

Acknowledging patient heterogeneity in colorectal cancer screening: An example from Norway

MATHYN A.M. VERVAART¹

EMILY A. BURGER^{1,2}

ELINE AAS^{1,*}

¹Institute of Health and Society, Department of Health Management and Health Economics, University of Oslo, Oslo, Norway

²Harvard T.H. Chan School of Public Health, Center for Health Decision Science, Boston, MA, USA

Abstract: Different sources of patient heterogeneity or personal characteristics may contribute to differential cost-effectiveness profiles of national screening programs for colorectal cancer (CRC). To motivate the use of subgroup analyses when individual level data are unavailable, we provide a stylized example of the potential economic value of capturing patient heterogeneity in CRC screening. We developed a Markov model to capture the impacts of patient heterogeneity on the cost-effectiveness of CRC screening involving once-only sigmoidoscopy compared to no screening. We simulated cohorts of Norwegian men, women, and six comorbidity subgroups that differentially influenced the relative treatment effect, the risks of developing CRC, dying from CRC, dying from background mortality or screening-related adverse events and baseline quality of life. We calculated the discounted (4%) incremental cost-effectiveness ratio (ICER), defined as the cost per quality-adjusted life year (QALY) gained, and the net monetary benefit (NMB) gained by stratification, from a societal perspective. Screening in men was cost-effective at any threshold value, while screening in women only provides good value for money from threshold values of €50,000 per QALY gained and above. Comorbidities unrelated to CRC development yielded generally less attractive cost-effectiveness ratios (i.e., increased the ICER), while related comorbidities improved the cost-effectiveness profiles of screening for CRC. A stratified policy that accounts for different screening outcomes between men and women could potentially improve the value of screening by €5.8 million annually. Accounting for patient heterogeneity in CRC screening will likely improve the value of screening strategies, as a single screening approach for the entire population can result in inefficient use of resources.

JEL classification: D61, D63, I1

Key words: patient heterogeneity, stratified analysis, cost-effectiveness, screening for colorectal cancer

* *Correspondence to:* Eline Aas, University of Oslo, Institute of Health and Society, Department of Health Management and Health Economics, Postboks 1089, Blindern, 0317 Oslo, Norway. E-mail: eline.aas@medisin.uio.no.

1 Introduction

Evaluations of the health and economic trade-offs of alternative screening programs typically use a population-based approach, yielding health policy recommendations and reimbursement decisions based on outcomes for the average population. National screening guidelines, which can vary by screening test, screening frequency, and triage algorithm, often identify a target population for whom screening is most appropriate. However, these recommendations can mask important sources of patient heterogeneity (sex, age, comorbidity status), often because economic evaluations of important sub-groups for whom screening benefits may differ are not undertaken (Bala & Zarkin, 2004; Grutters et al., 2013). Recommending patients to attend screening for whom it is not effective (or even harmful), or withholding screening from patients in the population who can benefit is potentially unethical and is an inefficient use of healthcare resources. In order to motivate more wide-spread use of sub-group analyses when individual-level data are unavailable, the objective of this paper was to provide a stylized example of the potential economic value of capturing patient heterogeneity in colorectal cancer (CRC) screening using the most up-to-date sources of heterogeneity from a recent randomized controlled trial and literature reviews. Underscoring the role of patient heterogeneity is particularly important for Norway, which is currently outlining a national CRC screening program for all men and women aged 55+ years, irrespective of potentially important patient heterogeneity (Cancer Registry of Norway, 2017).

CRC is the 4th most prevalent cause of cancer mortality in the world (Lozano et al., 2012), and an increasing number of Western countries have implemented (or are considering implementing) organized population-based screening programs for the prevention or early detection of CRC (Basu et al., 2018; Schreuders et al., 2015). Several factors are known to increase the risk of developing and dying from CRC, such as diabetes and obesity, lifestyle-related factors (e.g. smoking and heavy alcohol use), and familial history and genetic predisposition (American Cancer Society, 2018). In addition, the risk of screening-related adverse events may be different across patient subgroups (Warren, 2009). Similarly, these factors may also increase the monetary costs associated with screening, treatment, and complications, which together, will influence the value (i.e., cost-effectiveness) of a generic screening program. In reviewing recent economic evaluations of screening for CRC, studies explored common sources of patient heterogeneity such as age and sex, and few have explored such factors as generic comorbidity status (including diabetes) and race on the value of CRC screening (Dinh et al., 2012; Lansdorp-Vogelaar et al., 2014; Lansdorp-Vogelaar et al., 2009; van Hees et al., 2015). With the exception of diabetes (Dinh et al., 2012), studies have not provided more concrete examples of patient subgroups for which screening recommendations may differ. Even so, and despite several studies identifying important differences between sub-groups (e.g., increasing harms due to overdiagnosis with increasing generic comorbidity index, reduced benefits of screening with decreasing life expectancy (Lansdorp-Vogelaar et al., 2014; van Hees et al., 2015)), national guidelines for CRC do not account for differences in patient heterogeneity, apart from age (Basu et al., 2018; Schreuders et al., 2015). To our knowledge, no studies have assessed the value (and opportunity costs) of stratifying CRC screening guidelines for specific sub-groups, which may resonate with stakeholders considering the trade-offs of implementing a more generic CRC screening program. Similarly, no studies have included evidence from a forthcoming randomized trial suggesting that CRC screening using sigmoidoscopy has little or no effect on women (Holme et al., 2018).

2 Methods

2.1 Stylized example for colorectal cancer screening

We developed a decision-analytic Markov model that simulates cohorts of men and women aged 60 years in order to provide estimates the cost-effectiveness of screening for CRC involving once-only sigmoidoscopy compared to no screening for alternative sub-groups. Recent RCTs have shown that sigmoidoscopy reduces the future incidence of CRC by identifying and removing adenomas before they have an opportunity to progress to invasive cancer (Holme et al., 2018; Lin et al., 2016). The role of sigmoidoscopy to identify early stage CRC was not considered in this stylized analysis.

