates on screening cost-effectiveness

Despite at least 10 percentage points increase in FIT participation rates compared to Scenario 2, the pure FIT strategies were still strongly dominated by the strategies involving colonoscopies in Scenario 3. If focusing on the more efficient strategies starting at age 45 years, noticeable improvement in terms of benefits (especially LYs) in FIT1y45+COL10y50 and FIT1y45+COL10y50-3X relative to COL10y45-3X can be seen, although they came along with an increase of overall costs. This yielded an ICER of €68,273 per LYG between the two efficient strategies, FIT1y45+COL10y50 and COL10y45-3X, in Scenario 3.

A major part of the cost increase in Scenario 3 comes from the mail-out but unused FITs, which contained the postage and FIT kit costs. The wasted resource costs in both FIT1y45+COL10y50 and FIT1y45+COL10y50-3X strategies were around €44,000 in lifetime per 1,000 40-year-olds (discounted with 3% annual rate). If the mailed-out FITs were not wasted, the ICER between FIT1y45+COL10y50 and COL10y45-3X in Scenario 3 would have reduced from €68,273 per LYG to €3,769 per LYG, making the efforts to increasing FIT participation rates more cost-effective.

In Scenario 4, a high adherence scenario was designed to mimic the invitation approach implemented in the Netherlands, which is among one the few European countries with a participation rate >70% in the organized biennial FIT screening program (Toes-Zoutendijk et al., 2020). In addition, the participation in colonoscopy screening was also increased to 42%, informed by a US study (Singal et al., 2017) evaluating the invitation approach on colonoscopy screening. As expected, with the FIT participation rates increased more than two folds, both the number of wasted FITs and costs were reduced to about 40% of the figures in Scenario 3. The strong increase in FIT participation and moderate elevation in colonoscopy participation also improved the cost-effectiveness of FIT1y45+COL10y50 comparing to COL10y45-3X: the ICER in scenario 4 was €5,470 per LYG and €29,774 per QALYG.

In a randomized study by Gruner et al. (2020), the authors not only examined the effect of mail-out FIT, but also the effect of an alternative option, in which only a FIT request form is enclosed together with the invitation letter. Participants who are willing to receive a mail-out FIT can reply via an online-form, e-mail, or return letter. Interestingly, the 1-year participation rates between the alternative approach and direct mail-out FITs method were very close: 31.9% vs. 33.5%. In a systematic review, although only a few trials were found to examine the same alternative, the review revealed consistent results that the FIT request option could lead to a similar usage rate as directly mailing out FIT (Gruner et al., 2021). According to the estimation in the Scenario 3 of the present analysis, the alternative approach could then provide similar level of benefits and potentially save up more than €40,000 per 1,000 40-year-olds thanks to avoiding unused posted FIT kits, which would make the FIT-colonoscopy combined strategies even more attractive cost-effectively.

4.3 Strengths, limitations, and future research direction

4.3.1 Study strengths

There are several strengths in the studies of the present thesis. First, DECAS is the only CRC microsimulation model in the literature additionally considering serrated neoplasia pathway and

at the same time calibrated with a Bayesian approach. Although DECAS is not the first model accounting for the second important CRC tumorigenesis pathway, the serrated neoplasia pathway, it is still among the only three such models in CRC screening simulation (the other two being ASCCA (Greuter et al., 2014) and Policy1-Bowel (Lew et al., 2017)). This nature positions DECAS among the limited number of models capable of estimating the implications of serrated polyps, which are usually located in proximal colon and more difficult to detect under endoscopy, in CRC screening.

Additionally, DECAS stands out from the other two serrated polyp models in a way that it is calibrated with a Bayesian method, APMC (Lenormand et al., 2013), using a large screening dataset from the German screening colonoscopy program in 2003-2014 (Pox et al., 2012). The Bayesian calibration outputs allow DECAS to work with 1,000 sets of posterior parameters, which help capture the parameter uncertainty and provide a distribution of inputs to be used in PSA. Also, given that the priors used in Bayesian calibration require information from observed data of good quality, the long-term and large amount of screening data in Germany serve as a solid basis to inform the priors and as strong anchoring points in Bayesian calibration.

Secondly, DECAS has been robustly validated against randomized control trials of single FS and biennial FOBT, as well as a cohort study of screening colonoscopy. In addition, cross validation with the three most widely used CRC models, CISNET models, was also done. These exercises confirmed the prediction validity of DECAS as well as pointed out the caveats to interpret the screening results when comparing to other modeling studies in specific screening modalities. This also gives the confidence to trust the CEA outputs from the vetted DECAS.

Thirdly, this thesis provides the first extensive cost-effectiveness analysis on the current German screening program since the change into an organized framework in 2019 and on various combinations of FIT and colonoscopy screening strategies as well as starting ages. The modeling study provides a lifetime perspective to elucidate the effectiveness and cost-effectiveness of CRC screening in Germany, which supplements the to-date 16-year German colonoscopy screening program data.

No studies prior to this thesis assessed the change in CRC protective effects and cost-effectiveness when screening colonoscopy is extended 5 years earlier only for men. The results of present analysis strengthened the evidence from a health economic perspective to support the sex-differentiated screening strategy introduced in Germany in 2019.

Furthermore, no studies to date in the German context have explored if screening starting at age 45 can strike a good balance between benefits and burdens/harms as have been elucidated in the US context. This is especially important against the background that the CRC incidence among

younger people aged 20-49, including the population in Germany, has been increasing in the recent decade (Siegel et al., 2019). The results of the present thesis are in line with the recent changes in major CRC screening recommendations (Wolf et al., 2018; Shaukat et al., 2021; US Preventive Services Task Force, 2021), and this can serve a basis to spark a discussion if similar changes should be made in Germany.

Lastly, the present thesis presents the first economic analysis to quantify the wasted resources due to unused mailed-out FIT. This piece of evidence could be used in conjunction with the findings from the randomized trials by Gruner et al. (2020), in which they found the option to let participants request for mail-out FIT to have similar effect to boost participation rate as seen in the direct mail-out FIT approach. Taken together, the option to request for mail-out FIT could be a cost-effective way in Germany to enhance the FIT participation rate.

4.3.2 Study limitations and future research direction

There are a few limitations in the present thesis. Firstly, in the DECAS calibration, only the assumption that 15% of the CRCs developing through the serrated pathway was used. The decision was taken so that DECAS predictions can compare with other studies under the same assumption. However, given the uncertainty around the proportion of CRC arising from the serrated pathway is high (15-30% or even higher) (Bettington et al., 2013; Ijspeert et al., 2015), and it is a critical parameter influencing the estimation of screening effectiveness in CRC screening simulation, DECAS will need to explore calibration with different proportions in the future.

Secondly, as extensively discussed in the challenges for DECAS development in Chapter 4.1.2, there are still boundary-hitting issues for some parameters and deviation in prevalence calibration of precancerous lesions at late ages. Besides, the 95% CrIs of the DECAS predictions are generally wide. Improvement to these pitfalls could potentially be achieved if one can explore the priors with ranges which are extensively wider than literature suggests. This might allow a full search of the parameter spaces, and hopefully, even better fitted sets of posterior samples. However, this will take considerable amount of time to execute – the present study already spent about 6 months to look for proper prior spaces to achieve the current global fit in the APMC algorithm. Future endeavors can explore if there are more efficient algorithms developed in other disciplines and adapt one of those to further improve the calibration in DECAS. Additionally, re-calibration might be needed if more robust data on serrated lesion prevalence are available in the future.

Thirdly, the computation time is a non-negligible limiting factor for the ABC calibration, and trade-off needed to be made between computational costs and posterior sample quality. According to

the authors (Lenormand et al., 2013), the quality of the approximation in APMC can be improved with smaller α and $p_{acc_{min}}$. However, it also means more simulation steps are required during the calibration, and thus much longer computation time. With the current setting in this thesis, it already took 10 days even with parallel computing on a 60-core cluster computer. Given that the fitting to the calibration targets and external validation yielded satisfactory results for the posterior samples, one can argue that a good balance was reached between the quality of approximation and the expense of computational time. If there is more advanced computational power in the future, one could consider recalibrating DECAS with more refined APMC parameters.

Fourthly, the screening participation rates assumed in Scenario 2 and 3 in the CEA were based on the real CRC screening participation rates recorded until 2017 and the participation rate increase found in two German RCTs (Hoffmeister et al., 2017; Gruner et al., 2020). It is still unknow if the real-world participation rates under the organized screening program behaves as in the RCTs. Therefore, a repeated scenario analysis should be conducted when the real-world participation rates post-organized CRC screening program is available, and it would provide an even more accurate economic evaluation of the German organized CRC screening program.

Lastly, in order to have a more focused research objective, the present thesis only assesses 10 strategies involving FIT and colonoscopy screening. Nevertheless, given the flexibility of DECAS, evaluation of other screening tests and modalities can also be considered in the future. For example, although not currently recommended in Germany, other major US CRC screening recommendations (Wolf et al., 2018; Shaukat et al., 2021; US Preventive Services Task Force, 2021) include annual to 3-yearly mtsDNA and 5-yearly CTC, which can also be evaluated together with the FIT and colonoscopy modalities used in the German context. Furthermore, the strategy with sex-differentiated screening starting age was shown to be cost-effective in comparison to that of the same starting age for both sexes in the present thesis, however, only colonoscopy for men starting at 50 years and women starting at 55 years was considered. Given screening staring at 45 years for both sexes was found to be the most cost-effective, sex-differentiated strategies starting at age 45 years could also be evaluated in the future. Ultimately, other than using sex as a risk-stratified factor to design and evaluate strategies, DECAS could also be adapted to play a possible role to assess other risk-based approaches, e.g., the *a priori* risk of developing CRC in one's lifetime considering family history or genetic information.

5 CONCLUSIONS

This doctoral thesis aimed to develop an individual-level CRC natural history and screening model to evaluate the cost-effectiveness of German CRC screening program. It started with the development of DECAS, a DES CRC natural history model, which accounts for both CRC carcinogenesis pathways – adenoma-carcinoma and serrated neoplasia pathways – and is calibrated with a Bayesian method to better tackle the parameter uncertainty. The two features make DECAS a unique model within the realm of CRC modeling. DECAS also withstood several external validations with large CRC screening trials and a cohort study, which consolidated its prediction credibility to assess CRC screening effectiveness and cost-effectiveness.

Being calibrated with the large dataset from the German screening colonoscopy registry, DECAS is the first CRC microsimulation model tailored for the German context, therefore, suitable to evaluate the current CRC screening program in Germany as well as alternative strategies and scenarios to inform future CRC screening policies. Through DECAS, the present thesis was able to evaluate benefits, burdens, and harms of CRC screening from a lifetime perspective, which are the evidence that current real-world clinical trials with limited time span are unable to provide. The analysis in the German context confirmed that CRC screening, regardless of modality, frequency and starting age, delivers good protective effects compared with no-screening, given the differences in benefits among various screening strategies are much smaller in comparison to those between screening and no-screening.

Additionally, the present thesis concluded that the recently implemented strategy with sex-differentiated starting age for screening colonoscopy strikes a good balance between the benefits, burdens, and harms under the current participation rate in Germany, both in terms of cost-effectiveness as well as number-needed-to-colonoscope to prevent one additional CRC death. Strategies involving colonoscopy, either pure or combined strategies following annual FITs, were found to be more cost-effective than pure FIT strategies in Germany based on currently available data. Colonoscopy or FIT-colonoscopy screening strategies starting from the age of 45 years appeared to be the most cost-effective even under the sensitivity tests with different willingness-to-pay thresholds or different discount rates. With the mail-out FIT invitation approach and assumed higher participation rates, it increased screening benefits but also the costs in the strategies involving FITs due to non-negligible additional costs incurred by unused posted FITs. The modeling study also confirmed that, higher participation rates resulted in higher benefits in CRC screening and fewer waste from unused mailed-out FITs, which in turn enhanced the cost-effectiveness of screening. Even with extensive sensitivity analyses varying different input assumptions, these findings still withhold.

To conclude, the modeling evidence from the present thesis can, despite the uncertainty, serve as a basis to inform future CRC screening policy-making in Germany in the absence of long-term FIT and colonoscopy screening effectiveness evidence from clinical trials. As a flexible platform, DECAS can be further used in providing modeling evidence for various screening modalities (e.g., mtsDNA or CTC) or risk-stratified screening strategies (e.g., with *a priori* individual risks), either in the German context or other geographic regions pending adaptation to the local CRC epidemiology. Last but not least, DECAS can also be used to analyze the public health and economic impacts of delayed CRC screening due to disruption by external forces, e.g., the COVID-19 pandemic.

6 SUMMARY

Colorectal cancer (CRC) screening has been shown to contribute to the reduction in CRC incidence and mortality. To inform the CRC screening recommendations, it usually relies on models that are flexible to predict the effectiveness of various screening modalities and strategies from a lifetime perspective. To date, there are only two CRC microsimulation models considering the two CRC carcinogenesis pathways (adenoma-carcinoma and serrated neoplasia pathways). However, both are calibrated with grid search optimization methods, limiting their capability to account for parameter uncertainty. Furthermore, there is no cost-effectiveness analysis from a microsimulation model to assess the current German CRC screening program and to explore alternative strategies under different invitation approaches to improve screening participation.

The aims of the present thesis are two-fold: (1) To construct an individual-level model encompassing both CRC development pathways, and to explore a Bayesian calibration method for CRC disease modeling; (2) To conduct an up-to-date cost-effectiveness analysis for evaluating the cost-effectiveness of various CRC screening strategies in the current German organized CRC screening program, and to thereby inform future CRC screening policies in Germany.

A discrete event simulation model, DECAS, was thus developed in the R software. DECAS simulates the CRC natural history from the state of no lesions to precancerous lesions (adenoma or serrated polyps) and to pre-clinical and clinical CRCs in individuals with an average CRC risk and follows them up from the age of 20 to 90 or death, whichever occurs first. The rates of event happening were lesion-, age-, sex- and location-specific, and they were calibrated with a likelihood-free approximate Bayesian computation method, adaptive population Monte Carlo (APMC). The calibration took advantage of 74 prevalence data points from the German screening colonoscopy program, which consisted of 5.2 million average-risk screening participants in 2003-2014. The Bayesian calibration rendered 1,000 sets of posterior parameter samples, with which DECAS successfully reproduced the CRC incidence data from the German national cancer registry.

After DECAS natural history model validation, the screening component was added to the DECAS model. If any lesions prior to the clinical cancer state are detected by the screening tests, individuals can be referred to or directly removed by colonoscopy. To further validate the predictive ability of DECAS regarding the CRC screening effects, external validations against two large randomized control trials on flexible sigmoidoscopy and guaiac fecal occult blood test and a large colonoscopy cohort study were performed. Additionally, cross validation against the three most widely used CRC screening models, the CISNET models, was conducted. DECAS demonstrated accurate predictions for CRC incidence and mortality reduction in the validation studies.

The validated DECAS model was then used to evaluate the benefits, burdens, and harms of CRC screening strategies in Germany, including annual fecal immunochemical tests (FIT) for aged 50-54 years followed by two 10-yearly colonoscopies or biennial FIT from age 55-75 years for both sexes, and the new strategy allowing men to start the two 10-yearly colonoscopies from the age of 50 years. Alternative strategies including biennial FITs or 10-yearly colonoscopies from the age of 45 or 50 years, and combined strategies with annual FIT from the age of 45 followed by 10-yearly colonoscopies from the age of 50 were also evaluated. All strategies were evaluated under four scenarios: perfect adherence, low adherence under the current organized program with an invitation letter, improved adherence with an invitation letter and mail-out FITs, high (but imperfect) adherence with an invitation letter, mail-out FITs, and an additional reminder.

All strategies were found to be cost-effective compared to no-screening across all four scenarios. Assuming perfect adherence and compared to no-screening, the screening strategies brought about a 34-75% CRC incidence reduction, a 52-80% CRC mortality reduction, 57-97 life-years gained, and 36-98 quality-adjusted life-years gained per 1,000 40-year-olds. All strategies were cost-saving, and they resulted in 809-3,240 colonoscopies needed and 1-4 colonoscopy complication cases per 1,000 40-year-olds. In scenarios with imperfect adherence, the benefits, burdens, and harms decreased with the participation rates. In the two mail-out FIT scenarios, the sent but unused FITs could amount up to 9,967 kits and caused an additional cost of €93,323 per 1,000 40-year-olds in the biennial FIT strategy starting at age 45 in the lower adherence scenario.

Additionally, the strategy with sex-differentiated starting age for colonoscopy appeared to be more cost-effective than the equal-starting-age strategy. Both pure colonoscopy and FIT-colonoscopy combined strategies appeared to be more cost-effective than pure FIT ones. Three-time 10-yearly colonoscopies strategy starting from the age of 45 was deemed the most cost-effective across scenarios given the willingness-to-pay thresholds of €5,000-100,000. Overall, strategies starting from the age of 45 provided the best balance between benefits, burdens, and harms, which is consistent with recent recommendation changes from major US guidelines.

The modeling evidence from the present thesis can, despite the uncertainty, serve as a basis to inform future policy making for CRC screening in Germany in the absence of long-term evidence for FIT and colonoscopy screening from clinical trials. Future research directions include a recalibration of DECAS with more efficient Bayesian algorithms and with more robust serrated polyp data when available. Moreover, the cost-effectiveness for more risk-stratified screening strategies other than sex-specific ones (e.g., with a priori individual risks) and alternative screening modalities (e.g., multitarget stool DNA test or computed tomography colonography) can be explored. Lastly, DECAS can also be used to analyze the public health and economic impacts of delayed CRC screening due to disruption by external forces, e.g., the COVID-19 pandemic.

7 ZUSAMMENFASSUNG

Darmkrebsfrüherkennung tragen nachweislich zur Senkung der Inzidenz und Mortalität von Darmkrebs bei. Als Basis für Empfehlungen für Darmkrebsfrüherkennung sind Modelle nötig, die auf flexible Art und Weise die Effektivität verschiedener Screening-Modalitäten und -Strategien aus einer Lebenszeitperspektive heraus vorhersagen können. Bis heute gibt es nur zwei Darmkrebs-Mikrosimulationsmodelle, die die beiden Darmkrebs-Karzinogenesewege (den Adenom-Karzinom-Weg und den serratierten Neoplasie-Weg) berücksichtigen, wobei beide mit Gittersuch-Optimierungsmethoden kalibriert sind, was ihre Fähigkeit einschränkt Parameterunsicherheiten zu berücksichtigen. Außerdem gibt es keine Kosteneffektivitätsanalyse durch ein Mikrosimulationsmodell, um das gegenwärtige deutsche Darmkrebs-Früherkennungsprogramm zu bewerten und um alternative Strategien für verschiedene Einladungsschemata zu untersuchen, die eine bessere Teilnahme am Programm bewirken.

Die Ziele der vorliegenden Arbeit sind zweierlei: (1) Die Konstruktion eines Modells auf Individualebene, das beide Darmkrebs-Entwicklungswege umfasst, und die Untersuchung einer bayesschen-Kalibrierungsmethode für die Modellierung von Darmkrebsfrüherkrankungen; (2) die Durchführung einer aktuellen Kosteneffektivitätsanalyse für verschiedene Darmkrebs-Screening-Strategien im gegenwärtigen deutschen organisierten Darmkrebs-Screening-Programm und damit die Bereitstellung von Informationen für zukünftige Darmkrebs-Screening-Richtlinien in Deutschland.

Dazu wurde ein diskretes Ereignissimulationsmodell (DECAS) in der Software R entwickelt. DECAS simuliert den natürlichen Verlauf des Darmkrebses vom Zustand ohne Läsionen über präkanzeröse Läsionen (Adenome oder serratierte Polypen) bis hin zu präklinischem und klinischem Darmkrebs in Individuen mit durchschnittlichem Darmkrebsrisiko und verfolgt diese von 20 bis 90 Jahren oder bis zum Tod, je nachdem, was zuerst eintritt. Die erfolgten Zustandsübergänge waren läsions-, alters-, geschlechts- und ortsspezifisch und wurden mit einer wahrscheinlichkeitsfreien approximativen bayesschen Berechnungsmethode, der adaptiven Population-Monte-Carlo-Methode (APMC), kalibriert. Für die Kalibrierung wurden 74 Prävalenzdatenpunkte aus dem deutschen Koloskopie-Früherkennungsprogramm verwendet, welches von 2003 bis 2014 aus 5,2 Millionen untersuchten Teilnehmern mit durchschnittlichem Risiko bestand. Die bayessche Kalibrierung lieferte 1.000 Sätze an posterioren Parameterausprägungen, mit denen DECAS erfolgreich die Darmkrebs-Inzidenzdaten vom Zentrum für Krebsregisterdaten reproduzierte.

Nachdem bestätigt wurde, dass DECAS valide ist, um den natürlichen Verlauf von Darmkrebs zu simulieren, wurde die Screening-Komponente zu DECAS hinzugefügt. Falls Läsionen vor dem

klinischen Krebsstadium durch die Screening-Tests entdeckt werden, können sie durch Koloskopie weiterverfolgt oder direkt entfernt werden. Um die Vorhersagefähigkeit von DECAS bezüglich des Ergebnisses eines Darmkrebsfrüherkennungs weiter zu validieren, wurden externe Validierungen gegen zwei große, randomisierte kontrollierte Studien über flexible Sigmoidoskopie und Guajak-Test und eine große Koloskopie-Kohortenstudie durchgeführt. Zusätzlich wurde eine Kreuzvalidierung gegen die drei am häufigsten verwendeten Darmkrebs-Screening-Modelle, die CISNET-Modelle, durchgeführt. DECAS zeigte in den Validierungsstudien eine genaue Vorhersage von Darmkrebs-Inzidenz und Mortalitätsreduktion.