2.2 Stratification

In a systematic review, Grutters et al. (2013) categorized patient heterogeneity into three sources that potentially affect parameter values in a decision model: demographics (e.g., age, sex and socio-economic status), preferences (e.g., attitude, beliefs and risk tolerance) and clinical characteristics (e.g. disease severity, disease history and genetic profile). Although differences in treatment effects are one important aspect of heterogeneity, the cost-effectiveness profiles of interventions can also be affected through differential parameter values for baseline risk, health state utility and resource utilization (Grutters et al., 2013). For example, Braithwaite (2011) defined payoff time as the time until a guideline's cumulative benefits first exceed its cumulative incremental harms. When life expectancy of a patient not following the guideline is shorter than payoff time, the guideline should not be followed. This payoff time is especially relevant in the presence of severe comorbidities that affect life expectancy (Braithwaite, 2011). Therefore, factors 'unrelated' to the development of CRC (e.g., background life expectancy, and quality of life) may also result in differing cost-effectiveness profiles.

We reviewed the literature to derive comorbidity-specific risks of CRC incidence, screening-related adverse events, and overall- and cancer-specific mortality (Botteri et al., 2008; Carter et al., 2015; Cho et al., 2013; De Bruijn et al., 2013; Flegal et al., 2013; Mills et al., 2013; Moghaddam et al., 2007; Phipps et al., 2011; Reichle et al., 2015; Warren, 2009), and have been summarized in Table 1. Patient heterogeneity was represented in our analysis by two types of comorbidity that affect baseline risk and quality of life, i.e., comorbidities related and unrelated to CRC development and screening performance. 'Related comorbidities' included diabetes mellitus, obesity and smoking, as it is well established that these factors influence baseline quality of life, the risk of developing CRC, dying from CRC and dying from background mortality. 'Unrelated comorbidities' influenced baseline quality of life and the risk of dying from background mortality and were selected from the literature to explore varying degrees of severity and burden in the general population. Unrelated comorbidities included chronic obstructive pulmonary disease (COPD), dementia, and chronic renal failure. Both diabetes mellitus and COPD additionally increased the risk of screening-related adverse events. Furthermore, we included sex-based subgroups, as recently published results from a randomized trial suggested differences between men and women in the relative effect of sigmoidoscopy screening, screening resource use and baseline risk (Holme et al., 2018).

We used the stratified analysis (SA) framework, introduced by Coyle et al. (2003), to identify the optimal treatment for each subgroup in the absence of individual-level data and to compare the cost-effectiveness between subgroups. Within the SA framework, we compared two approaches to patient management. The 'population-based approach', in which the optimal decision assumes all patients receive the same treatment, is compared

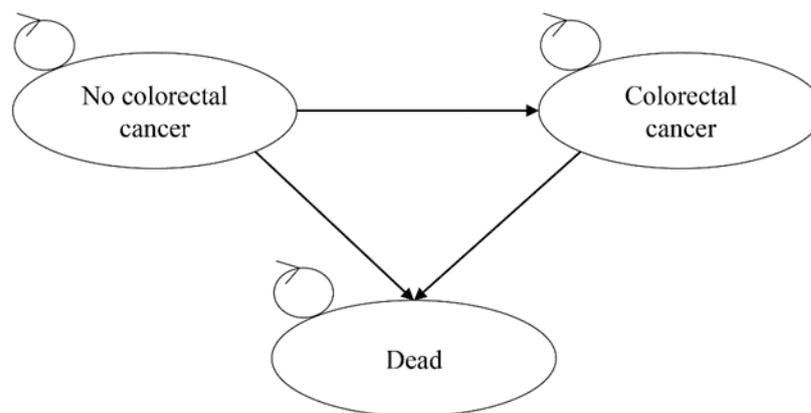
with a ‘stratified approach’, which assumes the total eligible population is stratified into relevant patient subgroups and the optimal decision is determined for each respective subgroup. The authors define this concept as the ‘limited use criteria’ (LUC) policy, which is intended to improve efficiency in healthcare by restricting treatment to subgroups for whom it provides good value for money. By calculating the net monetary benefit (NMB) gained with stratification (i.e., dividing the population into subgroups of patients) as inherent in LUC, the efficiency gains for considering patient heterogeneity can be quantified, allowing decision makers to examine the value of subgroup-specific policies. Quantification of the benefit of using patient heterogeneity to identify the optimal treatment for each subgroup involves calculating the cost-effectiveness of treatment in terms of the incremental net monetary benefit (INMB), conditioned on a specific monetary threshold value (i.e., willingness-to-pay threshold) assigned to an additional health effect, for each subgroup. Potential for efficiency gains exist when optimal treatment decisions for patient subgroups are different from an average population-based treatment decision (Coyle et al., 2003).

Model outcomes included monetary costs (2017 Euros) and QALYs for a cohort of a 1,000 screen-eligible individuals. For each of the eight subgroups, we calculated the incremental cost-effectiveness ratio (ICER), defined as the cost per quality-adjusted life year (QALY) gained. We adopted a societal perspective and discounted costs and health effects at 4% annually, consistent with Norwegian guidelines (Norwegian Medicines Agency, 2018). We calculated the NMB gained assuming stratification to quantify the potential efficiency gains for each patient subgroup compared with the general population.

3 Model structure

A cohort of Norwegian men or women enter the model at age 60 years without colorectal cancer (Figure 1). Each year, the cohort faces risks of developing and dying from CRC, dying due to other causes (i.e., background mortality) or remaining CRC-free. Screening, involving once-only sigmoidoscopy occurs at age 60 years (i.e., model start), followed by a colonoscopy for the proportion of screened individuals considered at risk of developing CRC. Adverse events associated with the sigmoidoscopy and colonoscopy procedures included intestinal bleeding and bowel perforation, which could result in hospitalization or death.