Das validierte DECAS-Model wurde dann verwendet, um Nutzen, Belastung und Schaden von Darmkrebs-Früherkennungsstrategien in Deutschland zu bewerten. Dies schloss die Strategie eines jährlichen FITs für Individuen von 50 bis 54 Jahren gefolgt von zwei Koloskopien im Abstand von 10 Jahren oder einem zweijährlichen FIT für Individuen von 55 bis 75 Jahren für beide Geschlechter ein sowie die neue Strategie, die es Männern erlaubt, die zwei Koloskopien im Abstand von 10 Jahren ab dem Alter von 50 Jahren zu beginnen. Alternative Strategien wie die eines zweijährlichen FITs oder Koloskopien alle 10 Jahre ab einem Alter von 45 oder 50 Jahren und kombinierte Strategien mit einem jährlichen FIT ab 45 gefolgt von zwei oder drei Koloskopien im Abstand von 10 Jahren ab einem Alter von 50 Jahren wurden ebenfalls bewertet. Alle Strategien wurden unter vier Szenarien evaluiert: perfekte Adhärenz, niedrige Adhärenz unter dem aktuellen organisierten Programm mit einem Einladungsschreiben, verbesserte Adhärenz mit einem Einladungsschreiben und Mail-out-FITs, hohe (aber unvollständige) Adhärenz mit einem Einladungsschreiben, Mail-out-FITs und einer zusätzlichen Erinnerung.

Alle Strategien erwiesen sich in allen vier Szenarien als kosteneffektiv im Vergleich zu keinem Screening. Unter der Annahme einer perfekten Adhärenz brachten die Screening-Strategien im Vergleich zu keinem Screening eine Reduktion der Darmkrebs-Inzidenz von 34-75 %, eine Reduktion der Darmkrebs-Mortalität von 52-80 %, 57-97 gewonnene Lebensjahre pro 1.000 40-Jährige und 36-98 gewonnene qualitätsadjustierte Lebensjahre pro 1.000 40-Jährige. Alle Strategien waren sparten Kosten und führten zu 809 bis 3.240 notwendigen Koloskopien sowie 1 bis 4 Koloskopie-Komplikationsfällen pro 1.000 40-Jährige. In den Szenarien mit unvollständiger Adhärenz sanken Nutzen, Belastung und Schaden mit den Teilnahmeraten. In den beiden Mail-out-FIT-Szenarien konnten sich die verschickten, aber nicht genutzten FITs auf bis zu 9.967 Kits summieren und verursachten in der zweijährlichen FIT-Strategie ab einem Alter von 45 Jahren im Szenario mit geringerer Adhärenz einen zusätzlichen Schaden von 93.323 € pro 1.000 40-Jährige.

Zusätzlich schien eine Strategie mit einem geschlechtsdifferenzierten Startalter für die Koloskopie kosteneffektiver zu sein als eine Strategie mit gleichem Startalter. Reine Koloskopie-

oder FIT-Koloskopie-Kombinationsstrategien schienen kosteneffektiver zu sein als reine FIT-Strategien. Die Strategie der dreimaligen Koloskopie alle 10 Jahre ab einem Alter von 45 Jahren erwies sich als die kosteneffektivste über alle Szenarien hinweg, wenn man die Zahlungsbereitschaftsschwellen von 5.000-100.000 € berücksichtigt. Im Allgemeinen zeigte sich, dass Strategien, die mit dem Alter von 45 Jahren beginnen, die beste Balance zwischen Nutzen, Belastung und Schaden bieten, was mit den jüngsten Empfehlungsänderungen der wichtigsten US-Leitlinien übereinstimmt.

Die Modellierungsevidenz aus der vorliegenden Arbeit kann trotz der Unsicherheiten als Grundlage für zukünftige Maßnahmen zur Darmkrebsfrüherkennung in Deutschland dienen, während keine Langzeitevidenz für FIT und Koloskopie-Screenings aus klinischen Studien vorliegt. Zukünftige Forschungsrichtungen schließen eine Neukalibrierung von DECAS mit effizienteren bayesschen Algorithmen und mit robusteren Daten zu serratierten Polypen ein, wenn diese verfügbar sind. Weiterhin kann die Kosteneffektivität für risikostratifizierte Screening-Strategien, die nicht geschlechtsspezifisch sind (z.B. mit a priori individuellen Risiken) und alternative Screening-Modalitäten (z.B. Multitarget-Stuhl-DNA-Test oder Computertomographie-Kolonographie) untersucht werden, entweder im Kontext von Deutschland oder in anderen geographischen Regionen, bis DECAS an die lokale Darmkrebs-Epidemiologie angepasst ist. Schließlich kann DECAS auch verwendet werden, um die Auswirkungen auf die öffentliche Gesundheit und die Wirtschaft zu analysieren, die sich aus einer verzögerten Darmkrebs-Früherkennung aufgrund von Störungen durch externe Kräfte ergeben, z. B. durch die COVID-19-Pandemie.

8 REFERENCES

Ait Ouakrim, D., Pizot, C., Boniol, Magali, Malvezzi, M., Boniol, Mathieu, Negri, E., Bota, M., Jenkins, M.A., Bleiberg, H., Autier, P., 2015. Trends in colorectal cancer mortality in Europe: retrospective analysis of the WHO mortality database. The BMJ 351. https://doi.org/10.1136/bmj.h4970

American Cancer Society, 2022. Colorectal Cancer Stages [WWW Document]. URL https://www.cancer.org/cancer/colon-rectal-cancer/detection-diagnosis-staging/staged.html (accessed 2.23.22).

Amos-Landgraf, J.M., Heijmans, J., Wielenga, M.C.B., Dunkin, E., Krentz, K.J., Clipson, L., Ederveen, A.G., Groothuis, P.G., Mosselman, S., Muncan, V., Hommes, D.W., Shedlovsky, A., Dove, W.F., Brink, G.R. van den, 2014. Sex disparity in colonic adenomagenesis involves promotion by male hormones, not protection by female hormones. Proc. Natl. Acad. Sci. 111, 16514–16519. https://doi.org/10.1073/pnas.1323064111

Atkin, W., Wooldrage, K., Parkin, D.M., Kralj-Hans, I., MacRae, E., Shah, U., Duffy, S., Cross, A.J., 2017. Long term effects of once-only flexible sigmoidoscopy screening after 17 years of follow-up: the UK Flexible Sigmoidoscopy Screening randomised controlled trial. The Lancet 389, 1299–1311. https://doi.org/10.1016/s0140-6736(17)30396-3

Atkin, W.S., Edwards, R., Kralj-Hans, I., Wooldrage, K., Hart, A.R., Northover, J.M.A., Parkin, D.M., Wardle, J., Duffy, S.W., Cuzick, J., UK Flexible Sigmoidoscopy Trial Investigators, 2010. Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. Lancet Lond. Engl. 375, 1624–1633. https://doi.org/10.1016/S0140-6736(10)60551-X

Atkin, W.S., Valori, R., Kuipers, E.J., Hoff, G., Senore, C., Segnan, N., Jover, R., Schmiegel, W., Lambert, R., Pox, C., International Agency for Research on Cancer, 2012. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Colonoscopic surveillance following adenoma removal. Endoscopy 44 Suppl 3, SE151-163. https://doi.org/10.1055/s-0032-1309821

Bagnardi, V., Rota, M., Botteri, E., Tramacere, I., Islami, F., Fedirko, V., Scotti, L., Jenab, M., Turati, F., Pasquali, E., Pelucchi, C., Galeone, C., Bellocco, R., Negri, E., Corrao, G., Boffetta, P., La Vecchia, C., 2015. Alcohol consumption and site-specific cancer risk: a comprehensive dose-response meta-analysis. Br. J. Cancer 112, 580–593. https://doi.org/10.1038/bjc.2014.579

Bakhoum, S.F., Silkworth, W.T., Nardi, I.K., Nicholson, J.M., Compton, D.A., Cimini, D., 2014. The mitotic origin of chromosomal instability. Curr. Biol. CB 24, R148–R149. https://doi.org/10.1016/j.cub.2014.01.019

Basu, A., Ganiats, T.G., 2016. Discounting in Cost-Effectiveness Analysis, in: Cost-Effectiveness in Health and Medicine. Oxford University Press.

Beaumont, M.A., 2019. Approximate Bayesian Computation. Annu. Rev. Stat. Its Appl. 6, 379–403. https://doi.org/10.1146/annurev-statistics-030718-105212

Bénard, F., Barkun, A.N., Martel, M., von Renteln, D., 2018. Systematic review of colorectal cancer screening guidelines for average-risk adults: Summarizing the current global recommendations. World J. Gastroenterol. 24, 124–138. https://doi.org/10.3748/wjg.v24.i1.124

Bettington, M., Walker, N., Clouston, A., Brown, I., Leggett, B., Whitehall, V., 2013. The serrated pathway to colorectal carcinoma: current concepts and challenges. Histopathology 62, 367–86. https://doi.org/10.1111/his.12055

Bonassi, F.V., West, M., 2015. Sequential Monte Carlo with Adaptive Weights for Approximate Bayesian Computation. Bayesian Anal. 10, 171–187. https://doi.org/10.1214/14-ba891

Botteri, E., Iodice, S., Bagnardi, V., Raimondi, S., Lowenfels, A.B., Maisonneuve, P., 2008. Smoking and colorectal cancer: a meta-analysis. JAMA 300, 2765–2778. https://doi.org/10.1001/jama.2008.839

Bouwens, M.W., van Herwaarden, Y.J., Winkens, B., Rondagh, E.J., de Ridder, R., Riedl, R.G., Driessen, A., Dekker, E., Masclee, A.A., Sanduleanu, S., 2014. Endoscopic characterization of sessile serrated adenomas/polyps with and without dysplasia. Endoscopy 46, 225–35. https://doi.org/10.1055/s-0034-1364936

Boyle, T., Keegel, T., Bull, F., Heyworth, J., Fritschi, L., 2012. Physical activity and risks of proximal and distal colon cancers: a systematic review and meta-analysis. J. Natl. Cancer Inst. 104, 1548–1561. https://doi.org/10.1093/jnci/djs354

Bradley, C.J., Yabroff, K.R., Warren, J.L., Zeruto, C., Chawla, N., Lamont, E.B., 2016. Trends in the Treatment of Metastatic Colon and Rectal Cancer in Elderly Patients. Med. Care 54, 490–497. https://doi.org/10.1097/MLR.0000000000000510

Brenner, H., Altenhofen, L., Hoffmeister, M., 2010. Sex, age, and birth cohort effects in colorectal neoplasms: a cohort analysis. Ann. Intern. Med. 152, 697–703. https://doi.org/10.7326/0003-4819-152-11-201006010-00002

Brenner, H., Altenhofen, L., Katalinic, A., Lansdorp-Vogelaar, I., Hoffmeister, M., 2011. Sojourn time of preclinical colorectal cancer by sex and age: estimates from the German national screening colonoscopy database. Am J Epidemiol 174, 1140–6. https://doi.org/10.1093/aje/kwr188

Brenner, H., Altenhofen, L., Stock, C., Hoffmeister, M., 2014a. Incidence of colorectal adenomas: birth cohort analysis among 4.3 million participants of screening colonoscopy. Cancer Epidemiol Biomark. Prev 23, 1920–7. https://doi.org/10.1158/1055-9965.EPI-14-0367

Brenner, H., Altenhofen, L., Stock, C., Hoffmeister, M., 2013. Natural history of colorectal adenomas: birth cohort analysis among 3.6 million participants of screening colonoscopy. Cancer Epidemiol Biomark. Prev 22, 1043–51. https://doi.org/10.1158/1055-9965.EPI-13-0162

Brenner, H., Chen, C., 2018. The colorectal cancer epidemic: challenges and opportunities for primary, secondary and tertiary prevention. Br. J. Cancer 119, 785–792. https://doi.org/10.1038/s41416-018-0264-x

Brenner, H., Hoffmeister, M., Stegmaier, C., Brenner, G., Altenhofen, L., Haug, U., 2007. Risk of progression of advanced adenomas to colorectal cancer by age and sex: estimates based on 840,149 screening colonoscopies. Gut 56, 1585–1589. https://doi.org/10.1136/gut.2007.122739

Brenner, H., Schrotz-King, P., Holleczek, B., Katalinic, A., Hoffmeister, M., 2016. Declining Bowel Cancer Incidence and Mortality in Germany. Dtsch Arztebl Int 113, 101–6. https://doi.org/10.3238/arztebl.2016.0101

Brenner, H., Stock, C., Hoffmeister, M., 2014b. Effect of screening sigmoidoscopy and screening colonoscopy on colorectal cancer incidence and mortality: systematic review and meta-analysis of randomised controlled trials and observational studies. BMJ 348, g2467. https://doi.org/10.1136/bmj.g2467

Brenner, H., Zwink, N., Ludwig, L., Hoffmeister, M., 2017. Should Screening Colonoscopy Be Offered From Age 50? Dtsch Arztebl Int 114, 94–100. https://doi.org/10.3238/arztebl.2017.0094

Briggs, A.D., Wolstenholme, J., Blakely, T., Scarborough, P., 2016. Choosing an epidemiological model structure for the economic evaluation of non-communicable disease public health interventions. Popul Health Metr 14, 17. https://doi.org/10.1186/s12963-016-0085-1

Briggs, A.H., Weinstein, M.C., Fenwick, E.A., Karnon, J., Sculpher, M.J., Paltiel, A.D., Force, I.-S.M.G.R.P.T., 2012. Model parameter estimation and uncertainty: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--6. Value Health 15, 835–42. https://doi.org/10.1016/j.jval.2012.04.014

Carethers, J.M., Jung, B.H., 2015. Genetics and Genetic Biomarkers in Sporadic Colorectal Cancer. Gastroenterology 149, 1177-1190.e3. https://doi.org/10.1053/j.gastro.2015.06.047

Caro, J.J., Möller, J., 2016. Advantages and disadvantages of discrete-event simulation for health economic analyses. Expert Rev. Pharmacoecon. Outcomes Res. 16, 327–329. https://doi.org/10.1586/14737167.2016.1165608

Chang, L.C., Shun, C.T., Hsu, W.F., Tu, C.H., Tsai, P.Y., Lin, B.R., Liang, J.T., Wu, M.S., Chiu, H.M., 2017. Fecal Immunochemical Test Detects Sessile Serrated Adenomas and Polyps With a Low Level of Sensitivity. Clin Gastroenterol Hepatol 15, 872-879 e1. https://doi.org/10.1016/j.cgh.2016.07.029

Chen, C., Stock, C., Hoffmeister, M., Brenner, H., 2019. Optimal age for screening colonoscopy: a modeling study. Gastrointest. Endosc. 89, 1017-1025.e12. https://doi.org/10.1016/j.gie.2018.12.021

Chen, C., Stock, C., Hoffmeister, M., Brenner, H., 2018. How long does it take until the effects of endoscopic screening on colorectal cancer mortality are fully disclosed?: a Markov model study. Int. J. Cancer 143, 2718–2724. https://doi.org/10.1002/ijc.31716

Cheng, C.-Y., Calderazzo, S., Schramm, C., Schlander, M., 2022. Modeling the natural history and screening effects for colorectal cancer using both adenoma and serrated neoplasia pathways: the development, calibration, and validation of a discrete event simulation model. Submitt. Med. Decis. Mak.

Cheng, C.-Y., Datzmann, T., Hernandez, D., Schmitt, J., Schlander, M., 2021. Do certified cancer centers provide more cost-effective care? A health economic analysis of colon cancer care in Germany using administrative data. Int. J. Cancer 149, 1744–1754. https://doi.org/10.1002/ijc.33728

Conteduca, V., Sansonno, D., Russi, S., Dammacco, F., 2013. Precancerous colorectal lesions (Review). Int. J. Oncol. 43, 973–984. https://doi.org/10.3892/ijo.2013.2041

Crockett, S.D., Nagtegaal, I.D., 2019. Terminology, Molecular Features, Epidemiology, and Management of Serrated Colorectal Neoplasia. Gastroenterology 157, 949-966.e4. https://doi.org/10.1053/j.gastro.2019.06.041

Daly, A.C., Cooper, J., Gavaghan, D.J., Holmes, C., 2017. Comparing two sequential Monte Carlo samplers for exact and approximate Bayesian inference on biological models. J R Soc Interface 14. https://doi.org/10.1098/rsif.2017.0340

Dekker, E., Tanis, P.J., Vleugels, J.L.A., Kasi, P.M., Wallace, M.B., 2019. Colorectal cancer. Lancet Lond. Engl. 394, 1467–1480. https://doi.org/10.1016/S0140-6736(19)32319-0

Del Moral, P., Doucet, A., Jasra, A., 2006. Sequential Monte Carlo samplers. J R Stat. Soc B 68, 411–436.

DeYoreo, M., Lansdorp-Vogelaar, I., Knudsen, A.B., Kuntz, K.M., Zauber, A.G., Rutter, C.M., 2020. Validation of Colorectal Cancer Models on Long-term Outcomes from a Randomized Controlled Trial. Med. Decis. Mak. Int. J. Soc. Med. Decis. Mak. 40, 1034–1040. https://doi.org/10.1177/0272989X20961095

Diamond, S.J., Enestvedt, B.K., Jiang, Z., Holub, J.L., Gupta, M., Lieberman, D.A., Eisen, G.M., 2011. Adenoma detection rate increases with each decade of life after 50 years of age. Gastrointest. Endosc. 74, 135–140. https://doi.org/10.1016/j.gie.2011.03.1178

Dominitz, J.A., Robertson, D.J., Ahnen, D.J., Allison, J.E., Antonelli, M., Boardman, K.D., Ciarleglio, M., Del Curto, B.J., Huang, G.D., Imperiale, T.F., Larson, M.F., Lieberman, D., O'Connor, T., O'Leary, T.J., Peduzzi, P., Provenzale, D., Shaukat, A., Sultan, S., Voorhees, A., Wallace, R., Guarino, P.D., 2017. Colonoscopy vs. Fecal Immunochemical Test in Reducing Mortality From Colorectal Cancer (CONFIRM): Rationale for Study Design. Am. J. Gastroenterol. 112, 1736–1746. https://doi.org/10.1038/ajg.2017.286

Drummond, M.F., Sculpher, M.J., Claxton, K., Stoddart, Gerg.L., Torrance, G.W., 2015. Methods for the Economic Evaluation of Health Care Programmes, 4th ed. Oxford University Press.

East, J.E., Vieth, M., Rex, D.K., 2015. Serrated lesions in colorectal cancer screening: detection, resection, pathology and surveillance. Gut 64, 991–1000. https://doi.org/10.1136/gutjnl-2014-309041

Eddy, D.M., Hollingworth, W., Caro, J.J., Tsevat, J., McDonald, K.M., Wong, J.B., Force, I.-S.M.G.R.P.T., 2012. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--7. Value Health 15, 843–50. https://doi.org/10.1016/j.jval.2012.04.012

Erichsen, R., Baron, J.A., Hamilton-Dutoit, S.J., Snover, D.C., Torlakovic, E.E., Pedersen, L., Frøslev, T., Vyberg, M., Hamilton, S.R., Sørensen, H.T., 2016. Increased Risk of Colorectal Cancer Development Among Patients With Serrated Polyps. Gastroenterology 150, 895-902.e5. https://doi.org/10.1053/j.gastro.2015.11.046

European Colorectal Cancer Screening Guidelines Working Group, von Karsa, L., Patnick, J., Segnan, N., Atkin, W., Halloran, S., Lansdorp-Vogelaar, I., Malila, N., Minozzi, S., Moss, S., Quirke, P., Steele, R.J., Vieth, M., Aabakken, L., Altenhofen, L., Ancelle-Park, R., Antoljak, N., Anttila, A., Armaroli, P., Arrossi, S., Austoker, J., Banzi, R., Bellisario, C., Blom, J., Brenner, H., Bretthauer, M., Camargo Cancela, M., Costamagna, G., Cuzick, J., Dai, M., Daniel, J., Dekker, E., Delicata, N., Ducarroz, S., Erfkamp, H., Espinàs, J.A., Faivre, J., Faulds Wood, L., Flugelman, A., Frkovic-Grazio, S., Geller, B., Giordano, L., Grazzini, G., Green, J., Hamashima, C., Herrmann, C., Hewitson, P., Hoff, G., Holten, I., Jover, R., Kaminski, M.F., Kuipers, E.J., Kurtinaitis, J., Lambert, R., Launoy, G., Lee, W., Leicester, R., Leja, M., Lieberman, D., Lignini, T., Lucas, E., Lynge, E., Mádai, S., Marinho, J., Maučec Zakotnik, J., Minoli, G., Monk, C., Morais, A., Muwonge, R., Nadel, M., Neamtiu, L., Peris Tuser, M., Pignone, M., Pox, C., Primic-Zakelj, M., Psaila, J., Rabeneck, L., Ransohoff, D., Rasmussen, M., Regula, J., Ren, J., Rennert,

G., Rey, J., Riddell, R.H., Risio, M., Rodrigues, V., Saito, H., Sauvaget, C., Scharpantgen, A., Schmiegel, W., Senore, C., Siddiqi, M., Sighoko, D., Smith, R., Smith, S., Suchanek, S., Suonio, E., Tong, W., Törnberg, S., Van Cutsem, E., Vignatelli, L., Villain, P., Voti, L., Watanabe, H., Watson, J., Winawer, S., Young, G., Zaksas, V., Zappa, M., Valori, R., 2013. European guidelines for quality assurance in colorectal cancer screening and diagnosis: overview and introduction to the full supplement publication. Endoscopy 45, 51–59. https://doi.org/10.1055/s-0032-1325997

Färkkilä, N., Sintonen, H., Saarto, T., Järvinen, H., Hänninen, J., Taari, K., Roine, R.P., 2013. Health-related quality of life in colorectal cancer. Colorectal Dis. Off. J. Assoc. Coloproctology G. B. Irel. 15, e215-222. https://doi.org/10.1111/codi.12143

Fearnhead, P., Prangle, D., 2012. Constructing summary statistics for approximate Bayesian computation: semi-automatic approximate Bayesian computation. J R Stat. Soc B 74, 419–474. https://doi.org/10.1111/j.1467-9868.2011.01010.x

Ferlitsch, M., Reinhart, K., Pramhas, S., Wiener, C., Gal, O., Bannert, C., Hassler, M., Kozbial, K., Dunkler, D., Trauner, M., Weiss, W., 2011. Sex-specific prevalence of adenomas, advanced adenomas, and colorectal cancer in individuals undergoing screening colonoscopy. JAMA 306, 1352–1358. https://doi.org/10.1001/jama.2011.1362

Fidler, M.M., Soerjomataram, I., Bray, F., 2016. A global view on cancer incidence and national levels of the human development index. Int. J. Cancer 139, 2436–2446. https://doi.org/10.1002/ijc.30382

G-BA, 2018. Beschluss des Gemeinsamen Bundesausschusses über eine Richtlinie für organisierte Krebsfrüherkennungsprogramme und eine Änderung der Krebsfrüherkennungs-Richtlinie. Gemeinsamer Bundesausschuss (G-BA), Berlin.