Figure 1: Model structure.



3.1 Model input parameters

Model parameter inputs included Norwegian national epidemiologic and economic data augmented with data obtained from the literature review as described earlier (Table 1).

Age- and sex-specific all-cause mortality was estimated from 2017 Norwegian life tables (Statistics Norway, 2018a). Age- and sex-specific CRC incidence rates and sex-specific 15-year relative survival was estimated using 2012-2016 data from the Norwegian Cancer Registry (Cancer Registry of Norway). We used tunnel states to capture the time-dependency of CRC survival and assumed CRC patients faced no excess CRC mortality after 15 years.

The comorbidity-specific risks of colorectal cancer incidence, screening-related adverse events, and overall- and cancer-specific mortality, identified in the previously-described literature review, were applied to the baseline event rates (Botteri et al., 2008; Carter et al., 2015; Cho et al., 2013; De Bruijn et al., 2013; Flegal et al., 2013; Mills et al., 2013; Moghaddam et al., 2007; Phipps et al., 2011; Reichle et al., 2015; Warren, 2009). To control for overestimation of CRC risks, increased event rates in subgroups were normalized by using age- and sex-specific Norwegian prevalence data (Norwegian Institute of Public Health, 2016a, 2016b; Norwegian Renal Registry, 2016; Solli, 2013; Statistics Norway, 2016; 2018a; 2018b; Waatevik et al., 2013).

The clinical effect of screening on CRC incidence, stratified for men and women, was based on a large (n= 100,210) randomized controlled trial of once-only screening involving sigmoidoscopy conducted in a sample of the Norwegian general population aged 55-64 years with a median follow-up of 14.8 years (Holme et al., 2018). The hazard ratio for incident CRC was applied in the model for the first 14 years following baseline screening at age 60 years. From years 15 to 19 after screening, incident CRC was assumed to linearly return to the age-specific baseline risk. As no comorbidity-specific data on the clinical effect of screening were available, we assumed a constant relative screening effect across comorbidity subgroups. We further assumed that the probability of a follow-up colonoscopy in screened individuals was proportional to CRC risk across the subgroups.

The probability of follow-up colonoscopy for men and women was based on the same Norwegian trial (Holme et al., 2018), and the baseline probability of screening-related adverse events was derived from a previous economic evaluation of colorectal cancer screening (Sharp et al., 2013).

As Norwegian data on health-related quality of life (HRQoL) do not exist, we derived the age- and sex-specific HRQoL values for the general population from the Swedish study by Burström et al. (2001), as recommended in the Norwegian guidelines (Norwegian Directorate of Health, 2018). We used Dutch studies to estimate the baseline values for comorbidity-specific HRQoL, elicited using the EQ-5D questionnaire (De Wit et al., 2000; Hakim et al., 2002; Redekop et al., 2002; Rutten-van Molken et al., 2006; Vogl et al., 2012; Wolfs et al., 2007). We applied an age-adjustment based on the general population values, in accordance with the Norwegian guidelines. The quality of life decrement for incident colorectal cancer was derived from the Global Burden of Disease study (Salomon et al., 2012) and subtracted from the baseline comorbidity-specific quality of life value. Similar to cancer survival, we assumed quality of life returned to its baseline pre-cancer value after 15 years.

Table 1: Values for key model parameters used in the analysis. Costs in 2017 Euros.

Model parameter	Value ^a	References
CRC (baseline)		(Cancer Registry of Norway, 2017)
CRC incidence, men (age 60-99)	0.001 – 0.006	(Cancer Registry of Norway, 2017)
CRC incidence, women (age 60-99)	0.001 – 0.004	(Cancer Registry of Norway, 2017)
CRC cumulative relative survival (year 1-15), men	0.836 – 0.540	(Cancer Registry of Norway, 2017)
CRC cumulative relative survival (year 1-15), women	0.834 – 0.584	(Cancer Registry of Norway, 2017)
Screening characteristics (baseline)		
RR for CRC incidence for sigmoidoscopy, men	0.66 (0.57 - 0.78)	(Holme et al., 2018)
RR for CRC incidence for sigmoidoscopy, women	0.92 (0.79 - 1.07)	(Holme et al., 2018)
Participation rate in, men	0.614	(Holme et al., 2018)
Participation rate in, women	0.647	(Holme et al., 2018)
Baseline proportion attending colonoscopy, men	0.229	(Holme et al., 2018)
Baseline proportion attending colonoscopy, women	0.162	(Holme et al., 2018)
Adverse screening events (baseline)		
Probability of perforation during sigmoidoscopy	0.00002	(Sharp et al., 2013)
Probability of death after perforation during sigmoidoscopy	0.065	(Sharp et al., 2013)
Probability of major bleeding during sigmoidoscopy	0.00003	(Sharp et al., 2013)
Probability of perforation during colonoscopy	0.002	(Sharp et al., 2013)
Probability of death after perforation during colonoscopy	0.052	(Sharp et al., 2013)
Probability of major bleeding during colonoscopy	0.004	(Sharp et al., 2013)
Risk for comorbidities		
RR for CRC incidence for obesity	1.19 (1.11 - 1.29)	(Moghaddam et al., 2007)
RR for conditional CRC mortality for obesity	3.17 (1.29 - 7.81)	(Reichle et al., 2015)
RR for overall mortality for obesity	1.18 (1.12 - 1.25)	(Flegal et al., 2013)
RR for CRC incidence for diabetes	1.26 (1.14 - 1.40)	(De Bruijn et al., 2013)
RR for conditional CRC mortality for diabetes	1.12 (1.01 - 1.24)	(Mills et al., 2013)
RR for overall mortality for diabetes	1.45 (1.43 - 1.47)	(Cho et al., 2013)
RR for adverse screening events for diabetes	1.262	(Warren, 2009)
RR for CRC incidence for smoking	1.18 (1.11 - 1.25)	(Botteri et al., 2008)
RR for conditional CRC mortality for smoking	1.30 (1.09 - 1.74)	(Phipps et al., 2011)
RR for overall mortality for smoking	2.80 (2.70 - 2.90)	(Carter et al., 2015)
RR for overall mortality for COPD	1.76 (1.74 - 1.79)	(Cho et al., 2013)
RR for adverse screening events for COPD	1.508	(Warren, 2009)
RR for overall mortality for chronic renal failure	2.30 (2.25 - 2.36)	(Cho et al., 2013)
RR for overall mortality for dementia	2.35 (2.31 - 2.40)	(Cho et al., 2013)
Quality of life		
Baseline quality of life for the general population	0.80 – 0.72	(Burstrom et al., 2001)
Baseline quality of life for patients with obesity	0.726	(Hakim et al., 2002)