G-BA, 2016. Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Krebsfrüherkennungsrichtlinie: Bewertung eines iFOBT-basierten Darmkrebs-screenings im Vergleich zu einem gFOBT-basierten Darmkrebsscreening. Gemeinsamer Bundesausschuss (G-BA), Berlin.

Gemeinsamer Bundesausschuss (G-BA), 2018. Richtlinie für organisierte Krebsfrüherkennungsprogramme – oKFE-RL. Gemeinsamer Bundesausschuss, Berlin.

Gesundheitsberichterstattung des Bundes (GBE), 2021. Gesundheitsversorgung Prävention Krebsfrüherkennung [WWW Document]. URL https://www.gbe-bund.de/gbe/trecherche.prc_them_rech?tk=14501&tk2=15110&p_uid=gast&p_aid=24380663&p_sprache=D&cnt_ut=8&ut=15140 (accessed 4.7.21).

GGPO, 2019. Evidenced-based Guideline for Colorectal Cancer. German Guideline Program in Oncology (GGPO), Berlin.

Gillespie, D.T., 1976. A General Method for Numerically Simulating the Stochastic Time Evolution of Coupled Chemical Reactions. J. Comput. Phys. 22, 403–434. https://doi.org/10.1016/0021-9991(76)90041-3

Goodwin, B.C., Ireland, M.J., March, S., Myers, L., Crawford-Williams, F., Chambers, S.K., Aitken, J.F., Dunn, J., 2019. Strategies for increasing participation in mail-out colorectal cancer screening programs: a systematic review and meta-analysis. Syst. Rev. 8, 257. https://doi.org/10.1186/s13643-019-1170-x

Greuter, M.J., Demirel, E., Lew, J.B., Berkhof, J., Xu, X.M., Canfell, K., Dekker, E., Meijer, G.A., Coupe, V.M., 2016. Long-Term Impact of the Dutch Colorectal Cancer Screening Program on

Cancer Incidence and Mortality-Model-Based Exploration of the Serrated Pathway. Cancer Epidemiol Biomark. Prev 25, 135–44. https://doi.org/10.1158/1055-9965.EPI-15-0592

Greuter, M.J., Xu, X.M., Lew, J.B., Dekker, E., Kuipers, E.J., Canfell, K., Meijer, G.A., Coupe, V.M., 2014. Modeling the Adenoma and Serrated pathway to Colorectal CAncer (ASCCA). Risk Anal 34, 889–910. https://doi.org/10.1111/risa.12137

Greuter, M.J.E., de Klerk, C.M., Meijer, G.A., Dekker, E., Coupé, V.M.H., 2017. Screening for Colorectal Cancer With Fecal Immunochemical Testing With and Without Postpolypectomy Surveillance Colonoscopy: A Cost-Effectiveness Analysis. Ann. Intern. Med. 167, 544–554. https://doi.org/10.7326/M16-2891

Gruner, L.F., Amitay, E.L., Heisser, T., Guo, F., Niedermaier, T., Gies, A., Hoffmeister, M., Brenner, H., 2021. The Effects of Different Invitation Schemes on the Use of Fecal Occult Blood Tests for Colorectal Cancer Screening: Systematic Review of Randomized Controlled Trials. Cancers 13, 1520. https://doi.org/10.3390/cancers13071520

Gruner, L.F., Hoffmeister, M., Ludwig, L., Meny, S., Brenner, H., 2020. The Effects of Differing Invitation Models on the Uptake of Immunological Fecal Occult Blood Testing. Dtsch. Arzteblatt Int. 117, 423–430. https://doi.org/10.3238/arztebl.2020.0423

Guo, F., Chen, C., Holleczek, B., Schöttker, B., Hoffmeister, M., Brenner, H., 2021. Strong Reduction of Colorectal Cancer Incidence and Mortality After Screening Colonoscopy: Prospective Cohort Study From Germany. Off. J. Am. Coll. Gastroenterol. ACG Publish Ahead of Print. https://doi.org/10.14309/ajg.0000000000001146

Hardcastle, J.D., Chamberlain, J.O., Robinson, M.H., Moss, S.M., Amar, S.S., Balfour, T.W., James, P.D., Mangham, C.M., 1996. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet Lond. Engl. 348, 1472–1477. https://doi.org/10.1016/S0140-6736(96)03386-7

Haug, U., 2018. Aktueller Stand der Evidenz und Entwicklung der Darmkrebsfrüherkennung in Deutschland. Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz 61, 1513–1516. https://doi.org/10.1007/s00103-018-2832-x

Haug, U., Engel, S., Verheyen, F., Linder, R., 2014. Estimating colorectal cancer treatment costs: a pragmatic approach exemplified by health insurance data from Germany. PloS One 9, e88407. https://doi.org/10.1371/journal.pone.0088407

Heisser, T., Hoffmeister, M., Brenner, H., 2021a. Effects of screening for colorectal cancer: Development, documentation and validation of a multistate Markov model. Int. J. Cancer 148, 1973–1981. https://doi.org/10.1002/ijc.33437

Heisser, T., Hoffmeister, M., Brenner, H., 2021b. Model based evaluation of long-term efficacy of existing and alternative colorectal cancer screening offers: A case study for Germany. Int. J. Cancer. https://doi.org/10.1002/ijc.33894

Helsingen, L.M., Vandvik, P.O., Jodal, H.C., Agoritsas, T., Lytvyn, L., Anderson, J.C., Auer, R., Murphy, S.B., Almadi, M.A., Corley, D.A., Quinlan, C., Fuchs, J.M., McKinnon, A., Qaseem, A., Heen, A.F., Siemieniuk, R.A.C., Kalager, M., Usher-Smith, J.A., Lansdorp-Vogelaar, I., Bretthauer, M., Guyatt, G., 2019. Colorectal cancer screening with faecal immunochemical testing, sigmoidoscopy or colonoscopy: a clinical practice guideline. BMJ 367, l5515. https://doi.org/10.1136/bmj.l5515

Hewitson, P., Glasziou, P., Watson, E., Towler, B., Irwig, L., 2008. Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult): an update. Am. J. Gastroenterol. 103, 1541–1549. https://doi.org/10.1111/j.1572-0241.2008.01875.x

Hoffmeister, M., Holleczek, B., Stock, C., Zwink, N., Stolz, T., Stegmaier, C., Brenner, H., 2019. Utilization and determinants of follow-up colonoscopies within 6 years after screening colonoscopy: Prospective cohort study. Int. J. Cancer 144, 402–410. https://doi.org/10.1002/ijc.31862

Hoffmeister, M., Holleczek, B., Zwink, N., Stock, C., Stegmaier, C., Brenner, H., 2017. Screening for Bowel Cancer: Increasing Participation via Personal Invitation. Dtsch. Arzteblatt Int. 114, 87–93. https://doi.org/10.3238/arztebl.2017.0087

Hol, L., van Leerdam, M.E., van Ballegooijen, M., van Vuuren, A.J., van Dekken, H., Reijerink, J.C.I.Y., van der Togt, A.C.M., Habbema, J.D.F., Kuipers, E.J., 2010. Screening for colorectal cancer: randomised trial comparing guaiac-based and immunochemical faecal occult blood testing and flexible sigmoidoscopy. Gut 59, 62–68. https://doi.org/10.1136/gut.2009.177089

Holme, O., Schoen, R.E., Senore, C., Segnan, N., Hoff, G., Loberg, M., Bretthauer, M., Adami, H.O., Kalager, M., 2017. Effectiveness of flexible sigmoidoscopy screening in men and women and different age groups: pooled analysis of randomised trials. BMJ 356, i6673. https://doi.org/10.1136/bmj.i6673

Human Mortality Database, 2018. Life tables 2010-2014, Germany. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany).

Ijspeert, J.E., Vermeulen, L., Meijer, G.A., Dekker, E., 2015. Serrated neoplasia-role in colorectal carcinogenesis and clinical implications. Nat Rev Gastroenterol Hepatol 12, 401–9. https://doi.org/10.1038/nrgastro.2015.73

Ijspeert, J.E.G., Bevan, R., Senore, C., Kaminski, M.F., Kuipers, E.J., Mroz, A., Bessa, X., Cassoni, P., Hassan, C., Repici, A., Balaguer, F., Rees, C.J., Dekker, E., 2017. Detection rate of serrated polyps and serrated polyposis syndrome in colorectal cancer screening cohorts: a European overview. Gut 66, 1225–1232. https://doi.org/10.1136/gutjnl-2015-310784

IJspeert, J.E.G., Tutein Nolthenius, C.J., Kuipers, E.J., van Leerdam, M.E., Nio, C.Y., Thomeer, M.G.J., Biermann, K., van de Vijver, M.J., Dekker, E., Stoker, J., 2016. CT-Colonography vs. Colonoscopy for Detection of High-Risk Sessile Serrated Polyps. Am. J. Gastroenterol. 111, 516–522. https://doi.org/10.1038/ajg.2016.58

Imperiale, T.F., Gruber, R.N., Stump, T.E., Emmett, T.W., Monahan, P.O., 2019. Performance Characteristics of Fecal Immunochemical Tests for Colorectal Cancer and Advanced Adenomatous Polyps: A Systematic Review and Meta-analysis. Ann. Intern. Med. 170, 319–329. https://doi.org/10.7326/M18-2390

Imperiale, T.F., Ransohoff, D.F., Itzkowitz, S.H., Levin, T.R., Lavin, P., Lidgard, G.P., Ahlquist, D.A., Berger, B.M., 2014. Multitarget stool DNA testing for colorectal-cancer screening. N. Engl. J. Med. 370, 1287–1297. https://doi.org/10.1056/NEJMoa1311194

InEK, 2021. German DRG System [WWW Document]. URL https://www.g-drg.de/ (accessed 3.29.21).

Institute for Quality and Efficiency in Health Care (IQWiG), 2020. General Methods Version 6.0. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, Cologne.

Jabot, F., Faure, T., Dumoulin, N., O'Hara, R.B., 2013. EasyABC: performing efficient approximate Bayesian computation sampling schemes using R. Methods Ecol. Evol. 4, 684–687. https://doi.org/10.1111/2041-210x.12050

Jasperson, K.W., Tuohy, T.M., Neklason, D.W., Burt, R.W., 2010. Hereditary and familial colon cancer. Gastroenterology 138, 2044–2058. https://doi.org/10.1053/j.gastro.2010.01.054

Jodal, H.C., Helsingen, L.M., Anderson, J.C., Lytvyn, L., Vandvik, P.O., Emilsson, L., 2019. Colorectal cancer screening with faecal testing, sigmoidoscopy or colonoscopy: a systematic review and network meta-analysis. BMJ Open 9, e032773. https://doi.org/10.1136/bmjopen-2019-032773

Kaminski, M.F., Bretthauer, M., Zauber, A.G., Kuipers, E.J., Adami, H.-O., van Ballegooijen, M., Regula, J., van Leerdam, M., Stefansson, T., Påhlman, L., Dekker, E., Hernán, M.A., Garborg, K., Hoff, G., 2012. The NordICC Study: rationale and design of a randomized trial on colonoscopy screening for colorectal cancer. Endoscopy 44, 695–702. https://doi.org/10.1055/s-0032-1306895

Karnon, J., Haji Ali Afzali, H., 2014. When to use discrete event simulation (DES) for the economic evaluation of health technologies? A review and critique of the costs and benefits of DES. PharmacoEconomics 32, 547–558. https://doi.org/10.1007/s40273-014-0147-9

Karnon, J., Stahl, J., Brennan, A., Caro, J.J., Mar, J., Moller, J., ISPOR-SMDM Modeling Good Research Practices Task Force, 2012. Modeling using discrete event simulation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--4. Value Health 15, 821–7. https://doi.org/10.1016/j.jval.2012.04.013

Keum, N., Giovannucci, E., 2019. Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies. Nat Rev Gastroenterol Hepatol. https://doi.org/10.1038/s41575-019-0189-8

Klabunde, C., Blom, J., Bulliard, J.-L., Garcia, M., Hagoel, L., Mai, V., Patnick, J., Rozjabek, H., Senore, C., Törnberg, S., 2015. Participation rates for organized colorectal cancer screening programmes: an international comparison. J. Med. Screen. 22, 119–126. https://doi.org/10.1177/0969141315584694

Knudsen, A.B., Hur, C., Gazelle, G.S., Schrag, D., McFarland, E.G., Kuntz, K.M., 2012. Rescreening of persons with a negative colonoscopy result: results from a microsimulation model. Ann Intern Med 157, 611–20. https://doi.org/10.7326/0003-4819-157-9-201211060-00005

Knudsen, A.B., Rutter, C.M., Peterse, E.F.P., Lietz, A.P., Seguin, C.L., Meester, R.G.S., Perdue, L.A., Lin, J.S., Siegel, R.L., Doria-Rose, V.P., Feuer, E.J., Zauber, A.G., Kuntz, K.M., Lansdorp-Vogelaar, I., 2021. Colorectal Cancer Screening: An Updated Modeling Study for the US Preventive Services Task Force. JAMA 325, 1998–2011. https://doi.org/10.1001/jama.2021.5746

Knudsen, A.B., Rutter, C.M., Peterse, E.F.P., Lietz, A.P., Seguin, C.L., Meester, R.G.S., Perdue, L.A., Lin, J.S., Siegel, R.L., Zauber, A.G., Kuntz, K.M., Lansdorp-Vogelaar, I., 2020. Colorectal Cancer Screening: A Decision Analysis for the U.S. Preventive Services Task Force (No. AHRQ Publication No. 20-05271-EF-2). Agency for Healthcare Research and Quality, Rockville, MD.

Kretschmann, J., El Mahi, C., Lichtner, F., Hagen, B., 2020. Früherkennungskoloskopie Jahresbericht 2018. Zentralinstitut für die kassenärztliche Versorgung in Deutschland, Cologne.

Kronborg, O., Jørgensen, O.D., Fenger, C., Rasmussen, M., 2004. Randomized study of biennial screening with a faecal occult blood test: results after nine screening rounds. Scand. J. Gastroenterol. 39, 846–851. https://doi.org/10.1080/00365520410003182

Kubisch, C., Crispin, A., Mansmann, U., Göke, B., Kolligs, F., 2016. Screening for Colorectal Cancer is Associated with Lower Disease Stage: A Population-Based Study. Clin. Gastroenterol. Hepatol. 14, 1612–1618. https://doi.org/10.1016/j.cgh.2016.04.008

Kuntz, K.M., Lansdorp-Vogelaar, I., Rutter, C.M., Knudsen, A.B., van Ballegooijen, M., Savarino, J.E., Feuer, E.J., Zauber, A.G., 2011. A systematic comparison of microsimulation models of colorectal cancer: the role of assumptions about adenoma progression. Med Decis Mak. 31, 530–9. https://doi.org/10.1177/0272989X11408730

KVB, 2021. Einheitlicher Bewertungsmasstab (EBM): Uniform Value Scale [WWW Document]. URL https://www.kbv.de/html/ebm.php (accessed 4.23.21).

Ladabaum, U., Alvarez-Osorio, L., Rosch, T., Brueggenjuergen, B., 2014. Cost-effectiveness of colorectal cancer screening in Germany: current endoscopic and fecal testing strategies versus plasma methylated Septin 9 DNA. Endosc Int Open 2, E96–E104. https://doi.org/10.1055/s-0034-1377182

Ladabaum, U., Mannalithara, A., Meester, R.G.S., 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. https://doi.org/10.1053/j.gastro.2019.03.023

Lansdorp-Vogelaar, I., Knudsen, A.B., Brenner, H., 2011. Cost-effectiveness of colorectal cancer screening. Epidemiol. Rev. 33, 88–100. https://doi.org/10.1093/epirev/mxr004

Lansdorp-Vogelaar, I., van Ballegooijen, M., Boer, R., Zauber, A., Habbema, J.D.F., 2009. A novel hypothesis on the sensitivity of the fecal occult blood test: Results of a joint analysis of 3 randomized controlled trials. Cancer 115, 2410–2419. https://doi.org/10.1002/cncr.24256

Lash, R.H., Genta, R.M., Schuler, C.M., 2010. Sessile serrated adenomas: prevalence of dysplasia and carcinoma in 2139 patients. J Clin Pathol 63, 681–6. https://doi.org/10.1136/jcp.2010.075507

Laudicella, M., Walsh, B., Burns, E., Smith, P.C., 2016. Cost of care for cancer patients in England: evidence from population-based patient-level data. Br. J. Cancer 114, 1286–1292. https://doi.org/10.1038/bjc.2016.77

Lenormand, M., Jabot, F., Deffuant, G., 2013. Adaptive approximate Bayesian computation for complex models. Comput. Stat. 28, 2777–2796. https://doi.org/10.1007/s00180-013-0428-3

Leshno, A., Moshkowitz, M., David, M., Galazan, L., Neugut, A.I., Arber, N., Santo, E., 2016. Prevalence of colorectal neoplasms in young, average risk individuals: A turning tide between East and West. World J Gastroenterol 22, 7365–72. https://doi.org/10.3748/wjg.v22.i32.7365

Levin, T.R., Corley, D.A., Jensen, C.D., Schottinger, J.E., Quinn, V.P., Zauber, A.G., Lee, J.K., Zhao, W.K., Udaltsova, N., Ghai, N.R., Lee, A.T., Quesenberry, C.P., Fireman, B.H., Doubeni, C.A., 2018. Effects of Organized Colorectal Cancer Screening on Cancer Incidence and Mortality in a Large Community-Based Population. Gastroenterology 155, 1383-1391.e5. https://doi.org/10.1053/j.gastro.2018.07.017

Lew, J.-B., St John, D.J.B., Macrae, F.A., Emery, J.D., Ee, H.C., Jenkins, M.A., He, E., Grogan, P., Caruana, M., Sarfati, D., Greuter, M.J.E., Coupé, V.M.H., Canfell, K., 2018. Evaluation of the benefits, harms and cost-effectiveness of potential alternatives to iFOBT testing for colorectal cancer screening in Australia. Int. J. Cancer 143, 269–282. https://doi.org/10.1002/ijc.31314

Lew, J.-B., St John, D.J.B., Xu, X.-M., Greuter, M.J.E., Caruana, M., Cenin, D.R., He, E., Saville, M., Grogan, P., Coupé, V.M.H., Canfell, K., 2017. Long-term evaluation of benefits, harms, and cost-effectiveness of the National Bowel Cancer Screening Program in Australia: a modelling study. Lancet Public Health 2, e331–e340. https://doi.org/10.1016/s2468-2667(17)30105-6

Lin, J.S., Perdue, L.A., Henrikson, N.B., Bean, S.I., Blasi, P.R., 2021. Screening for Colorectal Cancer: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. JAMA 325, 1978–1997. https://doi.org/10.1001/jama.2021.4417

Lin, J.S., Perdue, L.A., Henrikson, N.B., Bean, S.I., Blasi, P.R., 2020. Screening for Colorectal Cancer: An Evidence Update for the U.S. Preventive Services Task Force (No. AHRQ Publication No. 20-05271-EF-1). Agency for Healthcare Research and Quality, Rockville, MD.

Lindholm, E., Brevinge, H., Haglind, E., 2008. Survival benefit in a randomized clinical trial of faecal occult blood screening for colorectal cancer. Br. J. Surg. 95, 1029–1036. https://doi.org/10.1002/bjs.6136

Loeve, F., Boer, R., van Oortmarssen, G.J., van Ballegooijen, M., Habbema, J.D., 1999. The MISCAN-COLON simulation model for the evaluation of colorectal cancer screening. Comput Biomed Res 32, 13–33. https://doi.org/10.1006/cbmr.1998.1498

Morois, S., Cottet, V., Racine, A., Clavel-Chapelon, F., Carbonnel, F., Bastide, N., Boutron-Ruault, M.-C., 2014. Colonoscopy reduced distal colorectal cancer risk and excess cancer risk associated with family history. Cancer Causes Control CCC 25, 1329–1336. https://doi.org/10.1007/s10552-014-0438-7

Munich Canncer Registry, 2018. ICD-10 C18-C20: Colorectal Cancer Survival. Munich Cancer Registry, Munich, Germany.

Nagtegaal, I.D., Odze, R.D., Klimstra, D., Paradis, V., Rugge, M., Schirmacher, P., Washington, K.M., Carneiro, F., Cree, I.A., WHO Classification of Tumours Editorial Board, 2020. The 2019 WHO classification of tumours of the digestive system. Histopathology 76, 182–188. https://doi.org/10.1111/his.13975

Nguyen, S.P., Bent, S., Chen, Y.-H., Terdiman, J.P., 2009. Gender as a risk factor for advanced neoplasia and colorectal cancer: a systematic review and meta-analysis. Clin. Gastroenterol. Hepatol. Off. Clin. Pract. J. Am. Gastroenterol. Assoc. 7, 676-681.e1–3. https://doi.org/10.1016/j.cgh.2009.01.008

Nishihara, R., Wu, K., Lochhead, P., Morikawa, T., Liao, X., Qian, Z.R., Inamura, K., Kim, S.A., Kuchiba, A., Yamauchi, M., Imamura, Y., Willett, W.C., Rosner, B.A., Fuchs, C.S., Giovannucci, E., Ogino, S., Chan, A.T., 2013. Long-term colorectal-cancer incidence and mortality after lower endoscopy. N. Engl. J. Med. 369, 1095–1105. https://doi.org/10.1056/NEJMoa1301969

Nott, D.J., Ong, V.M.-H., Fan, Y., Sisson, S.A., 2019. High-dimensional ABC, in: Sisson, S.A., Fan, Y., Beaumont, M.A. (Eds.), Handbook of Approximate Bayesian Computation. CRC Press, Roca Raton.

Peterse, E.F.P., Meester, R.G.S., Siegel, R.L., Chen, J.C., Dwyer, A., Ahnen, D.J., Smith, R.A., Zauber, A.G., Lansdorp-Vogelaar, I., 2018. The impact of the rising colorectal cancer incidence in young adults on the optimal age to start screening: Microsimulation analysis I to inform the American Cancer Society colorectal cancer screening guideline. Cancer 124, 2964–2973. https://doi.org/10.1002/cncr.31543

Pickhardt, P.J., Hassan, C., Halligan, S., Marmo, R., 2011. Colorectal cancer: CT colonography and colonoscopy for detection--systematic review and meta-analysis. Radiology 259, 393–405. https://doi.org/10.1148/radiol.11101887

Pilonis, N.D., Bugajski, M., Wieszczy, P., Franczyk, R., Didkowska, J., Wojciechowska, U., Pisera, M., Rupinski, M., Regula, J., Kaminski, M.F., 2020. Long-Term Colorectal Cancer Incidence and Mortality After a Single Negative Screening Colonoscopy. Ann. Intern. Med. 173, 81–91. https://doi.org/10.7326/M19-2477

Pox, C.P., Altenhofen, L., Brenner, H., Theilmeier, A., Von Stillfried, D., Schmiegel, W., 2012. Efficacy of a nationwide screening colonoscopy program for colorectal cancer. Gastroenterology 142, 1460-1467.e2. https://doi.org/10.1053/j.gastro.2012.03.022

Prakash, M.K., Lang, B., Heinrich, H., Valli, P.V., Bauerfeind, P., Sonnenberg, A., Beerenwinkel, N., Misselwitz, B., 2017. CMOST: an open-source framework for the microsimulation of colorectal cancer screening strategies. BMC Med Inf. Decis Mak 17, 80. https://doi.org/10.1186/s12911-017-0458-9

Pritchard, J.K., Seielstad, M.T., Perez-Lezaun, A., Feldman, M.W., 1999. Population Growth of Human Y Chromosomes: A Study of Y Chromosome Microsatellites. Mol Biol Evol 16, 1791–1798.

Quintero, E., Castells, A., Bujanda, L., Cubiella, J., Salas, D., Lanas, Á., Andreu, M., Carballo, F., Morillas, J.D., Hernández, C., Jover, R., Montalvo, I., Arenas, J., Laredo, E., Hernández, V., Iglesias, F., Cid, E., Zubizarreta, R., Sala, T., Ponce, M., Andrés, M., Teruel, G., Peris, A., Roncales, M.-P., Polo-Tomás, M., Bessa, X., Ferrer-Armengou, O., Grau, J., Serradesanferm, A., Ono, A., Cruzado, J., Pérez-Riquelme, F., Alonso-Abreu, I., de la Vega-Prieto, M., Reyes-Melian, J.M., Cacho, G., Díaz-Tasende, J., Herreros-de-Tejada, A., Poves, C., Santander, C., González-Navarro, A., COLONPREV Study Investigators, 2012. Colonoscopy versus fecal immunochemical testing in colorectal-cancer screening. N. Engl. J. Med. 366, 697–706. https://doi.org/10.1056/NEJMoa1108895

Ran, T., Cheng, C.Y., Misselwitz, B., Brenner, H., Ubels, J., Schlander, M., 2019. Cost-Effectiveness of Colorectal Cancer Screening Strategies—a Systematic Review. Clin Gastroenterol Hepatol 17, 1969–1981. https://doi.org/10.1016/j.cgh.2019.01.014

Rex, D.K., Adler, S.N., Aisenberg, J., Burch, W.C., Carretero, C., Chowers, Y., Fein, S.A., Fern, S.E., Fernandez-Urien Sainz, I., Fich, A., Gal, E., Horlander, J.C., Isaacs, K.L., Kariv, R., Lahat, A., Leung, W.-K., Malik, P.R., Morgan, D., Papageorgiou, N., Romeo, D.P., Shah, S.S., Waterman, M., 2015. Accuracy of capsule colonoscopy in detecting colorectal polyps in a screening population. Gastroenterology 148, 948-957.e2. https://doi.org/10.1053/j.gastro.2015.01.025

Rex, D.K., Ahnen, D.J., Baron, J.A., Batts, K.P., Burke, C.A., Burt, R.W., Goldblum, J.R., Guillem, J.G., Kahi, C.J., Kalady, M.F., O'Brien, M.J., Odze, R.D., Ogino, S., Parry, S., Snover, D.C., Torlakovic, E.E., Wise, P.E., Young, J., Church, J., 2012. Serrated lesions of the colorectum: review and recommendations from an expert panel. Am J Gastroenterol 107, 1315–29; quiz 1314, 1330. https://doi.org/10.1038/ajg.2012.161

Robert Koch Institut, 2021. Krebs in Deutschland für 2017/2018. Rober Koch Institut, Berlin.

Robert Koch Institute, 2019. German Centre for Cancer Registry Data (ZfKD) [WWW Document]. URL https://www.krebsdaten.de/Krebs/DE/Home/homepage_node.html

Rubin, D.B., 2005. Causal Inference Using Potential Outcomes. J. Am. Stat. Assoc. 100, 322–331. https://doi.org/10.1198/016214504000001880

Rutter, C., Ozik, J., DeYoreo, M., Collier, N., 2018. Microsimulation Model Calibration using Incremental Mixture Approximate Bayesian Computation.

Rutter, C.M., Knudsen, A.B., Marsh, T.L., Doria-Rose, V.P., Johnson, E., Pabiniak, C., Kuntz, K.M., van Ballegooijen, M., Zauber, A.G., Lansdorp-Vogelaar, I., 2016. Validation of Models Used to Inform Colorectal Cancer Screening Guidelines. Med. Decis. Making 36, 604–614. https://doi.org/10.1177/0272989x15622642

Rutter, C.M., Miglioretti, D.L., Savarino, J.E., 2009. Bayesian Calibration of Microsimulation Models. J Am Stat Assoc 104, 1338–1350. https://doi.org/10.1198/jasa.2009.ap07466

Rutter, C.M., Zaslavsky, A.M., Feuer, E.J., 2011. Dynamic microsimulation models for health outcomes: a review. Med Decis Mak. 31, 10–8. https://doi.org/10.1177/0272989X10369005

Sarfati, D., Koczwara, B., Jackson, C., 2016. The impact of comorbidity on cancer and its treatment. CA. Cancer J. Clin. 66, 337–350. https://doi.org/10.3322/caac.21342

Scholefield, J.H., Moss, S.M., Mangham, C.M., Whynes, D.K., Hardcastle, J.D., 2012. Nottingham trial of faecal occult blood testing for colorectal cancer: a 20-year follow-up. Gut 61, 1036–1040. https://doi.org/10.1136/gutjnl-2011-300774

Schramm, C., Janhsen, K., Hofer, J.H., Toermer, H., Stelzer, A., Stenschke, F., Stollenwerk, M., Scheller, I., Lang, S., Goeser, T., Steffen, H.M., 2018. Detection of clinically relevant serrated polyps during screening colonoscopy: results from seven cooperating centers within the German colorectal screening program. Endoscopy 50, 993–1000. https://doi.org/10.1055/a-0598-4477

Schreuders, E.H., Ruco, A., Rabeneck, L., Schoen, R.E., Sung, J.J.Y., Young, G.P., Kuipers, E.J., 2015. Colorectal cancer screening: a global overview of existing programmes. Gut 64, 1637–1649. https://doi.org/10.1136/gutjnl-2014-309086

Senore, C., Basu, P., Anttila, A., Ponti, A., Tomatis, M., Vale, D.B., Ronco, G., Soerjomataram, I., Primic-Žakelj, M., Riggi, E., Dillner, J., Elfström, M.K., Lönnberg, S., Sankaranarayanan, R., Segnan, N., 2019. Performance of colorectal cancer screening in the European Union Member States: data from the second European screening report. Gut 68, 1232–1244. https://doi.org/10.1136/gutjnl-2018-317293

Senore, C., Inadomi, J., Segnan, N., Bellisario, C., Hassan, C., 2015. Optimising colorectal cancer screening acceptance: a review. Gut 64, 1158–1177. https://doi.org/10.1136/gutjnl-2014-308081

Shaukat, A., Kahi, C.J., Burke, C.A., Rabeneck, L., Sauer, B.G., Rex, D.K., 2021. ACG Clinical Guidelines: Colorectal Cancer Screening 2021. Am. J. Gastroenterol. 116, 458–479. https://doi.org/10.14309/ajg.0000000000001122

Shaukat, A., Mongin, S.J., Geisser, M.S., Lederle, F.A., Bond, J.H., Mandel, J.S., Church, T.R., 2013. Long-term mortality after screening for colorectal cancer. N Engl J Med 369, 1106–14. https://doi.org/10.1056/NEJMoa1300720

Siebert, U., Alagoz, O., Bayoumi, A.M., Jahn, B., Owens, D.K., Cohen, D.J., Kuntz, K.M., ISPOR-SMDM Modeling Good Research Practices Task Force, 2012. State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--3. Value Health J. Int. Soc. Pharmacoeconomics Outcomes Res. 15, 812–820. https://doi.org/10.1016/j.jval.2012.06.014

Siegel, R.L., Miller, K.D., Goding Sauer, A., Fedewa, S.A., Butterly, L.F., Anderson, J.C., Cercek, A., Smith, R.A., Jemal, A., 2020. Colorectal cancer statistics, 2020. CA. Cancer J. Clin. 70, 145–164. https://doi.org/10.3322/caac.21601

Siegel, R.L., Torre, L.A., Soerjomataram, I., Hayes, R.B., Bray, F., Weber, T.K., Jemal, A., 2019. Global patterns and trends in colorectal cancer incidence in young adults. Gut 68, 2179–2185. https://doi.org/10.1136/gutjnl-2019-319511

Siegel, R.L., Ward, E.M., Jemal, A., 2012. Trends in colorectal cancer incidence rates in the United States by tumor location and stage, 1992-2008. Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol. 21, 411–416. https://doi.org/10.1158/1055-9965.EPI-11-1020

Silva-Illanes, N., Espinoza, M., 2018. Critical Analysis of Markov Models Used for the Economic Evaluation of Colorectal Cancer Screening: A Systematic Review. Value Health J. Int. Soc. Pharmacoeconomics Outcomes Res. 21, 858–873. https://doi.org/10.1016/j.jval.2017.11.010

Singal, A.G., Gupta, S., Skinner, C.S., Ahn, C., Santini, N.O., Agrawal, D., Mayorga, C.A., Murphy, C., Tiro, J.A., McCallister, K., Sanders, J.M., Bishop, W.P., Loewen, A.C., Halm, E.A., 2017. Effect of Colonoscopy Outreach vs Fecal Immunochemical Test Outreach on Colorectal Cancer Screening Completion: A Randomized Clinical Trial. JAMA 318, 806–815. https://doi.org/10.1001/jama.2017.11389

Singal, A.G., Gupta, S., Tiro, J.A., Skinner, C.S., McCallister, K., Sanders, J.M., Bishop, W.P., Agrawal, D., Mayorga, C.A., Ahn, C., Loewen, A.C., Santini, N.O., Halm, E.A., 2016. Outreach invitations for FIT and colonoscopy improve colorectal cancer screening rates: A randomized controlled trial in a safety-net health system. Cancer 122, 456–463. https://doi.org/10.1002/cncr.29770

Sisson, S.A., Fan, Y., Tanaka, M.M., 2007. Sequential Monte Carlo without likelihoods. PNAS 104, 1760–1765. https://doi.org/10.1073/pnas.0607208104

Starker, A., Buttmann-Schweiger, N., Kraywinkel, K., Kuhnert, R., 2017. The utilization of colonoscopy in Germany. J. Health Monit. 2. https://doi.org/10.17886/RKI-GBE-2017-126

Statistische Bundesamt, 2021. Früherkennung von Darmkrebs: fäkaler okkulter Bluttest (FOBT), Koloskopie [WWW Document]. URL https://www.gbe-bund.de/gbe/ (accessed 3.23.21).

Statistisches Bundesamt, 2021. Consumer price indices [WWW Document]. URL https://www.destatis.de/EN/Themes/Economy/Prices/Consumer-Price-Index/Tables/_tabellen-innen-consumer-price-indices.html (accessed 12.29.20).

Stinnett, A.A., Mullahy, J., 1998. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. Med. Decis. Mak. Int. J. Soc. Med. Decis. Mak. 18, S68-80. https://doi.org/10.1177/0272989X98018002S09

Stock, C., Ihle, P., Sieg, A., Schubert, I., Hoffmeister, M., Brenner, H., 2013. Adverse events requiring hospitalization within 30 days after outpatient screening and nonscreening colonoscopies. Gastrointest. Endosc. 77, 419–429. https://doi.org/10.1016/j.gie.2012.10.028

Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A., Bray, F., 2021. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA. Cancer J. Clin. https://doi.org/10.3322/caac.21660

Swart, E., Gothe, H., Geyer, S., Jaunzeme, J., Maier, B., Grobe, T.G., Ihle, P., German Society for Social Medicine and Prevention, German Society for Epidemiology, 2015. [Good Practice of Secondary Data Analysis (GPS): guidelines and recommendations]. Gesundheitswesen Bundesverb. Arzte Offentlichen Gesundheitsdienstes Ger. 77, 120–126. https://doi.org/10.1055/s-0034-1396815

Toes-Zoutendijk, E., Portillo, I., Hoeck, S., de Brabander, I., Perrin, P., Dubois, C., van Leerdam, M., Lansdorp-Vogelaar, I., Bardou, M., 2020. Participation in faecal immunochemical testing-based colorectal cancer screening programmes in the northwest of Europe. J. Med. Screen. 27, 68–76. https://doi.org/10.1177/0969141319879712

Toni, T., Welch, D., Strelkowa, N., Ipsen, A., Stumpf, M.P., 2009. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. J R Soc Interface 6, 187–202. https://doi.org/10.1098/rsif.2008.0172

Trautmann, F., Reißfelder, C., Pecqueux, M., Weitz, J., Schmitt, J., 2018. Evidence-based quality standards improve prognosis in colon cancer care. Eur. J. Surg. Oncol. J. Eur. Soc. Surg. Oncol. Br. Assoc. Surg. Oncol. 44, 1324–1330. https://doi.org/10.1016/j.ejso.2018.05.013

US Preventive Services Task Force, 2021. Screening for Colorectal Cancer: US Preventive Services Task Force Recommendation Statement. JAMA 325, 1965–1977. https://doi.org/10.1001/jama.2021.6238

USPSTF, 2020. US Preventive Services Task Force Draft Recommendation: Colorectal Cancer Screening [WWW Document]. URL https://uspreventiveservicestaskforce.org/uspstf/draft-recommendation/colorectal-cancer-screening3 (accessed 4.2.21).

van Ballegooijen, M., Rutter, C.M., Knudsen, A.B., Zauber, A.G., Savarino, J.E., Lansdorp-Vogelaar, I., Boer, R., Feuer, E.J., Habbema, J.D., Kuntz, K.M., 2011. Clarifying differences in natural history between models of screening: the case of colorectal cancer. Med Decis Mak. 31, 540–9. https://doi.org/10.1177/0272989X11408915

van Roon, A.H.C., Goede, S.L., van Ballegooijen, M., van Vuuren, A.J., Looman, C.W.N., Biermann, K., Reijerink, J.C.I.Y., Mannetje, H. 't, van der Togt, A.C.M., Habbema, J.D.F., van Leerdam, M.E., Kuipers, E.J., 2013. Random comparison of repeated faecal immunochemical testing at different intervals for population-based colorectal cancer screening. Gut 62, 409–415. https://doi.org/10.1136/gutjnl-2011-301583

Vanni, T., Karnon, J., Madan, J., White, R.G., Edmunds, W.J., Fossa, A.M., Legood, R., 2011. Calibrating Models in Economic Evaluation: A Seven-Step Approach. Pharmacoeconomics 29, 35–49. https://doi.org/10.2165/11584600-000000000-00000

Vieira, A.R., Abar, L., Chan, D.S.M., Vingeliene, S., Polemiti, E., Stevens, C., Greenwood, D., Norat, T., 2017. Foods and beverages and colorectal cancer risk: a systematic review and meta-analysis of cohort studies, an update of the evidence of the WCRF-AICR Continuous Update Project. Ann. Oncol. Off. J. Eur. Soc. Med. Oncol. 28, 1788–1802. https://doi.org/10.1093/annonc/mdx171

Vuik, F.E., Nieuwenburg, S.A., Bardou, M., Lansdorp-Vogelaar, I., Dinis-Ribeiro, M., Bento, M.J., Zadnik, V., Pellisé, M., Esteban, L., Kaminski, M.F., Suchanek, S., Ngo, O., Májek, O., Leja, M., Kuipers, E.J., Spaander, M.C., 2019. Increasing incidence of colorectal cancer in young adults in Europe over the last 25 years. Gut 68, 1820–1826. https://doi.org/10.1136/gutjnl-2018-317592

Weisenberger, D.J., Siegmund, K.D., Campan, M., Young, J., Long, T.I., Faasse, M.A., Kang, G.H., Widschwendter, M., Weener, D., Buchanan, D., Koh, H., Simms, L., Barker, M., Leggett, B., Levine, J., Kim, M., French, A.J., Thibodeau, S.N., Jass, J., Haile, R., Laird, P.W., 2006. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. Nat. Genet. 38, 787–793. https://doi.org/10.1038/ng1834

Weiser, M.R., 2018. AJCC 8th Edition: Colorectal Cancer. Ann. Surg. Oncol. 25, 1454–1455. https://doi.org/10.1245/s10434-018-6462-1

Whitlock, E.P., Lin, J.S., Liles, E., Beil, T.L., Fu, R., 2008. Screening for colorectal cancer: a targeted, updated systematic review for the U.S. Preventive Services Task Force. Ann. Intern. Med. 149, 638–658. https://doi.org/10.7326/0003-4819-149-9-200811040-00245

Whyte, S., Walsh, C., Chilcott, J., 2011. Bayesian calibration of a natural history model with application to a population model for colorectal cancer. Med Decis Mak. 31, 625–41. https://doi.org/10.1177/0272989X10384738

Wilkinson, D.J., 2018. Stochastic Modelling for Systems Biology, Third. ed. CRC Press, Boca Raton.