Table 1 continued: Values for key model parameters used in the analysis. Costs in 2017 Euros.

Model parameter	Value ^a	References
Baseline quality of life for patients with diabetes	0.780	(Redekop et al., 2002)
Baseline quality of life for smokers	0.767	(Vogl et al., 2012)
Baseline quality of life for patients with COPD	0.760	(Rutten-van Molken et al., 2006)
Baseline quality of life for patients with chronic renal failure	0.690	(De Wit et al., 2000)
Baseline quality of life for patients with dementia	0.540	(Wolfs et al., 2007)
Quality of life decrement for CRC	0.294	(Salomon et al., 2012)
Unit costs (2017 €)		
Invitation incl. reminder for sigmoidoscopy	25	(Aas, 2015)
Invitation incl. reminder for colonoscopy	25	(Aas, 2015)
Sigmoidoscopy	337	(Aas, 2015)
Colonoscopy	495	(Aas, 2015)
Pathology for sigmoidoscopy	6	(Aas, 2015)
Pathology for colonoscopy	234	(Aas, 2015)
Treating bowel perforation	3 440	(Norwegian Directorate of Health, 2016)
Hospital admittance for bleeding	790	(Lonne et al., 2015)
CRC treatment	39 268	(Joranger et al., 2015)
Round-trip to the medical center	35	(Aas, 2015)

COPD: chronic obstructive pulmonary disease, CRC: colorectal cancer, RR: risk ratio.

^aFor some parameters, the confidence intervals/standard errors were either not available or difficult to present due to age-specific values.

Costs of the screening procedure, adverse events and average lifetime costs of colorectal cancer treatment were based on Norwegian cost studies (Aas, 2015; Joranger et al., 2015; Lonne et al., 2015; Norwegian Directorate of Health, 2016). In the absence of data we assumed unit costs of the screening procedure and per-patient average lifetime costs of cancer treatment were equal across subgroups, and adjusted to the participation rate in the screening trial. Per-patient average lifetime costs of CRC treatment were applied to all incident CRC cases. Costs of treating comorbidities were not included, as these were not directly affected by the screening procedure.

Direct non-healthcare and indirect costs included travel costs and productivity costs related to screening and CRC treatment. We applied the friction cost method to calculate productivity costs (Koopmanschap & Rutten, 1996). The friction period for Norway has previously been estimated to be 68 days, based on data from the Norwegian Labour and Welfare Administration on open and filled vacancies in 2010 and 2011 (Norwegian Labour and Welfare Administration, 2015). Absence from work for the screening procedure has previously been estimated at 2.5 hours for sigmoidoscopy and 5 hours for colonoscopy (Aas, 2015).

3.2 Sensitivity analysis

We performed a number of one-way and multi-way sensitivity analyses to evaluate the sensitivity of model outcomes for alternative parameter values (Table 3). We obtained the lower and upper values for the risk parameters from the corresponding 95% confidence

intervals as reported in the literature. For the lifetime cost of CRC treatment, we used lower and upper values corresponding to a 40% deviation from the mean value.

4 Results

4.1 Stratified cost-effectiveness analysis

For the average general population, once-only sigmoidoscopy screening at age 60 years to reduce the future incidence of CRC resulted in €8,241 per QALY gained (Table 2); however, the cost-effectiveness profiles varied considerably across the alternative subgroups. Screening men using sigmoidoscopy was considered cost-effective at any positive willingness-to-pay threshold value, while screening in women only provided good value for money for threshold values of €50,000 per QALY gained and above. With the exception of smokers, screening subgroups with comorbidities considered related to CRC yielded more attractive ICERs compared to the average general population. In contrast, the subgroups with comorbidities considered unrelated to CRC generally resulted in less favorable ICERs compared to the average population. For example, at a threshold value of €20,000 per QALY gained, screening using once-only sigmoidoscopy provided good value for money in the average general population, but not in subgroups of women and individuals with dementia.

Table 2: Discounted incremental costs, quality-adjusted life years (QALY), incremental cost-effectiveness ratios (ICER) and incremental net monetary benefit for screening for CRC by once-only sigmoidoscopy at age 60 years compared to no screening. Cohorts of 1,000 individuals, societal perspective. Costs in 2017 Euros.

	Incremental QALYs	Incremental cost (€)	ICER (€)	Incremental net monetary benefit (€)		
				λ =€20,000	λ =€40,000	λ =€60,000
General population	21.6	177,657	8,241	253,499	684,655	1,115,810
<i>Sex</i>						
Men	36.5	40,050	1,098	689,226	1,418,503	2,147,779
Women	6.4	317,186	49,225	-188,314	-59,441	69,431
<i>Comorbidities related to CRC</i>						
Obesity	42.9	155,261	3,619	702,718	1,560,697	2,418,676
Diabetes	26.8	136,153	5,076	400,341	936,834	1,473,328
Smoking	17.8	175,218	9,824	181,495	538,207	89,4920
<i>Comorbidities unrelated to CRC</i>						
COPD	18.8	132,068	7,040	243,144	618,355	993,567
Chronic renal failure	11.3	158,609	14,095	66,440	291,490	516,539
Dementia	7.5	212,098	28,139	-61,346	89,407	240,160

λ = monetary value of a quality-adjusted life year gained. COPD: chronic obstructive pulmonary disease, CRC: colorectal cancer.