Wolf, A.M.D., Fontham, E.T.H., Church, T.R., Flowers, C.R., Guerra, C.E., LaMonte, S.J., Etzioni, R., McKenna, M.T., Oeffinger, K.C., Shih, Y.-C.T., Walter, L.C., Andrews, K.S., Brawley, O.W., Brooks, D., Fedewa, S.A., Manassaram-Baptiste, D., Siegel, R.L., Wender, R.C., Smith, R.A., 2018. Colorectal cancer screening for average-risk adults: 2018 guideline update from the American Cancer Society. CA. Cancer J. Clin. 68, 250–281. https://doi.org/10.3322/caac.21457

Wong, M.C.S., Ching, J.Y.L., Chan, V.C.W., Lam, T.Y.T., Luk, A.K.C., Wong, S.H., Ng, S.C., Ng, S.S.M., Wu, J.C.Y., Chan, F.K.L., Sung, J.J.Y., 2016. Colorectal Cancer Screening Based on Age and Gender: A Cost-Effectiveness Analysis. Medicine (Baltimore) 95, e2739. https://doi.org/10.1097/MD.0000000000002739

Yabroff, K.R., Lamont, E.B., Mariotto, A., Warren, J.L., Topor, M., Meekins, A., Brown, M.L., 2008. Cost of care for elderly cancer patients in the United States. J. Natl. Cancer Inst. 100, 630–641. https://doi.org/10.1093/jnci/djn103

Yothers, G., O'Connell, M.J., Allegra, C.J., Kuebler, J.P., Colangelo, L.H., Petrelli, N.J., Wolmark, N., 2011. Oxaliplatin as adjuvant therapy for colon cancer: updated results of NSABP C-07 trial, including survival and subset analyses. J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol. 29, 3768–3774. https://doi.org/10.1200/JCO.2011.36.4539

Zhao, S., Wang, S., Pan, P., Xia, T., Chang, X., Yang, X., Guo, L., Meng, Q., Yang, F., Qian, W., Xu, Z., Wang, Y., Wang, Z., Gu, L., Wang, R., Jia, F., Yao, J., Li, Z., Bai, Y., 2019. Magnitude, Risk Factors, and Factors Associated With Adenoma Miss Rate of Tandem Colonoscopy: A

Systematic Review and Meta-analysis. Gastroenterology 156, 1661-1674 e11. https://doi.org/10.1053/j.gastro.2019.01.260

Zorzi, M., Senore, C., Da Re, F., Barca, A., Bonelli, L.A., Cannizzaro, R., de Pretis, G., Di Furia, L., Di Giulio, E., Mantellini, P., Naldoni, C., Sassatelli, R., Rex, D.K., Zappa, M., Hassan, C., 2017. Detection rate and predictive factors of sessile serrated polyps in an organised colorectal cancer screening programme with immunochemical faecal occult blood test: the EQuIPE study (Evaluating Quality Indicators of the Performance of Endoscopy). Gut 66, 1233–1240. https://doi.org/10.1136/gutjnl-2015-310587

Zwink, N., Holleczek, B., Stegmaier, C., Hoffmeister, M., Brenner, H., 2017. Complication Rates in Colonoscopy Screening for Cancer. Dtsch. Arzteblatt Int. 114, 321–327. https://doi.org/10.3238/arztebl.2017.0321

9 APPENDIX

9.1 **Appendix A.** Supplementary information of colon cancer treatment cost analysis

Supplementary Table 1. Case and variable definition of the colon cancer treatment cost analysis

Variable	Definition
Incident diagnosis of colon cancer	≥1 inpatient diagnosis with ICD-10-GM C18 or C19 **and** no inpatient diagnosis (C18/C19) within at least 3 previous years **and** no outpatient medical accounts (C18/C19) before diagnosis (except for 12 months prior to inpatient diagnosis)
Cancer severity	
Low (proxy for UICC stage 1&2)	Incident diagnosis of colon cancer without chemotherapy (for definition, see procedures) within 6 months prior to or 12 months after initial resection. No diagnosis of distant metastasis (ICD-10 C78.0, C78.6, C78.7) concurrently with initial surgery.
Moderate (proxy for UICC stage 3)	Incident diagnosis of colon cancer with chemotherapy (definition see procedures) within 3 months prior to or 12 months after initial resection. No diagnosis of distant metastasis (ICD-10 C78.0, C78.6, C78.7) concurrently with initial surgery.
Advanced (proxy for UICC stage 4)	Incident diagnosis of colon cancer with chemotherapy (for definition, see procedures) within 6 months prior to or 12 months after initial resection and ▪ inpatient diagnosis of distant metastasis (ICD-10 C78.0, C78.7) or peritoneal carcinomatosis (ICD-10 C78.6) concurrently with initial treatment or ▪ bypass surgery (OPS 5-456.%) or ostomy (5-460.%) or colectomy (OPS 5-456.%) concurrently with initial treatment.
Comorbidities	
Hypertension	≥1 inpatient diagnosis or ≥2 outpatient diagnosis within two consecutive quarters with ICD-10 I10 or I11
Other cardiovascular diseases	≥1 inpatient diagnosis or ≥2 outpatient diagnosis within two consecutive quarters with ICD-10 I20-I25, I30-I39, I42, I48, I49, I50, Z94.3, Z95
Type 2 diabetes	≥1 inpatient diagnosis or ≥2 outpatient diagnosis within two consecutive quarters with ICD-10 E11 or E11 and ≥1 prescription of insulin (ATC A10A)
Kidney failure/chronic kidney disease	≥1 inpatient diagnosis or ≥2 outpatient diagnosis within two consecutive quarters with ICD-10 N17, N18 or N19

Note: The table is adapted from the case definition table by Trautmann et al. (2018)

9.2 **Appendix B.** Supplementary information of DECAS natural history model development, calibration and validation

Supplementary Table 2. Location of adenoma and clinically relevant serrated polyps (crSP)

Location	Adenoma (%)	crSP (%)	Reference
Cecum	11	17	(Greuter et al., 2014; Schramm et al., 2018)
Ascending colon	23	36.3	
Transverse colon	16	27.5	
Descending colon	11	15.1	
Sigmoid	21	2.5	
Rectum	18	1.6	

Supplementary Table 3. Clinical/screening-detected CRC stage distribution and observed survival by cancer stage

	Stage I	Stage II	Stage III	Stage IV	Reference
Clinical cancer stage distribution (%)	24.4	31.5	25.5	18.6	(Kubisch et al., 2016)
Screening-detected cancer stage distribution (%)	43.7	24.4	21.2	10.7	(Pox et al. 2012)
CRC observed survival (%)					(Munich Canncer Registry, 2018)
0-year	100	100	100	100	
1-year	94.1	89.7	88	55.7	
2-year	91.3	84.3	78.8	35.8	
3-year	88.1	78.7	70.9	23.2	
4-year	84.4	73.8	63.8	17.3	
5-year	80.6	68.7	58.4	13.3	
6-year	76.9	64.2	53.9	9.5	
7-year	73.2	59.9	50.4	8.1	
8-year	69.5	56.2	47.5	7.3	
9-year	65.9	53.1	45	6.7	
10-year	62	50.4	42.8	6.3	

Supplementary Table 4. German life table corrected by CRC mortality (2010-2014)

	Male			Female		
Age	**l(x)**	**d(x)**	**m(x)**	**l(x)**	**d(x)**	**m(x)**
20	99279	47.97	0.000483	99466	19.98	0.000201
21	99229	49.97	0.000504	99447	18.98	0.000191
22	99180	47.97	0.000484	99427	18.98	0.000191
23	99130	51.97	0.000524	99406	21.98	0.000221
24	99078	50.97	0.000514	99385	19.98	0.000201
25	99026	51.94	0.000524	99364	20.96	0.000211
26	98973	53.94	0.000545	99343	20.96	0.000211
27	98918	55.94	0.000565	99321	23.96	0.000241
28	98861	56.94	0.000576	99296	24.96	0.000251
29	98803	60.94	0.000617	99270	26.96	0.000272
30	98742	60.84	0.000616	99242	26.88	0.000271
31	98677	66.84	0.000677	99214	28.88	0.000291
32	98608	71.84	0.000729	99183	31.88	0.000321
33	98535	73.84	0.000749	99148	37.88	0.000382
34	98459	77.84	0.000791	99109	38.88	0.000392
35	98379	81.71	0.000831	99069	41.75	0.000421
36	98295	86.71	0.000882	99024	45.75	0.000462
37	98206	90.71	0.000924	98976	49.75	0.000503
38	98107	106.71	0.001088	98923	55.75	0.000564
39	97997	112.71	0.00115	98863	63.75	0.000645
40	97880	121.48	0.001241	98797	67.48	0.000683
41	97752	132.48	0.001355	98724	77.48	0.000785
42	97610	149.48	0.001531	98641	86.48	0.000877
43	97451	168.48	0.001729	98550	94.48	0.000959
44	97273	185.48	0.001907	98449	106.48	0.001082
45	97076	208.64	0.002149	98338	115.03	0.00117
46	96856	228.64	0.002361	98214	132.03	0.001344
47	96611	258.64	0.002677	98071	152.03	0.00155
48	96336	287.64	0.002986	97912	163.03	0.001665
49	96027	327.64	0.003412	97736	188.03	0.001924
50	95681	360.24	0.003765	97538	205.07	0.002102
51	95291	414.24	0.004347	97318	231.07	0.002374
52	94853	456.24	0.00481	97072	256.07	0.002638
53	94370	504.24	0.005343	96801	283.07	0.002924
54	93834	563.24	0.006003	96507	301.07	0.00312
55	93245	605.54	0.006494	96189	328.88	0.003419
56	92607	659.54	0.007122	95843	356.88	0.003724
57	91908	728.54	0.007927	95472	378.88	0.003968
58	91141	793.54	0.008707	95074	411.88	0.004332
59	90313	852.54	0.00944	94637	453.88	0.004796
60	89420	917.28	0.010258	94158	496.98	0.005278
61	88459	986.28	0.01115	93641	527.98	0.005638
62	87434	1043.28	0.011932	93088	567.98	0.006101
63	86350	1107.28	0.012823	92489	619.98	0.006703
64	85189	1194.28	0.014019	91837	672.98	0.007328
65	83946	1266.41	0.015086	91136	717.12	0.007869
66	82617	1361.41	0.016479	90377	784.12	0.008676

67	81203	1434.41	0.017665	89567	821.12	0.009168
68	79702	1536.41	0.019277	88704	890.12	0.010035
69	78110	1616.41	0.020694	87779	943.12	0.010744
70	76425	1715.49	0.022447	86787	1021.99	0.011776
71	74633	1824.49	0.024446	85718	1091.99	0.012739
72	72730	1936.49	0.026626	84562	1194.99	0.014132
73	70706	2067.49	0.029241	83301	1303.99	0.015654
74	68544	2211.49	0.032264	81911	1450.99	0.017714
75	66221	2376.38	0.035886	80354	1632.51	0.020316
76	63723	2551.38	0.040039	78606	1825.51	0.023224
77	61044	2740.38	0.044892	76654	2041.51	0.026633
78	58169	2943.38	0.050601	74469	2291.51	0.030771
79	55082	3163.38	0.05743	72027	2556.51	0.035494
80	51814	3290.02	0.063497	69318	2812.38	0.040572
81	48362	3516.02	0.072702	66323	3116.38	0.046988
82	44727	3656.02	0.081741	63012	3445.38	0.054678
83	40971	3759.02	0.091748	59390	3736.38	0.062913
84	37137	3810.02	0.102594	55477	4028.38	0.072614
85	33278	3835.07	0.115243	51289	4299.14	0.083822
86	29429	3816.07	0.12967	46853	4536.14	0.096816
87	25636	3721.07	0.14515	42231	4672.14	0.110633
88	21970	3563.07	0.162179	37500	4754.14	0.126777
89	18487	3355.07	0.181483	32740	4728.14	0.144415
90	15239	3093.07	0.202971	28030	4656.14	0.166113

Note: d(x): Number of deaths between ages x and x+n; l(x): Number of survivors at exact age x, assuming l(0) = 100,000; m(x): Central death rate between ages x and x+n

Source: (Human Mortality Database, 2018)

Supplementary Table 5. Epidemiological data served as calibration targets in DECAS by age, sex and location

Variables	Observed data*		Reference
	Mean	95% CI	
Non-advanced pre-cancer lesion prevalence (%)			
Non-advanced adenoma			
Male, age 45-49 years	15.3		(Leshno et al., 2016; Kretschmann et al., 2020)***
Male, age 55-59 years	26	(24.6, 27.5)	
Male, age 60-64 years	29.7	(28.3, 31.1)	
Male, age 65-69 years	31.5	(29.8, 33.1)	
Male, age 70-74 years	31.6	(30, 33.2)	
Male, age 75-79 years	30.4	(28.5, 32.2)	
Male, age 80+ years	27.6	(26.5, 28.8)	
Female, age 45-49 years	9.6		
Female, age 55-59 years	16.3	(15.3, 17.3)	
Female, age 60-64 years	19.5	(18.4, 20.6)	
Female, age 65-69 years	21.3	(20.1, 22.4)	
Female, age 70-74 years	22.1	(20.8, 23.4)	
Female, age 75-79 years	21.8	(20.3, 23.4)	
Female, age 80+ years	20.4	(19.3, 21.5)	
Non-clinically relevant serrated polyp			
Male, age 40-49 years	12		(Leshno et al., 2016; Schramm et al., 2018)
Male, age 50-59 years	20.4	(5.1, 23.2)	
Male, age 60-69 years	17	(10.3, 28.5)	
Male, age 70-79 years	15.9	(12.2, 29.7)	
Male, age 80+ years	22.8	(2.2, 51.9)	
Female, age 40-49 years	10.4		
Female, age 50-59 years	17.7	(7.3, 32.1)	
Female, age 60-69 years	16.3	(10.7, 32.2)	
Female, age 70-79 years	12	(5.5, 20.8)	
Female, age 80+ years	9.3	(0, 14)	
Advanced pre-cancer lesion prevalence (%)			
Advanced adenoma			
Male, age 45-49 years	3.2		(Leshno et al., 2016; Kretschmann et al., 2020)***
Male, age 55-59 years	7.4	(7.1, 7.8)	
Male, age 60-64 years	9.9	(9.6, 10.2)	
Male, age 65-69 years	10.9	(10.5, 11.2)	
Male, age 70-74 years	11.2	(10.4, 12)	
Male, age 75-79 years	11.7	(10.9, 12.6)	
Male, age 80+ years	11.7	(10.7, 12.8)	
Female, age 45-49 years	1.8		
Female, age 55-59 years	4.1	(4, 4.3)	
Female, age 60-64 years	5.4	(5.3, 5.6)	
Female, age 65-69 years	6.2	(6, 6.5)	
Female, age 70-74 years	6.8	(6.3, 7.2)	
Female, age 75-79 years	7.5	(7, 8)	
Female, age 80+ years	8.5	(7.9, 9.1)	
Clinically relevant serrated polyp			
Male, age 40-49 years	3.3		(Leshno et al., 2016; Schramm et al., 2018)
Male, age 50-59 years	7.8	(0.7, 15.4)	
Male, age 60-69 years	10	(0.9, 15.4)	
Male, age 70-79 years	7.9	(1.4, 13.7)	
Male, age 80+ years	7.3	(0, 28.7)	
Female, age 40-49 years	4.1		
Female, age 50-59 years	9.6	(2, 15.3)	
Female, age 60-69 years	8	(0.4, 11.6)	
Female, age 70-79 years	9.5	(1.4, 16.7)	
Female, age 80+ years	6	(0, 20.1)	

(Continue Supplementary Table 5)

Variables (continued.)	Observed data		Reference
	Mean	95% CI	
CRC prevalence (per 100,000)			
Colorectal cancer			
Male, age 45-49 years	284.5		(Leshno et al., 2016)***
Female, age 45-49 years	139.4		
Colon cancer (proximal + distal colon)**			
Male, age 55-59 years	393.2	(168.2, 618.2)	
Male, age 60-64 years	586.1	(362.4, 809.8)	
Male, age 65-69 years	781.8	(601.5, 962)	
Male, age 70-74 years	1107.2	(924.9, 1289.4)	
Male, age 75-79 years	1605	(1257.6, 1952.5)	
Male, age 80+ years	2141.2	(1624.2, 2658.2.8)	(Kretschmann et al., 2020)***
Female, age 55-59 years	228.8	(134.2, 323.5)	
Female, age 60-64 years	301.5	(219.5, 383.6)	
Female, age 65-69 years	411.4	(314.5, 508.2)	
Female, age 70-74 years	662.3	(555.9, 768.8)	
Female, age 75-79 years	1020.8	(852.5, 1189)	
Female, age 80+ years	1727.1	(1291.7, 2162.5)	
Rectal cancer**			
Male, age 55-59 years	269.7	(217.4, 322)	
Male, age 60-64 years	378.7	(302.6, 454.8)	
Male, age 65-69 years	453.9	(343.5, 564.3)	
Male, age 70-74 years	604.1	(467.9, 740.2)	
Male, age 75-79 years	653.1	(557.5, 748.7)	
Male, age 80+ years	1075.9	(810.6, 1341.2)	(Kretschmann et al., 2020)***
Female, age 55-59 years	95.9	(74, 117.7)	
Female, age 60-64 years	151.6	(94.8, 208.4)	
Female, age 65-69 years	206.3	(112.2, 300.4)	
Female, age 70-74 years	281.1	(150.7, 411.4)	
Female, age 75-79 years	424.6	(262.8, 586.5)	
Female, age 80+ years	790.1	(584.8, 995.4)	

Note: CI = confidence interval

* Prevalence were corrected with colonoscopy miss rate from a meta-analysis (Zhao et al., 2019)

** Eighty-five percent of the CRC prevalence was used to calibrate the transition from advanced adenoma to CRC, and the other 15% was used for clinically relevant serrated polyps to CRC

*** Adenoma prevalence were taken from the data in the period 2007-2014 from German screening colonoscopy registry, while CRC prevalence were from the period 2003-2006

Supplementary Table 1. CRC incidence used for DECAS natural history model validation (per 100,000)

Variables	CRC incidence	Reference
Male		
Age 20-24 years	1.2	
Age 25-29 years	2.2	
Age 30-34 years	4.2	
Age 35-39 years	7.8	
Age 40-44 years	14.0	
Age 45-49 years	29.3	
Age 50-54 years	56.7	(Robert Koch Institute, 2019)*
Age 55-59 years	113.9	
Age 60-64 years	188.4	
Age 65-69 years	269.8	
Age 70-74 years	370.2	
Age 75-79 years	459.3	
Age 80+ years	565	
Female		
Age 20-24 years	1.9	
Age 25-29 years	2.4	
Age 30-34 years	4.5	
Age 35-39 years	7.6	
Age 40-44 years	13.3	
Age 45-49 years	23.9	
Age 50-54 years	43.9	(Robert Koch Institute, 2019)*
Age 55-59 years	71.9	
Age 60-64 years	114.8	
Age 65-69 years	153.9	
Age 70-74 years	212.3	
Age 75-79 years	295.6	
Age 80+ years	416.3	

* CRC incidence 2003-2006 from the German Centre for Cancer Registry Data

9.3 **Appendix C.** Supplementary information of the cost-effectiveness analysis of German CRC screening program

Supplementary Table 6. Modeled life-year, quality adjusted life-years and costs using 0% discount rates per 1,000 people for a cohort of 40 years old

Strategy[1,2]	dLYs	dQALYs	dCosts	Wasted dCosts due to unused FITs
No Screening	36,005 (35,621-36,326)	33,307 (32,344-34,274)	2,626,782(1,633,300-3,980,411)	-
Scenario 1 (perfect adherence)				
FIT1y50+FIT2y55	36,189 (35,884-36,467)	33,445 (32,514-34,371)	1,838,668 (1,233,951-2,676,209)	-
FIT1y50+COL10y55	36,221 (35,929-36,490)	33,484 (32,547-34,397)	1,489,420 (1,056,858-2,107,263)	-
mCOL50/fFIT50+COL55	36,258 (35,955-36,514)	33,505 (32,564-34,446)	1,374,787 (972,988-1,958,548)	-
FIT2y50	36,184 (35,882-36,460)	33,441 (32,509-34,381)	1,857,638 (1,243,754-2,663,880)	-
COL10y50	36,262 (35,962-36,521)	33,521 (32,584-34,441)	1,287,924 (908,407-1,850,327)	-
COL10y50-3X	36,261 (35,976-36,520)	33,523 (32,583-34,455)	1,337,142 (989,131-1,860,213)	-
FIT1y45+COL10y50	36,272 (35,996-36,533)	33,524 (32,585-34,444)	1,394,490 (1,010,288-1,947,709)	-
FIT1y45+COL10y50-3X	36,270 (35,992-36,535)	33,523 (32,590-34,451)	1,458,042 (1,098,174-2,001,738)	-
FIT2y45	36,219 (35,954-36,848)	33,468 (32,520-34,387)	1,779,983 (1,201,066-2,578,431)	-
COL10y45-3X	36,296 (36,023-36,532)	33,550 (32,605-34,474)	1,327,711 (995,935-1,809,848)	-
Scenario 2 (current program)				
FIT1y50+FIT2y55	36,066 (35,707-36,367)	33,358 (32,410-34,307)	2,281,753 (1,463,867-3,445,946)	-
FIT1y50+COL10y55	36,095 (35,741-36,395)	33,383 (32,441-34,317)	2,094,987 (1,343,376-3,136,872)	-
mCOL50/fFIT50+COL55	36,106 (35,754-36,410)	33,387 (32,446-34,332)	2,089,260 (1,346,939-3,119,333)	-
FIT2y50	36,069 (35,707-36,377)	33,355 (32,406-34,301)	2,297,653 (1,469,789-3,448,979)	-
COL10y50	36,102 (35,757-36,405)	33,386 (32,438-34,323)	2,092,204 (1,345,466-3,123,954)	-
COL10y50-3X	36,107 (35,751-36,406)	33,389 (32,444-34,338)	2,090,488 (1,359,201-3,103,488)	-
FIT1y45+COL10y50	36,106 (35,767-36,393)	33,390 (32,449-34,332)	2,080,001 (1,349,215-3,121,392)	-
FIT1y45+COL10y50-3X	36,109 (35,767-36,393)	33,393 (32,453-34,332)	2,080,686 (1,344,832-3,074,003)	-
FIT2y45	36,071 (35,734-36,388)	33,357 (32,419-34,303)	2,288,299 (1,469,164-3,427,234)	-
COL10y45-3X	36,115 (35,776-36,413)	33,395 (32,447-34,341)	2,092,704 (1,356,606-3,129,569)	-

(Continue Supplementary Table 6)

Strategy[1,2]	dLYs	dQALYs	dCosts	Wasted dCosts due to unused FITs
Scenario 3 (mail-out FIT)				
FIT1y50+FIT2y55	36,079 (35,744-36,380)	33,363 (32,427-34,305)	2,459,887 (1,668,130-3,591,184)	158,165 (155,711-159,770)
FIT1y50+COL10y55	36,103 (35,759-36,403)	33,385 (32,440-34,321)	2,151,768 (1,417,931-3,178,955)	53,706 (53,227-54,067)
mCOL50/fFIT50+COL55	36,104 (35,772-36,409)	33,385 (32,450-34,315)	2,121,290 (1,371,236-3,162,219)	26,012 (25,721-26,260)
FIT2y50	36,079 (35,729-36,376)	33,361 (32,409-34,306)	2,442,382 (1,638,123-3,553,971)	127,452 (125,423-128,861)
COL10y50	36,102 (35,757-36,405)	33,386 (32,438-34,323)	2,092,204 (1,345,466-3,123,954)	-
COL10y50-3X	36,107 (35,751-36,406)	33,389 (32,444-34,338)	2,090,488 (1,359,201-3,103,488)	-
FIT1y45+COL10y50	36,117 (35,785-36,412)	33,396 (32,445-34,328)	2,134,930 (1,405,151-3,156,460)	54,638 (54,248-54,919)
FIT1y45+COL10y50-3X	36,115 (35,774-36,417)	33,396 (32,461-34,336)	2,136,212 (1,423,410-3,138,715)	54,640 (54,260-54,920)
FIT2y45	36,081 (35,709-36,384)	33,364 (32,414-34,303)	2,472,935 (1,674,735-3,609,673)	158,681 (156,210-160,342)
COL10y45-3X	36,115 (35,776-36,413)	33,395 (32,447-34,341)	2,092,704 (1,356,606-3,129,569)	-
Scenario 4 (high adherence)				
FIT1y50+FIT2y55	36,140 (35,830-36,440)	33,409 (32,471-34,335)	2,256,561 (1,546,333-3,245,599)	60,581 (59,635-61,357)
FIT1y50+COL10y55	36,153 (35,840-36,429)	33,424 (32,493-34,369)	1,939,484 (1,305,706-2,804,960)	20,562 (20,287-20,809)
mCOL50/fFIT50+COL55	36,139 (35,815-36,426)	33,420 (32,478-34,361)	1,926,741 (1,285,439-2,840,137)	9,852 (9,683-10,012)
FIT2y50	36,127 (35,780-36,426)	33,397 (32,455-34,345)	2,261,704 (1,535,871-3,290,777)	48,767 (47,834-49,504)
COL10y50	36,139 (35,805-36,427)	33,417 (32,476-34,346)	1,907,841 (1,252,697-2,837,455)	-
COL10y50-3X	36,144 (35,832-36,430)	33,421 (32,483-34,362)	1,905,340 (1,264,564-2,798,360)	-
FIT1y45+COL10y50	36,176 (35,882-36,455)	33,444 (32,495-34,373)	1,906,581 (1,301,166-2,790,273)	20,918 (29,649-21,143)
FIT1y45+COL10y50-3X	36,178 (35,866-36,453)	33,445 (32,508-34,371)	1,916,207 (1,331,370-2,780,280)	20,919 (20,681-21,142)
FIT2y45	36,147 (35,833-36,427)	33,412 (32,476-34,340)	2,257,748 (1,564,517-3,230,424)	60,739 (59,704-61,642)
COL10y45-3X	36,157 (35,846-36,439)	33,431 (32,493-34,365)	1,910,283 (1,291,985-2,790,800)	-

Note: COL = colonoscopy; dLY = discounted life-years; dQALY = discounted quality-adjusted life-years; FIT = fecal immunochemical test.

(1) The life-years and quality-adjusted life-years were discounted with 0% annual rates.

(2) Results are presented as mean and 95% credible interval.

Supplementary Table 7. Modeled life-year, quality adjusted life-years and costs using 5% discount rates per 1,000 people for a cohort of 40 years old

Strategy[1,2]	dLYs	dQALYs	dCosts	Wasted dCosts due to unused FITs
No Screening	16,742 (16,657-16,793)	14,290 (13,885-14,689)	624,976 (347,696-1,079,426)	-
Scenario 1 (perfect adherence)				
FIT1y50+FIT2y55	16,774 (16,714-16,811)	14,306 (13,908-14,708)	541,065 (348,721-850,622)	-
FIT1y50+COL10y55	16,779 (16,714-16,816)	14,317 (13,921-14,712)	538,814 (374,591-820,147)	-
mCOL50/fFIT50+COL55	16,783 (16,727-16,821)	14,327 (13,925-14,730)	513,741 (360,808-782,790)	-
FIT2y50	16,773 (16,710 (16,811)	14,305 (13,908-14,707)	534,722 (339,093-847,007)	-
COL10y50	16,787 (16,732-16,822)	14,335 (13,934-14,730)	497,271 (355,537-748,335)	-
COL10y50-3X	16,787 (16,729-16,823)	14,335 (13,931-14,731)	516,902 (375,318-775,680)	-
FIT1y45+COL10y50	16,790 (16,742-16,823)	14,331 (13,928-14,725)	556,531 (416,286-804,281)	-
FIT1y45+COL10y50-3X	16,789 (16,736-16,823)	14,330 (13,927-14,728)	575,516 (436,371-833,494)	-
FIT2y45	16,781 (16,727-16,816)	14,312 (13,907-14,711)	539,477 (363,003-829,400)	-
COL10y45-3X	16,795 (16,746-16,825)	14,345 (13,943-14,739)	565,466 (441,426-790,600)	-
Scenario 2 (current program)				
FIT1y50+FIT2y55	16,757 (16,687-16,803)	14,297 (13,893-14,702)	546,959 (317,513-906,651)	-
FIT1y50+COL10y55	16,760 (16,695-16,806)	14,303 (13,897-14,706)	531,544 (312,581-891,598)	-
mCOL50/fFIT50+COL55	16,762 (16,690-16,805)	14,305 (13,902-14,711)	530,537 (314,736-885,247)	-
FIT2y50	16,756 (16,685-16,799)	14,297 (13,891-14,696)	546,791 (312,584-915,523)	-
COL10y50	16,762 (16,694-16,806)	14,306 (13,900-14,705)	529,049 (311,279-875,461)	-
COL10y50-3X	16,762 (16,695-16,805)	14,306 (13,901-14,710)	532,302 (318,971-877,599)	-
FIT1y45+COL10y50	16,762 (16,693-16,806)	14,306 (13,902-14,705)	535,428 (322,791-881,249)	-
FIT1y45+COL10y50-3X	16,763 (16,695-16,807)	14,306 (10,904-14,711)	539,359 (324,820-883,661)	-
FIT2y45	16,756 (16,687-16,802)	14,297 (13,894-14,699)	548,816 (319,914-915,595)	-
COL10y45-3X	16,764 (16,695-16,807)	14,308 (13,903-14,707)	544,141 (334,464-886,674)	-

(Continue Supplementary Table 7)

Strategy[1,2]	dLYs	dQALYs	dCosts	Wasted dCosts due to unused FITs
Scenario 3 (mail-out FIT)				
FIT1y50+FIT2y55	16,757 (16,686-16,802)	14,297 (13,897-14,698)	631,954 (401,825-986,528)	63,156 (62,319-63,693)
FIT1y50+COL10y55	16,761 (16,696-16,807)	14,303 (13,899-14,702)	573,040 (358,578-925,894)	29,989 (29,724-30,190)
mCOL50/fFIT50+COL55	16,761 (16,694-16,804)	14,304 (10,903-14,702)	551,686 (337,584-897,408)	14,523 (14,361-14,662)
FIT2y50	16,757 (16,689-16,804)	14,297 (13,891-14,699)	609,780 (381,633-978,239)	47,870 (47,197-48,334)
COL10y50	16,762 (16,694-16,806)	14,306 (13,900-14,705)	529,049 (311,279-875,461)	-
COL10y50-3X	16,762 (16,695-16,805)	14,306 (13,901-14,710)	532,302 (318,971-877,599)	-
FIT1y45+COL10y50	16,764 (16,697-16,805)	14,307 (13,898-14,706)	588,053 (378,945-932,984)	38,932 (38,657-39,132)
FIT1y45+COL10y50-3X	16,764 (16,696-16,806)	14,306 (13,905-14,707)	591,579 (383,055-924,450)	38,934 (38,665-39,132)
FIT2y45	16,758 (16,689-16,805)	14,298 (13,894-14,704)	639,536 (416,153-1,004,224)	68,516 (67,644-69,106)
COL10y45-3X	16,764 (16,695-16,807)	14,308 (13,903-14,707)	544,141 (334,464-886,674)	-
Scenario 4 (high adherence)				
FIT1y50+FIT2y55	16,767 (16,705-16,809)	14,302 (13,898-14,701)	630,120 (423,995-962,895)	24,184 (23,837-24,463)
FIT1y50+COL10y55	16,769 (16,706-16,808)	14,308 (13,905-14,711)	572,528 (379,868-891,883)	11,482 (11,329-11,620)
mCOL50/fFIT50+COL55	16,768 (16,700-16,808)	14,310 (13,909-14,711)	547,570 (352,108-870,736)	5,501 (5,406-5,590)
FIT2y50	16,764 (16,700-16,808)	14,300 (13,897-14,702)	609,092 (390,726-390,726)	18,318 (18,011-18,570)
COL10y50	16,767 (16,699-16,808)	14,312 (13,908-14,714)	520,472 (325,006-849,327)	-
COL10y50-3X	16,768 (16,704-16,810)	14,313 (13,914 (14,711)	526,827 (330,653-843,144)	-
FIT1y45+COL10y50	16,775 (16,717-16,814)	14,316 (13,909-14,719)	590,666 (410,718-887,128)	14,906 (14,714-15,065)
FIT1y45+COL10y50-3X	16,774 (16,714-16,813)	14,315 (13,914-14,711)	598,031 (415,437-904,181)	14,906 (14,738-15,064)
FIT2y45	16,769 (16,706-16,809)	14,304 (13,903-14,700)	638,534 (437,932-957,840)	26,229 (25,841-26,580)
COL10y45-3X	16,771 (16,709-16,812)	14,317 (13,915-14,716)	547,886 (362,209-853,432)	-

Note: COL = colonoscopy; dLY = discounted life-years; dQALY = discounted quality-adjusted life-years; FIT = fecal immunochemical test.

(1) The life-years and quality-adjusted life-years were discounted with 5% annual rates.

(2) Results are presented as mean and 95% credible interval.

Supplementary Table 8. Incremental cost-effectiveness ratio and incremental number-needed-to colonoscope between screening colonoscopy strategies with and without differentiated starting age for different sex – with 0% discount rate

Strategy[1,2,3,4]	ΔCosts	By LY		By QALY	
		LYG	ICER	QALYG	ICER
Scenario 1 (perfect adherence)					
FIT1y50+COL10y55	--	--	--	--	--
mCOL50/fFIT50+COL55	CS (CS-CS)	37 (25-26)	Dominant (D-D)	21 (17-49)	Dominant (D-D)
Scenario 2 (current program)					
FIT1y50+COL10y55	--	--	--	--	--
mCOL50/fFIT50+COL55	CS (CS-3,563)	11 (13-15)	Dominant (276-D)	5 (4-15)	Dominant (809-D)
Scenario 3 (mail-out FIT)					
FIT1y50+COL10y55	--	--	--	--	--
mCOL50/fFIT50+COL55	CS (CS-CS)	1 (6-13)	Dominant (D-D)	1 (-6-10)	Dominant (2,599-D)
Scenario 4 (high adherence)					
FIT1y50+COL10y55	--	--	--	--	--
mCOL50/fFIT50+COL55	CS (CS-35,177)	-14 (-4- -25)	929 (D'ed-809)	-3 (-8- -16)	3,820 (D'ed-1,287)

Note: -- = reference; Δ = difference; CS = cost-saving; D = dominant; D'ed = dominated; ICER = incremental cost-effectiveness ratio; LY = life-years; QALY= Quality-adjusted life-years.

(1) The costs, life-years and quality-adjusted life-years are expressed as per 1,000 40-year-olds and were discounted with 0% annual rates. Costs were in 2021 Euro.

(2) FIT1y50+COL10y55 was used as the reference.

(3) The results are presented as: difference between means of the two strategies and difference between the upper value and lower values of 95% credible intervals of the two strategies.

(4) "Dominant" denotes a strategy being more effective and cost-saving; "Dominated" denotes a strategy being less effective and costing more.

Supplementary Table 9. Incremental cost-effectiveness ratio and incremental number-needed-to colonoscope between screening colonoscopy strategies with and without differentiated starting age for different sex – with 5% discount rate

Strategy[1,2,3,4]	ΔCosts	By LY		By QALY	
		LYG	ICER	QALYG	ICER
Scenario 1 (perfect adherence)					
FIT1y50+COL10y55	--	--	--	--	--
mCOL50/fFIT50+COL55	CS (CS-CS)	4 (5-9)	Dominant (D-D)	10 (4-18)	Dominant (D-D)
Scenario 2 (current program)					
FIT1y50+COL10y55	--	--	--	--	--
mCOL50/fFIT50+COL55	CS (CS-1,149)	1 (-5-0)	Dominant (D'ed-22,147)	2 (4-5)	Dominant (486-D)
Scenario 3 (mail-out FIT)					
FIT1y50+COL10y55	--	--	--	--	--
mCOL50/fFIT50+COL55	CS (CS-CS)	0 (-2- -3)	Dominant (8,680-8,761)	2 (0-4)	Dominant (113,806-D)
Scenario 4 (high adherence)					
FIT1y50+COL10y55	--	--	--	--	--
mCOL50/fFIT50+COL55	CS (CS-CS)	-1 (-6-0)	27,071 (4,286-D)	2 (0-4)	Dominant (81,515-D)

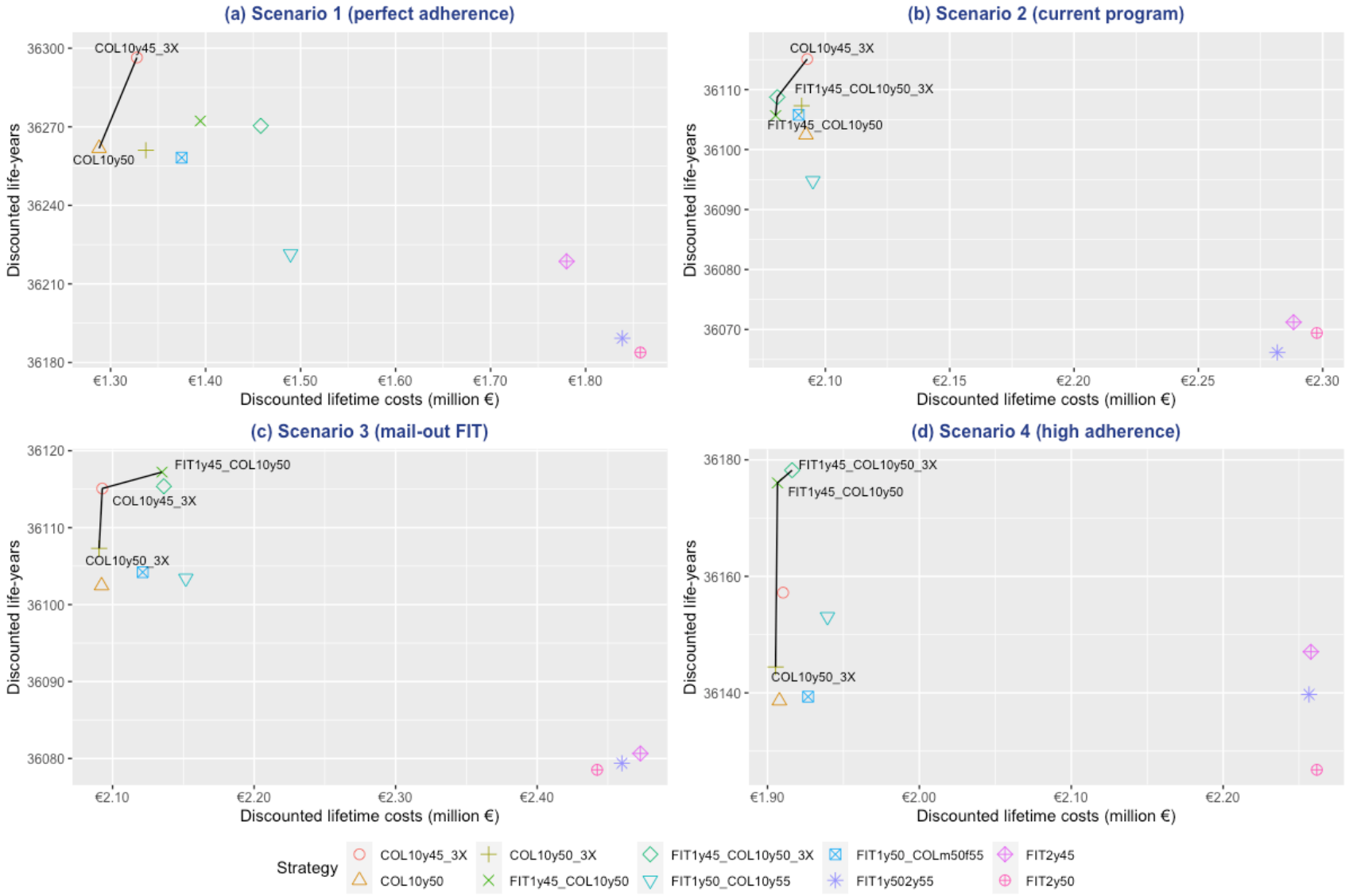
Note: -- = reference; Δ = difference; CS = cost-saving; D = dominant; D'ed = dominated; ICER = incremental cost-effectiveness ratio; LY = life-years; QALY= Quality-adjusted life-years.

(1) The costs, life-years and quality-adjusted life-years are expressed as per 1,000 40-year-olds and were discounted with 5% annual rates. Costs were in 2021 Euro.

(2) FIT1y50+COL10y55 was used as the reference.

(3) The results are presented as: difference between means of the two strategies and difference between the upper value and lower values of 95% credible intervals of the two strategies.

(4) "Dominant" denotes a strategy being more effective and cost-saving; "Dominated" denotes a strategy being less effective and costing more.



Supplementary Figure 1. Efficiency frontier based on cost-effectiveness (life-years as benefits, 0% discount rate).
Texts shown in the graphs mark the dominant strategies in each scenario, which will be used to calculate incremental cost-effectiveness ratios (ICER).

Supplementary Table 10. Cost-effectiveness and incremental cost-effectiveness ratios (ICERs) between efficient strategies (life-years as benefits, 0% discount rate)

Strategy[1,2,3,4]	Compared to no screening			ICER		
	ΔCosts	LYG	CER	ΔCosts	LYG	ICER
Scenario 1 (perfect adherence)						
COL10y50	Cost-saving (CS-CS)	257 (195-341)	Dominant (D-D)	--	--	--
COL10y45-3X	Cost-saving (CS-CS)	291 (207-402)	Dominant (D-D)	39,787 (CS-87,528)	35 (17-60)	1,148 (1,447-D)
Scenario 2 (current program)						
FIT1y45+COL10y50	Cost-saving (CS-CS)	100 (67-147)	Dominant (D-D)	--	--	--
FIT1y45+COL10y50-3X	Cost-saving (CS-CS)	103 (67-146)	Dominant (D-D)	685 (CS-CS)	3 (0-0)	222 (10,904-D)
COL10y45-3X	Cost-saving (CS-CS)	110 (87-155)	Dominant (D-D)	12,018 (11,775-55,567)	6 (9-20)	1,894 (1,343-2,827)
Scenario 3 (mail-out FIT)						
COL10y50-3X	Cost-saving (CS-CS)	102 (80-131)	Dominant (D-D)	--	--	--
COL10y45-3X	Cost-saving (CS-CS)	110 (87-155)	Dominant (D-D)	2,216 (CS-26,081)	8 (7-24)	2,216 (D'ed-26,081)
FIT1y45+COL10y50	Cost-saving (CS-CS)	112 (86-164)	Dominant (D-D)	42,226 (26,891-48,545)	2 (-1-9)	42,226 (26,891-48,545)
Scenario 4 (high adherence)						
COL10y45-3X	Cost-saving (CS-CS)	139 (104-211)	Dominant (D-D)	--	--	--
FIT1y45+COL10y50	Cost-saving (CS-CS)	171 (129-261)	Dominant (D-D)	1,240 (CS-36,601)	31 (25-50)	1,240 (D'ed-36,601)
FIT1y45+COL10y50-3X	Cost-saving (CS-CS)	173 (128-246)	Dominant (D-D)	9,626 (CS-30,204)	2 (-15- -2)	9,626 (D'ed-30,204)

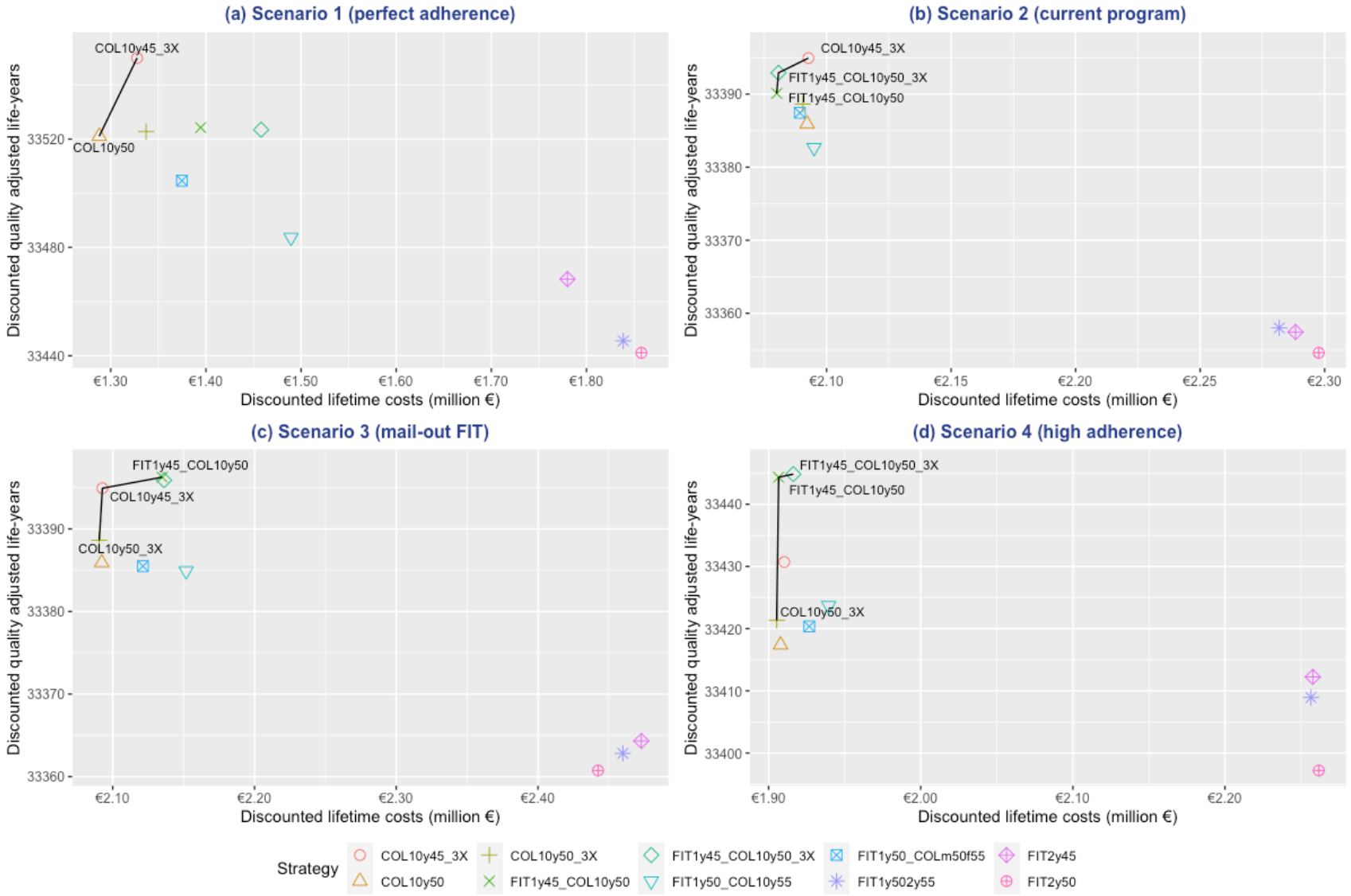
Note: -- = reference; Δ = difference; CER = cost-effectiveness ratio; CS = cost-saving; D = dominant; D'ed = dominated; ICER = incremental cost-effectiveness ratio; LYG = life-years gained.