4.2 Gain from stratification

Dependent on the threshold willingness-to-pay value, we found that the optimal decision to screen using once-only sigmoidoscopy was different for the respective subgroups than for the average population, yielding positive gains associated with stratified recommendations. For a cohort of 1,000 women, the NMB gained by stratification was €188,314 at a threshold value of €20,000 per QALY gained (Table 2). When we scaled to the Norwegian population of 31,007 60-year-old women in 2017, the value of a stratified policy based on a sex-specific policy could amount to approximately €5.8 million annually. The NMB gained by stratification was positive at €61,346 for a cohort of 1,000 individuals with dementia (Table 2), which approximately equals the potential annual value of a stratified policy based on this characteristic in the Norwegian 60-year old population.

4.3 Sensitivity

In sensitivity analysis, we found that the sex-specific relative screening effect had the largest impact on the model outcomes, as the upper limit for the relative effect in women was above 1 (i.e., a higher CRC incidence in the screening arm) (Table 3). Alternative values for the baseline risk across the comorbidity subgroups and cost of CRC treatment had a limited impact on the ICER, and would generally not change the optimal treatment decision given a threshold value of €20,000 per QALY gained. An exception was the dementia subgroup when the higher bound for the cost of CRC treatment was used; the ICER dropped to just below €20,000 per QALY gained, which would result in a positive screening recommendation at the given threshold. The results of the sensitivity analysis for the comorbidity subgroups not presented in Table 3 were very similar.

Table 3: Discounted incremental costs, quality-adjusted life years (QALY) and incremental cost-effectiveness ratios (ICER) associated with the one-way and multi-way analyses for screening for CRC involving once-only sigmoidoscopy at age 60 years compared to no screening for the comorbidity subgroups related and unrelated to CRC with the highest expected benefit of screening within their category, i.e., obesity and COPD.

	Value	Incremental QALYs	Incremental cost (€)	ICER (€)
<i>Men</i>				
Relative screening effect (lower)	0.57	46.4	-59,778	Dominant
Relative screening effect (upper)	0.78	23.3	172,639	7,403
<i>Women</i>				
Relative screening effect (lower)	0.79	17.8	207,267	11,673
Relative screening effect (upper)	1.07	-6.6	443,500	Dominated
<i>Obesity</i>				
Relative screening effect (lower)	0.57	79.0	-57,919	Dominant
Relative screening effect (upper)	1.07	-12.9	498,607	Dominated
Relative risk for CRC incidence (lower)	1.11	40.8	159,616	3,914
Relative risk for CRC incidence (upper)	1.29	45.4	150,512	3,312
Relative risk for CRC mortality (lower)	1.29	27.7	155,261	5,595
Relative risk for CRC mortality (upper)	7.81	49.9	155,261	3,114
Relative risk for other mortality (lower)	1.12	43.6	154,521	3,546
Relative risk for other mortality (upper)	1.25	42.2	156,105	3,703
Relative risks for CRC incidence, CRC mortality, other mortality (lower)	1.11, 1.29, 1.12	27.0	158,902	5,889
Relative risks for CRC incidence, CRC mortality, other mortality (upper)	1.29, 7.81, 1.25	52.0	151,391	2,911
Relative risks for CRC incidence, CRC mortality, other mortality (lower), CRC treatment cost	1.11, 1.29, 1.12, €54,975	27.0	64,328	2,384
Relative risks for CRC incidence, CRC mortality, other mortality (upper), CRC treatment cost	1.29, 7.81, 1.25, €23,561	52.0	255,904	4,921
<i>COPD</i>				
Relative screening effect (lower)	0.57	29.2	-18,157	Dominant
Relative screening effect (upper)	1.07	-4.9	475,122	Dominated
Relative risk for other mortality (lower)	1.74	18.9	131,773	6,965
Relative risk for other mortality (upper)	1.79	18.5	132,507	7,152
Relative risk for other mortality (lower), CRC treatment cost	1.74, €54,975	18.9	29,695	1,570
Relative risk for other mortality (upper), CRC treatment cost	1.79, €23,561	18.5	234,300	12,647

COPD: chronic obstructive pulmonary disease, CRC: colorectal cancer.

5 Discussion

Despite evidence of important heterogeneity in CRC screening and existing methodology to account for this in decision analyses, few economic evaluations of population-based CRC screening programs systematically address patient heterogeneity, often because individual-level data are not available. We applied the SA framework developed by Coyle et al. (2003), which does not require individual-level data, to a cost-effectiveness analysis of once-only sigmoidoscopy screening for CRC to demonstrate that decisions in favor of providing the same intervention to the entire population can result in inefficient use of resources. Our results generally indicate that screening women and patients with comorbidities considered unrelated to CRC led to less attractive cost-effectiveness profiles, while screening men and comorbidities considered related to CRC led to more attractive cost-effectiveness profiles. Reducing future CRC incidence may therefore outweigh an increase in potential adverse events and reduced life expectancy. These findings argue against screening recommendations that consider individuals of similar ages as a homogenous population. Instead, we argue that analyses and policy recommendations should consider heterogeneity in screening outcomes not only for age, but also for sex, comorbidities and other characteristics, as this may improve health and reduce costs in society. This is especially relevant for Norway, which is currently rolling-out a national CRC screening program without consideration of personal characteristics other than age, despite evidence from a large Norwegian randomized trial suggesting sigmoidoscopy-screening may not be effective for women. Although alternative CRC screening technologies may be used in population-based screening (e.g., fecal occult blood testing, colonoscopy, computed tomography colonography), our findings suggest that women may yield greater benefits if different screening algorithms catered to their risk profiles were recommended (not a one-size-fits-all national recommendation).