(1) The costs and life-years are expressed as per 1,000 40-year-olds were discounted with 0% annual rates. Costs were in 2021 Euro.

(2) CERs were calculated by comparing each strategy with no-screening strategy. ICERs were calculated by comparing each strategy with the next most effective one (the rows next to each other).

(3) The results are presented as: difference between means of the two strategies and difference between the upper value and lower values of 95% credible intervals of the two strategies.

(4) "Dominant" denotes a strategy being more effective and cost-saving; "Dominated" denotes a strategy being less effective and costing more.



Supplementary Figure 2. Efficiency frontier based on cost-effectiveness (quality adjusted life-years as benefits, 0% discount rate).
Texts shown in the graphs mark the dominant strategies in each scenario, which will be used to calculate incremental cost-effectiveness ratios (ICER).

Supplementary Table 11. Cost-effectiveness and incremental cost-effectiveness ratios (ICERs) between efficient strategies (quality-adjusted life-years as benefits, 0% discount rate)

Strategy[1,2,3,4]	Compared to no screening			Per QALY		
	ΔCosts	QALYG	CER	ΔCosts	QALYG	ICER
Scenario 1 (perfect adherence)						
COL10y50	Cost-saving (CS-CS)	214 (167-240)	Dominant (D-D)	--	--	--
COL10y45-3X	Cost-saving (CS-CS)	243 (200-261)	Dominant (D-D)	39,787 (CS-87,528)	29 (21-33)	1,378 (4,072-D)
Scenario 2 (current program)						
FIT1y45+COL10y50	Cost-saving (CS-CS)	83 (58-104)	Dominant (D-D)	--	--	--
FIT1y45+COL10y50-3X	Cost-saving (CS-CS)	86 (58-108)	Dominant (D-D)	685 (CS-CS)	3 (0-4)	243 (D-D)
COL10y45-3X	Cost-saving (CS-CS)	88 (68-102)	Dominant (D-D)	12,018 (11,775-55,567)	5 (-2-9)	2,626 (D'ed-887)
Scenario 3 (mail-out FIT)						
COL10y50-3X	Cost-saving (CS-CS)	82 (64-100)	Dominant (D-D)	--	--	--
COL10y45-3X	Cost-saving (CS-CS)	88 (68-102)	Dominant (D-D)	2,216 (CS-26,081)	6 (2-3)	351 (7,529-D)
FIT1y45+COL10y50	Cost-saving (CS-CS)	89 (55-101)	Dominant (D-D)	42,226 (26,891-48,545)	1 (-13- -1)	31,545 (D'ed-D'ed)
Scenario 4 (high adherence)						
COL10y45-3X	Cost-saving (CS-CS)	115 (89-138)	Dominant (D-D)	--	--	--
FIT1y45+COL10y50	Cost-saving (CS-CS)	138 (99-150)	Dominant (D-D)	1,240 (CS-36,601)	23 (10-12)	54 (3,147-D)
FIT1y45+COL10y50-3X	Cost-saving (CS-CS)	138 (97-163)	Dominant (D-D)	9,626 (CS-30,204)	1 (-2-13)	19,196 (2,273-5,495)

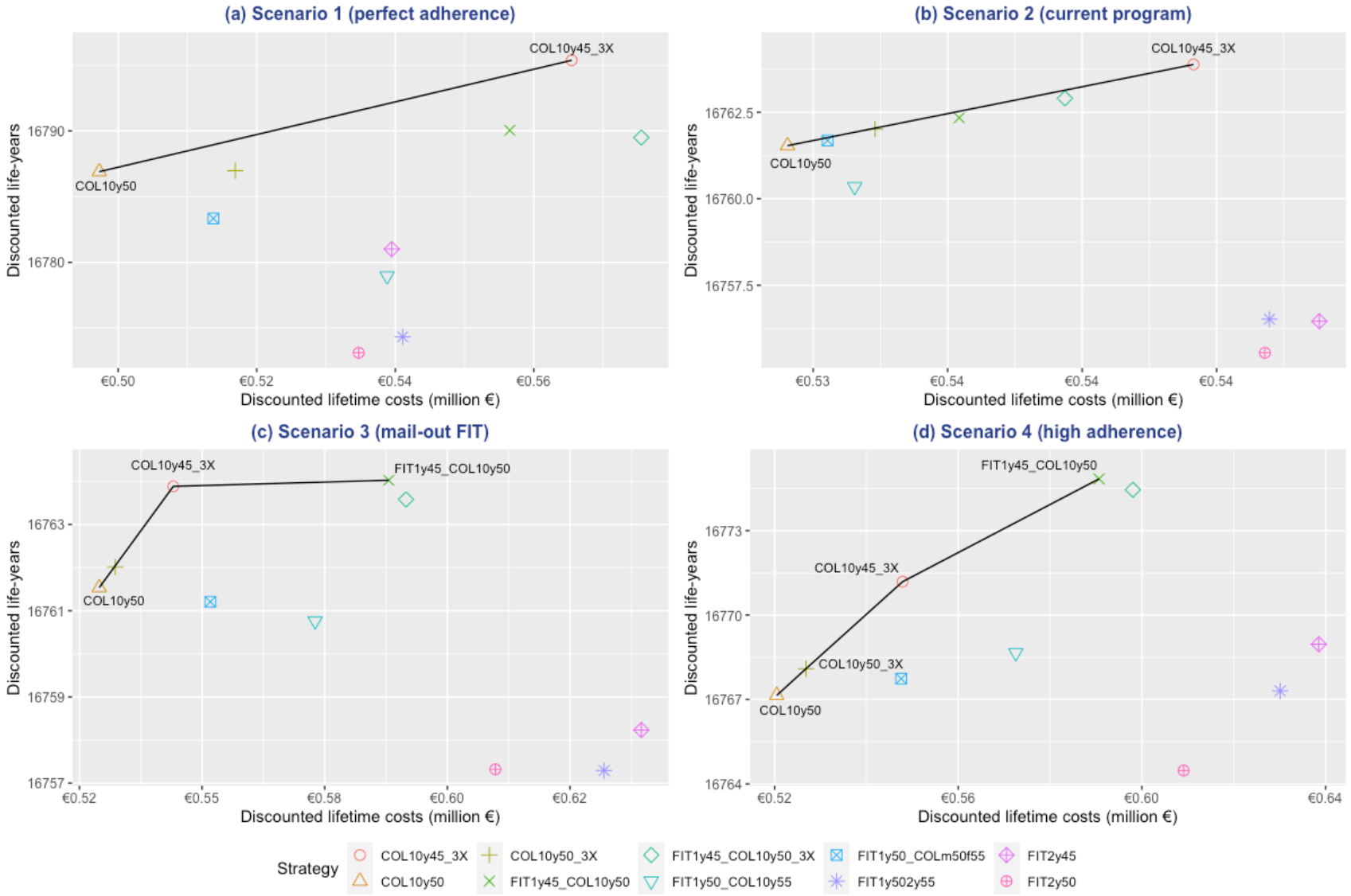
Note: -- = reference; Δ = difference; CER = cost-effectiveness ratio; CS = cost-saving; D = dominant; D'ed = dominated; ICER = incremental cost-effectiveness ratio; QALYG= quality-adjusted life-years gained.

(1) The costs and quality-adjusted life-years are expressed as per 1,000 40-year-olds were discounted with 0% annual rates. Costs were in 2021 Euro.

(2) CERs were calculated by comparing each strategy with no-screening strategy. ICERs were calculated by comparing each strategy with the next most effective one (the rows next to each other).

(3) The results are presented as: difference between means of the two strategies and difference between the upper value and lower values of 95% credible intervals of the two strategies.

(4) "Dominant" denotes a strategy being more effective and cost-saving; "Dominated" denotes a strategy being less effective and costing more.



Supplementary Figure 3. Efficiency frontier based on cost-effectiveness (life-years as benefits, 5% discount rate).
Texts shown in the graphs mark the dominant strategies in each scenario, which will be used to calculate incremental cost-effectiveness ratios (ICER).

Supplementary Table 12. Cost-effectiveness and incremental cost-effectiveness ratios (ICERs) between efficient strategies (life-years as benefits, 5% discount rate)

Strategy[1,2,3,4]	Compared to no screening			ICER		
	ΔCosts	LYG	CER	ΔCosts	LYG	ICER
Scenario 1 (perfect adherence)						
COL10y50	Cost-saving (CS-7,840))	45 (29-75)	Dominant (104-D)	--	--	--
COL10y45-3X	Cost-saving (CS-93,930)	53 (31-89)	Dominant (1,053-D)	68,195 (42,264-86,089)	8 (3-14)	8,046 (6,287-14,116)
Scenario 2 (current program)						
COL10y50	Cost-saving (CS-CS)	19 (13-37)	Dominant (D-D)	--	--	--
COL10y45-3X	Cost-saving (CS-CS)	22 (14-39)	Dominant (D-D)	15,092 (11213-23,185)	2 (2-2)	6,428 (6,909-15,114)
Scenario 3 (mail-out FIT)						
COL10y50	Cost-saving (CS-CS)	19 (13-37)	Dominant (D-D)	--	--	--
COL10y45-3X	Cost-saving (CS-CS)	22 (14-39)	Dominant (D-D)	15,092 (11213-23,185)	2 (2-2)	6,428 (6,909-15,114)
FIT1y45+COL10y50	Cost-saving (CS-31,248)	22 (12-40)	Dominant (777-D)	43,912 (44,481-46,310)	0(-2-2)	30,8421 (D'ed-26,046)
Scenario 4 (high adherence)						
COL10y50	Cost-saving (CS-CS)	25 (15-43)	Dominant (D-D)	--	--	--
COL10y45-3X	Cost-saving (CS-14,513)	29 (20-52)	Dominant (277-D)	27,414 (4,105-37,203)	4 (4-10)	6,785 (936-3,820)
FIT1y45+COL10y50	Cost-saving (CS-63,021)	32 (21-60)	Dominant (1,046-D)	42,780 (33,696-48,508)	4 (2-8)	11,711 (6,128-20,818)

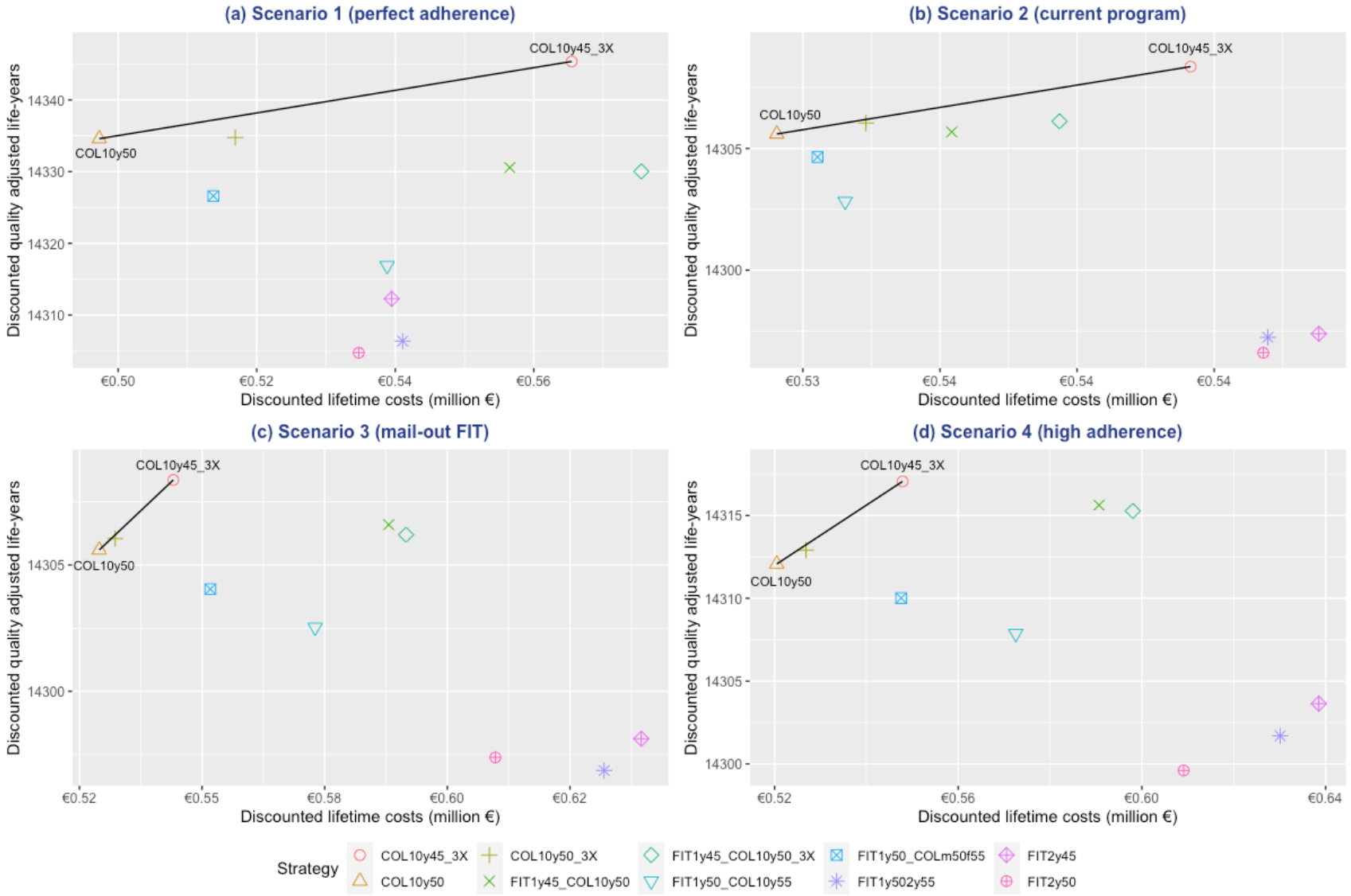
Note: -- = reference; Δ = difference; CER = cost-effectiveness ratio; CS = cost-saving; D = dominant; D'ed = dominated; ICER = incremental cost-effectiveness ratio; LYG = life-years gained.

(1) The costs and life-years are expressed as per 1,000 40-year-olds were discounted with 5% annual rates. Costs were in 2021 Euro.

(2) CERs were calculated by comparing each strategy with no-screening strategy. ICERs were calculated by comparing each strategy with the next most effective one (the rows next to each other).

(3) The results are presented as: difference between means of the two strategies and difference between the upper value and lower values of 95% credible intervals of the two strategies.

(4) "Dominant" denotes a strategy being more effective and cost-saving; "Dominated" denotes a strategy being less effective and costing more.



Supplementary Figure 4. Efficiency frontier based on cost-effectiveness (quality adjusted life-years as benefits, 5% discount rate).
Texts shown in the graphs mark the dominant strategies in each scenario, which will be used to calculate incremental cost-effectiveness ratios (ICER).

Supplementary Table 13. Cost-effectiveness and incremental cost-effectiveness ratios (ICERs) between efficient strategies (quality-adjusted life-years as benefits, 5% discount rate)

Strategy[1,2,3,4]	Compared to no screening			Per QALY		
	ΔCosts	QALYG	CER	ΔCosts	QALYG	ICER
Scenario 1 (perfect adherence)						
COL10y50	Cost-saving (CS-7,840))	44 (41-49)	Dominant (161-D)	--	--	--
COL10y45-3X	Cost-saving (CS-93,930)	55 (50-57)	Dominant (1,637-D)	68,195 (42,264-86,089)	11 (9-9)	6,333 (4,605-9,995)
Scenario 2 (current program)						
COL10y50	Cost-saving (CS-CS)	15 (14-16)	Dominant (D-D)	--	--	--
COL10y45-3X	Cost-saving (CS-CS)	18 (18-18)	Dominant (D-D)	15,092 (11213-23,185)	3 (2-4)	5,448 (5,375-6,187)
Scenario 3 (mail-out FIT)						
COL10y50	Cost-saving (CS-CS)	15 (14-16)	Dominant (D-D)	--	--	--
COL10y45-3X	Cost-saving (CS-CS)	18 (18-18)	Dominant (D-D)	15,092 (11213-23,185)	3 (2-4)	5,448 (5,375-6,187)
Scenario 4 (high adherence)						
COL10y50	Cost-saving (CS-CS)	22 (22-25)	Dominant (D-D)	--	--	--
COL10y45-3X	Cost-saving (CS-14,513)	27 (27-29)	Dominant (496-D)	27,414 (4,105-37,203)	5 (2-7)	5,496 (2,636-5,410)

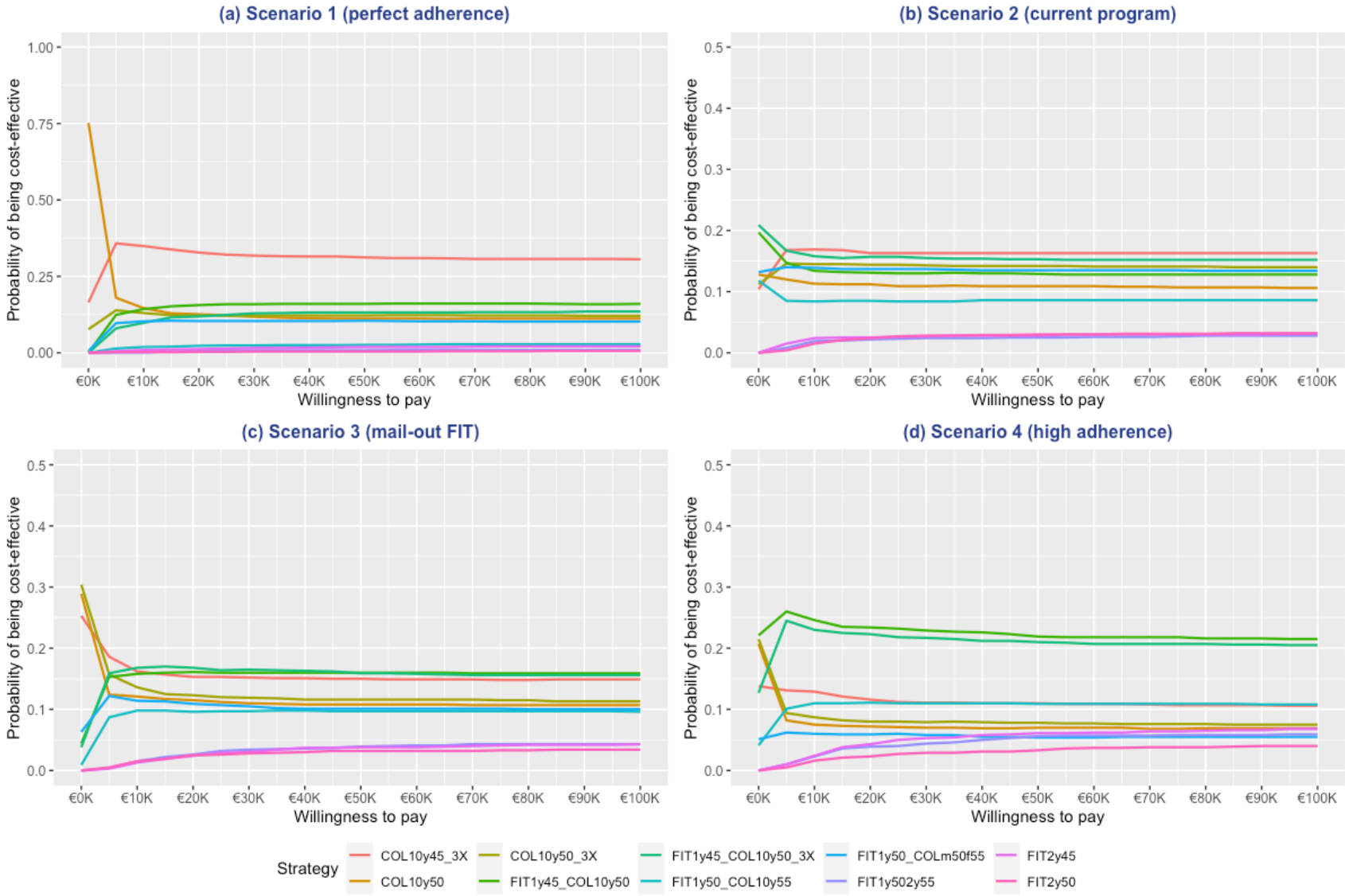
Note: -- = reference; Δ = difference; CER = cost-effectiveness ratio; CS = cost-saving; D = dominant; D'ed = dominated; ICER = incremental cost-effectiveness ratio; QALYG= quality-adjusted life-years gained.