Our analysis included clinically defined subgroups, for which differential baseline risks for CRC development and all-cause mortality are well-established in the literature, in conjunction with a normalization of the applied relative risks that enabled matching of model outcomes with national epidemiologic statistics. Similar to the previous studies that considered patient heterogeneity in CRC screening, we found that CRC risk and life expectancy are key variables that affect the cost-effectiveness of CRC screening in opposite directions (Braithwaite, 2011; Lansdorp-Vogelaar et al., 2014; van Hees et al., 2015). In contrast to our findings, Dinh et al. (2012) found that the cost-effectiveness of screening for CRC involving colonoscopy was negatively affected by the presence of diabetes. Unfortunately, we were unable to directly compare the parameter values for the baseline risks of CRC development and all-cause mortality in diabetes patients with our study. None of the previous studies explicitly assessed the value of stratification for specifically defined subgroups, nor incorporated the most recent evidence on heterogeneity in screening outcomes from a large randomized controlled trial (Holme et al., 2018).

There are several frameworks for accounting for patient heterogeneity in decision analyses that differ with respect to the types of data (i.e., individual level vs aggregate/secondary) required to perform a subgroup analysis, which inherently dictate the type of framework that can be employed in an analysis. The expected value of individualized care (EVIC) framework, developed by Basu and Meltzer (2007), is based on calculating the monetary gain obtained by determining the optimal treatment for specific individual patients compared with providing all patients with the optimal treatment identified on the population average. The authors make a distinction in calculating the EVIC with and without cost internalization (i.e. based on cost-effectiveness versus based on effectiveness alone). Espinoza et al. (2014) extended the SA and EVIC frameworks into the value of

heterogeneity (VoH) framework, by including an efficiency frontier to determine the optimal level of stratification that accounts for the transaction costs associated with identifying additional subgroups in practice. This framework also distinguishes between the value of a subgroup policy based on current information and the potential value of performing additional research to reduce the uncertainty surrounding parameter values within subgroups. The SA framework that we used in our analysis does not require individual level data, and provides insight in whether a subgroup policy may be valuable by quantifying the efficiency gains obtained from stratification. Unlike the EVIC and VoH frameworks, the SA framework does not allow for estimating the optimal level of stratification or consideration of cost-internalization.

There are some important limitations to consider. The structure of the model is simplified and does not include cancer stages or consideration of alternative screening technologies and algorithms, as we did not aim to determine the optimal method and timing of CRC screening, but rather demonstrate the potential importance of capturing heterogeneity in CRC screening in general. More complex models can evaluate more refined algorithms that may identify optimal screening approaches for each subgroup that vary by screening intensity (e.g., age to start, screening frequency), rather than a choice between an ‘all or nothing’ approach. We reflected patient heterogeneity by adjusting baseline risks of overall mortality, relative treatment effect for men and women, CRC-specific incidence and mortality, adverse screening events and baseline health-related quality of life. However, patient heterogeneity may also exist in the relative treatment effect across comorbidity subgroups, resource utilization and the quality of life decrement for CRC (Grutters et al., 2013). As data on these other potential sources of heterogeneity are limited, we could only include explore the impact of these other factors in a sensitivity analysis.

Improving expected population health benefits by further stratification may have to be weighed against an increase in uncertainty as a result of using smaller subsets of data (Sculpher, 2008). Well-defined subgroups may also explain, in part, some differences in screening outcomes between patients and thereby reduce uncertainty. We were however unable to explore this trade-off in the absence of individual level data.

The comorbidities that were selected are not mutually exclusive; therefore, it was not possible to calculate the full efficiency gains yielded from a subgroup policy accounting for multiple comorbidities compared to using average population estimates. However, the potential gain of using a subgroup policy will be equal or greater than the gain in the individual comorbidity subgroup. Overlap between comorbidities (e.g. patients with both diabetes and COPD) should be further explored, as the effect of having multiple comorbidities is not clear.

We expect the transaction costs related to developing and implementing tools to identify subgroups in clinical practice to decrease efficiency associated with a stratified policy (Espinoza et al., 2014). In addition, leakage or non-adherence to a LUC policy, defined as the provision of treatment to patients for whom it is not recommended, will necessarily result in a loss in the NMBs gained from stratification (Coyle et al., 2003). Transaction costs required to identify subgroups, and leakage could potentially offset the efficiency gained from stratification, and should be considered when determining the optimal subgroup policy. Importantly, several comorbidity subgroups included in the analysis could in practice easily be identified and selected for screening by linking general practitioner’s registry data, resulting in a relatively low transaction cost. In contrast, restricting organized screening for specific subgroups (e.g. smoking and obesity) may not be possible due to identification problems *inter alia*.

There are several ethical considerations associated with stratified cost-effectiveness analyses that are not only relevant for screening, but also for healthcare initiatives such as

genetic testing, e-health and other interventions that are targeted at heterogeneous populations. An implicit assumption of the SA framework is that the gains in efficiency are calculated by restricting treatment to subgroups with a positive INMB. However, subgroups with a negative INMB do not necessarily have to be excluded from a screening program altogether. A subgroup-specific policy may also involve restricting interventions aimed at increasing screening uptake in those subgroups that are expected to benefit most, or to include subgroup-specific screening algorithms that vary by technology, intensity and cost (including ages to initiate screening, ages to terminate screening, and frequency of screening) for different patient groups. Some factors (e.g. a small relative treatment effect, or high mortality due to other causes than CRC) may ethically be more acceptable as basis for a stratified policy than other factors (e.g. a low quality of life in life years saved). Importantly, enumerating the costs and benefits associated with subgroups and evaluating what factors affect screening outcomes allows decision makers to make an explicit consideration of the potential trade-off between equity and efficiency.

6 Conclusion

Accounting for patient heterogeneity in CRC screening will likely improve the value of national screening policies. Decision makers should consider the role of patient heterogeneity when evaluating population-based screening programs, as decisions in favor of providing or rejecting interventions for the entire population can result in inefficient use of resources.

Acknowledgements

We acknowledge The Norwegian Cancer Society (grant no: 58140001) and the Research Council of Norway (grant no: 238042) for grant support. Further we thank participants and discussants at the Eu-HEA meeting in Hamburg (July 2016), Nordic Health Economics Study Group Meeting in Odense (August 2016), and at Voksenåsen (March 2016), and SMDM in Vancouver (October 2016).

References

- Aas, E. (2015). Costs of screening for colorectal cancer: Results of the pilot study in Østfold and Bærum. Available at: https://www.kreftregisteret.no/globalassets/publikasjoner-og-rapporter/tarmkreftscreening/kostnadsanalyse_pilot_screening_tykk-endetarmskreft.pdf.
- American Cancer Society. (2018). Colorectal cancer risk factors. Available at: <http://www.cancer.org/cancer/colonandrectumcancer/detailedguide/colorectal-cancer-risk-factors>.
- Bala, M. V., & Zarkin, G. A. (2004). Pharmacogenomics and the Evolution of Healthcare. Is it Time for Cost-Effectiveness Analysis at the Individual Level? *Pharmacoeconomics*, 22(8), 495-498.
- Basu, A., & Meltzer, D. (2007). Value of information on preference heterogeneity and individualized care. *Medical Decision Making*, 27(2), 112-127.
- Basu, P., Ponti, A., Anttila, A., Ronco, G., Senore, C., Vale, D. B., . . . Primic Žakelj, M. (2018). Status of implementation and organization of cancer screening in The European Union Member States—Summary results from the second European screening report. *International journal of cancer*, 142(1), 44-56.
- Botteri, E., Iodice, S., Bagnardi, V., Raimondi, S., Lowenfels, A. B., & Maisonneuve, P. (2008). Smoking and colorectal cancer: a meta-analysis. *Jama*, 300(23), 2765-2778.
- Braithwaite, R. S. (2011). Can life expectancy and QALYs be improved by a framework for deciding whether to apply clinical guidelines to patients with severe comorbid disease? *Medical Decision Making*, 31(4), 582-595.

- Burström, K., Johannesson, M., & Diderichsen, F. (2001). Swedish population health-related quality of life results using the EQ-5D. *Quality of life research*, 10(7), 621-635.
- Cancer Registry of Norway. (2017). Cancer in Norway 2016 - Cancer incidence, mortality, survival and prevalence in Norway. 2017. Available at: <https://www.kreftregisteret.no/globalassets/cancer-in-norway/2016/cin-2106.pdf>.
- Carter, B. D., Abnet, C. C., Feskanich, D., Freedman, N. D., Hartge, P., Lewis, C. E., . . . Jacobs, E. J. (2015). Smoking and mortality--beyond established causes. *New England Journal of Medicine*, 372(7), 631-640.
- Cho, H., Klabunde, C. N., Yabroff, K. R., Wang, Z., Meekins, A., Lansdorp-Vogelaar, I., & Mariotto, A. B. (2013). Comorbidity-adjusted life expectancy: a new tool to inform recommendations for optimal screening strategies. *Annals of Internal Medicine*, 159(10), 667-676.
- Coyle, D., Buxton, M. J., & O'Brien, B. J. (2003). Stratified cost-effectiveness analysis: a framework for establishing efficient limited use criteria. *Health Economics*, 12(5), 421-427.
- De Bruijn, K. M., Arends, L. R., Hansen, B. E., Leeftang, S., Ruiter, R., & van Eijck, C. H. (2013). Systematic review and meta-analysis of the association between diabetes mellitus and incidence and mortality in breast and colorectal cancer. *British Journal of Surgery*, 100(11), 1421-1429.
- De Wit, G. A., Busschbach, J. J., & De Charro, F. T. (2000). Sensitivity and perspective in the valuation of health status: whose values count? *Health Economics*, 9(2), 109-126.
- Dinh, T. A., Alperin, P., Walter, L. C., & Smith, R. (2012). Impact of comorbidity on colorectal cancer screening cost-effectiveness study in diabetic populations. *Journal of General Internal Medicine*, 27(6), 730-738.
- Espinoza, M. A., Manca, A., Claxton, K., & Sculpher, M. J. (2014). The value of heterogeneity for cost-effectiveness subgroup analysis: conceptual framework and application. *Medical Decision Making*, 34(8), 951-964.
- Flegal, K. M., Kit, B. K., Orpana, H., & Graubard, B. I. (2013). Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis. *Jama*, 309(1), 71-82.
- Grutters, J. P., Sculpher, M., Briggs, A. H., Severens, J. L., Candel, M. J., Stahl, J. E., . . . Joore, M. A. (2013). Acknowledging patient heterogeneity in economic evaluation : a systematic literature review. *Pharmacoeconomics*, 31(2), 111-123.
- Hakim, Z., Wolf, A., & Garrison, L. P. (2002). Estimating the Effect of Changes in Body Mass Index on Health State Preferences. *Pharmacoeconomics*, 20 (6), 393-404.
- Holme, Ø., Løberg, M., Kalager, M., Bretthauer, M., Hernán, M. A., Aas, E., . . . Schneede, J. (2018). Long-Term Effectiveness of Sigmoidoscopy Screening on Colorectal Cancer Incidence and Mortality in Women and Men: A Randomized Trial. *Annals of Internal Medicine*. 168(11), s 775- 782 .
- Joranger, P., Nesbakken, A., Hoff, G., Sorbye, H., Oshaug, A., & Aas, E. (2015). Modeling and validating the cost and clinical pathway of colorectal cancer. *Medical Decision Making*, 35(2), 255-265.
- Koopmanschap, M. A., & Rutten, F. F. (1996). A practical guide for calculating indirect costs of disease. *Pharmacoeconomics*, 10(5), 460-466.
- Lansdorp-Vogelaar, I., Gulati, R., Mariotto, A. B., Schechter, C. B., de Carvalho, T. M., Knudsen, A. B., . . . Mandelblatt, J. S. (2014). Personalizing age of cancer screening cessation based on comorbid conditions: model estimates of harms and benefits. *Annals of Internal Medicine*, 161(2), 104-112.
- Lansdorp-Vogelaar, I., van Ballegooijen, M., Zauber, A. G., Boer, R., Wilschut, J., Winawer, S. J., & Habbema, J. D. (2009). Individualizing colonoscopy screening by sex and race. *Gastrointestinal Endoscopy*, 70(1), 96-108.