(1) The costs and quality-adjusted life-years are expressed as per 1,000 40-year-olds were discounted with 5% annual rates. Costs were in 2021 Euro.

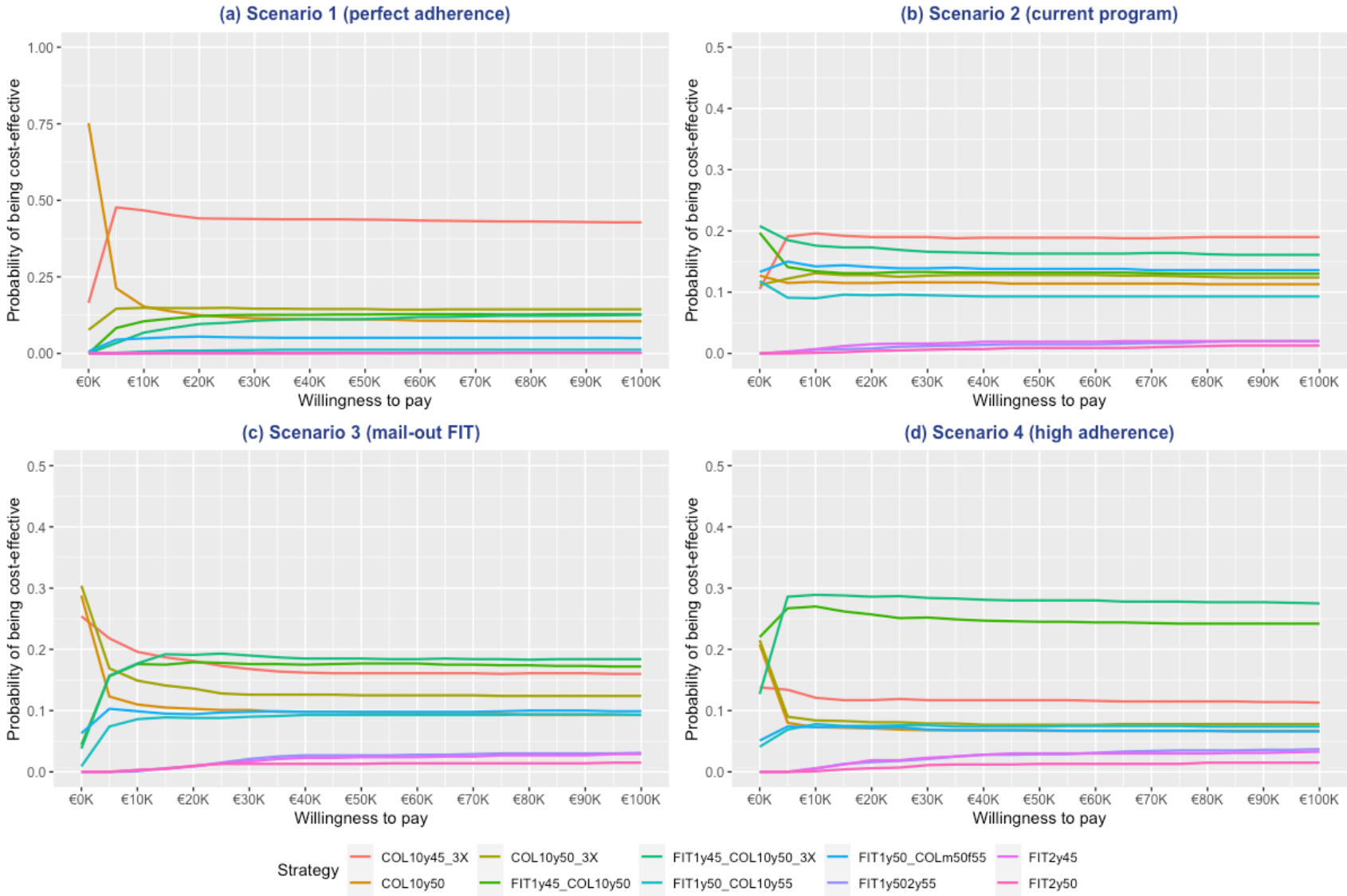
(2) CERs were calculated by comparing each strategy with no-screening strategy. ICERs were calculated by comparing each strategy with the next most effective one (the rows next to each other).

(3) The results are presented as: difference between means of the two strategies and difference between the upper value and lower values of 95% credible intervals of the two strategies.

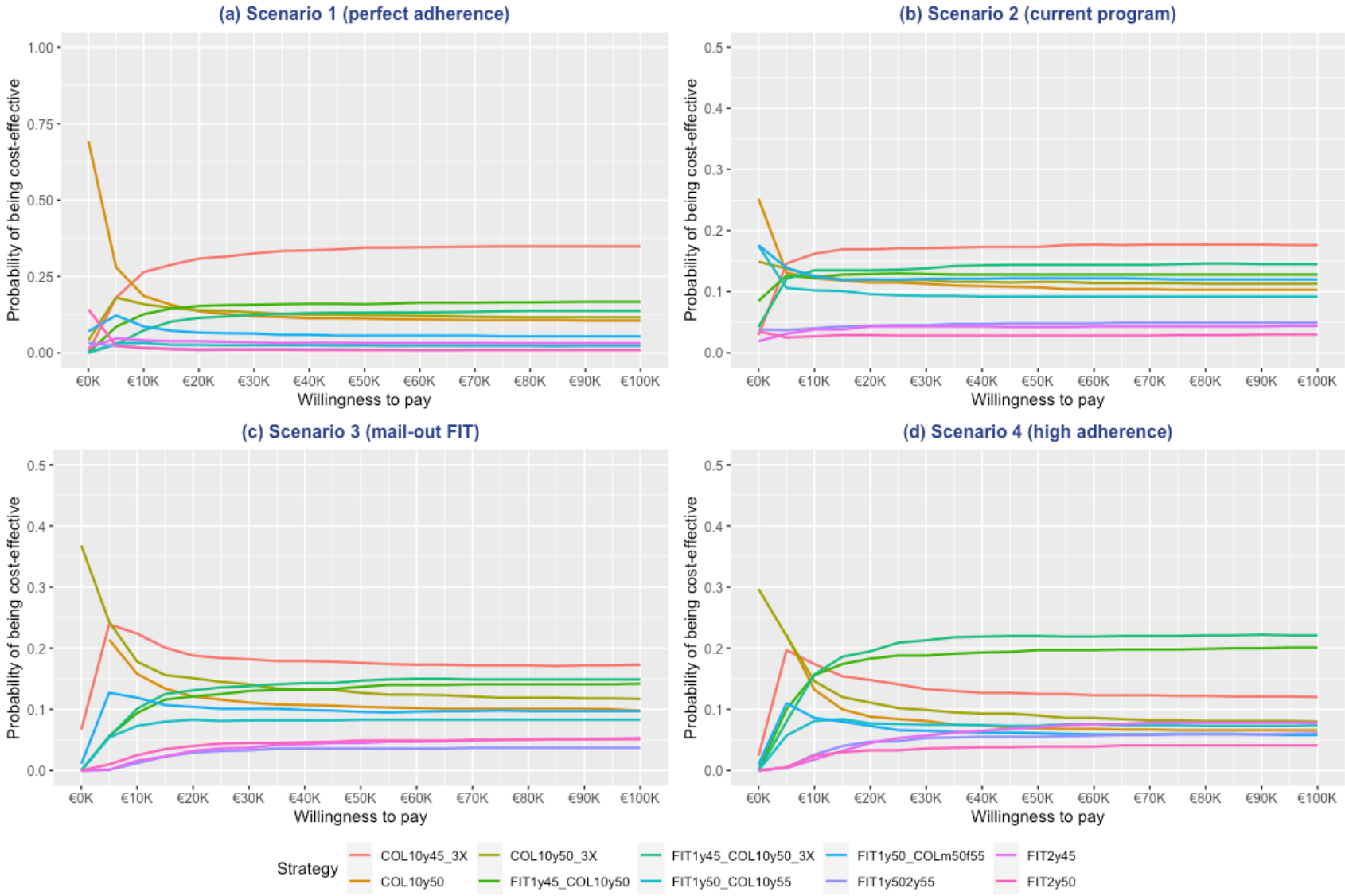
(4) "Dominant" denotes a strategy being more effective and cost-saving; "Dominated" denotes a strategy being less effective and costing more.



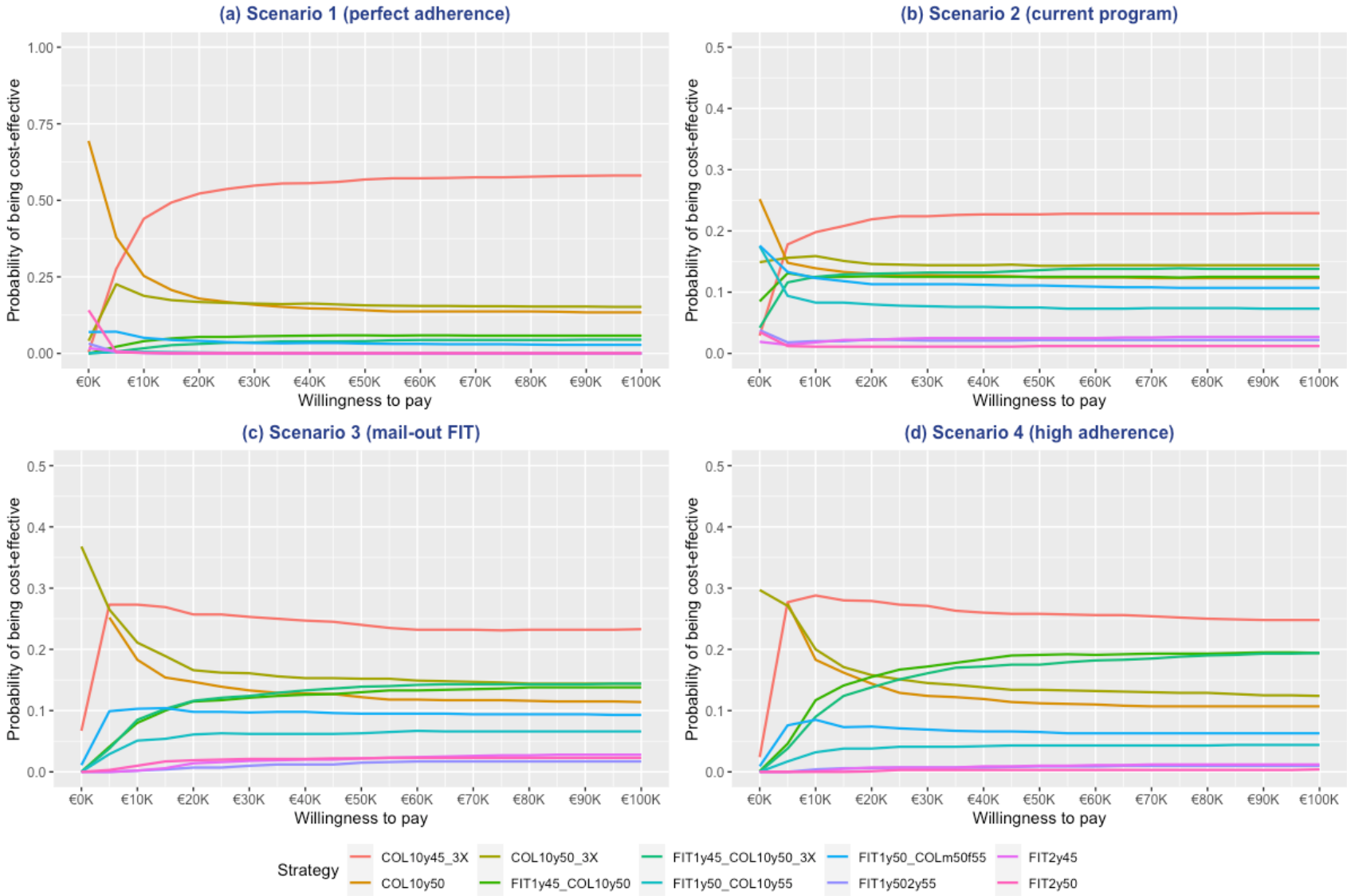
Supplementary Figure 5. Cost-effectiveness acceptance curve (life-years as benefits, 0% discount rate).
Each line denotes how likely each strategy could be the most cost-effective strategy under the given willingness-to-pay threshold.



Supplementary Figure 6. Cost-effectiveness acceptance curve (quality-adjusted life-years as benefits, 0% discount rate).
Each line denotes how likely each strategy could be the most cost-effective strategy under the given willingness-to-pay threshold.



Supplementary Figure 7. Cost-effectiveness acceptance curve (life-years as benefits, 5% discount rate).
Each line denotes how likely each strategy could be the most cost-effective strategy under the given willingness-to-pay threshold.



Supplementary Figure 8. Cost-effectiveness acceptance curve (quality-adjusted life-years as benefits, 5% discount rate).
Each line denotes how likely each strategy could be the most cost-effective strategy under the given willingness-to-pay threshold.

OWN CONTRIBUTION TO PUBLICATIONS

Below is the list of publications which stem from this PhD project (reverse chronologically):

1. Cheng, C.Y., Calderazzo, S., Schramm, C., Schlander, M., 2022. Modeling the natural history and screening effects for colorectal cancer using both adenoma and serrated neoplasia pathways: the development, calibration, and validation of a discrete event simulation model. Submitted to Medical Decision Making Policy & Practice
2. Cheng, C.Y., Datzmann, T., Hernandez, D., Schmitt, J., Schlander, M., 2021. Do certified cancer centers provide more cost-effective care? A health economic analysis of colon cancer care in Germany using administrative data. Int J Cancer. https://doi.org/10.1002/ijc.33728
3. Deibel, A., Deng, L., Cheng, C.Y., Schlander, M., Ran, T., Lang, B., Krupka, N., Beerenwinkel, N., Rogler, G., Wiest, R., Sonnenberg, A., Poleszczuk, J., Misselwitz, B., 2021. Evaluating key characteristics of ideal colorectal cancer screening modalities: the microsimulation approach. Gastrointest. Endosc. https://doi.org/10.1016/j.gie.2021.02.013
4. Ran, T., Cheng, C.Y., Misselwitz, B., Brenner, H., Ubels, J., Schlander, M., 2019. Cost-Effectiveness of Colorectal Cancer Screening Strategies—a Systematic Review. Clin Gastroenterol Hepatol 17, 1969–1981. https://doi.org/10.1016/j.cgh.2019.01.014
5. Schlander, M., Cheng, C.-Y., Ran, T., 2018. Gesundheitsökonomie der Krebsfrüherkennung in Deutschland: Welche Interventionen sind kosteneffektiv bei bevölkerungsweiter Umsetzung? Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 61, 1559–1568. https://doi.org/10.1007/s00103-018-2839-3

Part of the results from **Publication 1** are used in Chapters 2.1, 3.1 and 4.2.1, while **Publication 2** are used in Chapters 2.2, 2.3, 3.2, 3.3, and 4.1. I contributed to both publications by drafting the research proposals and conceptualization, study design, data analysis and interpretation, drafting and revising the manuscripts.

Part of the content in Chapter 1.2. and Chapter 1.3.1 are informed by **Publication 4 and 5**. In Publication 4, I was one the two reviewers to perform the full systematic review, including database searching, title, abstract and full-text review and selection, data extraction and analysis, drafting part of manuscript, critical review, and revision. In Publication 5, I contributed to the review of clinical evidence on CRC screening and the production of tables and figures with relevant information.

Conference presentations which stem from this PhD project (reverse chronologically):

1. Cheng, C.Y., Calderazzo, S., Schramm, C., Schlander, M. Modeling the long-term effectiveness of colonoscopy screening by a discrete event simulation model calibrated with German screening registry data, 2020 ESMO Virtual Conference, Online
2. Cheng, C.Y., Calderazzo, S., Schramm, C., Schlander, M. Modeling the duration of protective effects and resource use of colonoscopy screening by a discrete event simulation model calibrated with German screening registry data. 2020 Virtual ISPOR Europe, Online.
3. Cheng, C.Y., Ran, T., Schlander, M. A systematic review update on cost-effectiveness of colorectal cancer screening: identification of an optimal strategy in Europe. Poster and Silent Theater Oral Presentation, 2018 NCRI Cancer Conference, Glasgow, UK.

CURRICULUM VITAE

PERSONAL DATA

Name:	Chih-Yuan Cheng
Date of Birth:	26.12.1987
Place of Birth:	Tainan, Taiwan

SCHOOL EDUCATION

2000 – 2006	Li-Ming High School, Tainan, Taiwan
30.06.2006	**High School Diploma**

UNIVERSITY EDUCATION

2006-2013	**Medicine** National Cheng Kung University, Tainan, Taiwan
30.06.2013	**Degree: Doctor of Medicine (MD)** Grade: GPA 3.95/4
30.09.2013	**Certified as a Physician in Taiwan** Passed the National Qualifying Exam for Medical Doctors, Taiwan
2015-2016	**MSc Health Policy, Planning and Financing** London School of Economics and Political Science, London, UK London School of Hygiene and Tropical Medicine, London, UK
10.11.2016	**Degree: MSc in Health Policy, Planning and Financing** Grade: Distinction

PROFESSIONAL EXPERIENCE

2012-2015	**Physician in training** National Cheng Kung University Hospital, Tainan, Taiwan Duke University Hospital, Durham, NC, USA National Taiwan University Hospital, Taipei, Taiwan Taiwan Medical Mission, Koudougou, Burkina Faso
2016-2018	**Research Associate in Health Economics** Aquarius Population Health, London, UK
2020	**Intern in Inhouse Consulting** Merck KGaA, Darmstadt, Germany
2021-	**Postdoctoral Fellow, Health Economics Modeling** Johnson & Johnson, Beerse, Belgium

ACKNOWLEDGEMENTS

First of all, I give my big thank you to my supervisor *Prof. Dr. med. Michael Schlander.*

Michael: Thank you for accepting me to join the team and work on this interesting and important topic. It was a great pleasure to have this challenging yet fulfilling academic journey under your supervision. I am always very grateful for your support throughout, be it the obstacles faced during registration, the challenges in looking for collaborations, and my long-term professional development plan.

I also especially thank *Dr. Silvia Calderazzo* for her expertise in Bayesian statistics, the time and efforts in supporting this project, and the assistance in coding in R program. Without her professional input and tireless guidance, the modeling project would not have been possible.

I thank *Dr. Tao Ran* for the collaboration in the systematic review project for the cost-effectiveness of CRC, which paved the foundation for this modeling project.

I thank *Prof. Dr. med. Peter Kolominsky-Rabas* and *Priv.-Doz. Dr. sc. hum. Michael Hoffmeister* for being the advisors in the thesis committee and their valuable comments to improve the doctoral project.

I thank *Prof. Dr. med. Benjamin Misselwitz* for the initial modeling collaboration using the microsimulation model in their group, the research stay in Zurich, and the invitation to be a co-author in their publications, which helped me better understand modeling.

I thank *Prof. Dr. med. Jochen Schmitt* for providing the access to AOK database and *Dr. Thomas Datzmann* for processing the data from AOK database for us to realize the estimation of CRC treatment costs.

I thank all the *co-authors* of my publications for the input, support and feedback. I have learned a lot from the process to grow as a researcher.

I thank my *colleagues* that accompanied me through this journey. They both supported me professionally with inspiring idea exchange and personally with care and laughter. The years in Heidelberg would not have been the same without their company.

Finally, I thank my parents *Shu-Ing* and *His-Kun* as well as my partner *Tim* with all my heart for their love, patience and huge faith in myself. I also thank all my *friends,* local or afar*,* that made this journey more enjoyable.