- Lin, J. S., Piper, M. A., Perdue, L. A., Rutter, C. M., Webber, E. M., O'connor, E., . . . Whitlock, E. P. (2016). Screening for colorectal cancer: updated evidence report and systematic review for the US Preventive Services Task Force. *Jama*, 315(23), 2576-2594.
- Lonne, G., Johnsen, L. G., Aas, E., Lydersen, S., Andresen, H., Ronning, R., & Nygaard, O. P. (2015). Comparing cost-effectiveness of X-Stop with minimally invasive decompression in lumbar spinal stenosis: a randomized controlled trial. *Spine (Phila Pa 1976)*, 40(8), 514-520.
- Lozano, R., Naghavi, M., Foreman, K., Lim, S., Shibuya, K., Aboyans, V., . . . Memish, Z. A. (2012). Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*, 380(9859), 2095-2128.
- Mills, K. T., Bellows, C. F., Hoffman, A. E., Kelly, T. N., & Gagliardi, G. (2013). Diabetes mellitus and colorectal cancer prognosis: a meta-analysis. *Diseases of the Colon and Rectum*, 56(11), 1304-1319.
- Moghaddam, A. A., Woodward, M., & Huxley, R. (2007). Obesity and risk of colorectal cancer: a meta-analysis of 31 studies with 70,000 events. *Cancer Epidemiology, Biomarkers and Prevention*, 16(12), 2533-2547.
- Norwegian Directorate of Health. (2016). Activity-based funding 2014. Available at: <https://helsedirektoratet.no/finansieringsordninger/innsatsstyrt-finansiering-isf-og-drg-systemet>.
- Norwegian Institute of Public Health. (2016a). Dementia in Norway. Available at: <https://www.fhi.no/en/op/hin/health-disease/dementia-in-norway---public-health-/#how-much-will-the-occurrence-increase-during-the-years-to-come>.
- Norwegian Institute of Public Health. (2016b). Overweight and obesity in Norway - Public health report 2014. Available at: <https://www.fhi.no/en/op/public-health-report-2014/risk--protective-factors/overweight-and-obesity-in-norway---/>.
- Norwegian Labour and Welfare Administration. (2015). Labour market statistics. Available at: nav.no.
- Norwegian Medicines Agency. (2018). Guidelines for the submission of documentation for single technology assessment (STA) of pharmaceuticals.
- Norwegian Renal Registry. (2016). Annual Report 2014. Available at: https://legemiddelverket.no/Documents/English/Price%20and%20reimbursement/Application%20for%20reimbursement/Guidelines_april_2018.pdf.
- Phipps, A. I., Baron, J., & Newcomb, P. A. (2011). Prediagnostic smoking history, alcohol consumption, and colorectal cancer survival: the Seattle Colon Cancer Family Registry. *Cancer*, 117(21), 4948-4957.
- Redekop, W. K., Koopmanschap, M. A., Stolk, R. P., Rutten, G. E. H. M., R. Wolffenduttel, B. H., & Niessen, L. W. (2002). Health-Related Quality of Life and Treatment Satisfaction in Dutch Patients With Type 2 Diabetes. *Diabetes Care*, 25(3).
- Reichle, K., Peter, R. S., Concin, H., & Nagel, G. (2015). Associations of pre-diagnostic body mass index with overall and cancer-specific mortality in a large Austrian cohort. *Cancer Causes Control*, 26(11), 1643-1652.
- Rutten-van Molken, M. P., Oostenbrink, J. B., Tashkin, D. P., Burkhart, D., & Monz, B. U. (2006). Does quality of life of COPD patients as measured by the generic EuroQol five-dimension questionnaire differentiate between COPD severity stages? *Chest*, 130(4), 1117-1128.
- Salomon, J. A., Vos, T., Hogan, D. R., Gagnon, M., Naghavi, M., Mokdad, A., . . . Murray, C. J. L. (2012). Common values in assessing health outcomes from disease and injury: disability weights measurement study for the Global Burden of Disease Study 2010. *The Lancet*, 380(9859), 2129-2143.

- Schreuders, E. H., Ruco, A., Rabeneck, L., Schoen, R. E., Sung, J. J., Young, G. P., & Kuipers, E. J. (2015). Colorectal cancer screening: a global overview of existing programmes. *Gut*, *64*(10), 1637-1649.
- Sculpher, M. (2008). Subgroups and Heterogeneity in Cost-Effectiveness Analysis. *Pharmacoeconomics*, *26*(9), 799-806.
- Sharp, L., Tilson, L., Whyte, S., Ceilleachair, A. O., Walsh, C., Usher, C., . . . Comber, H. (2013). Using resource modelling to inform decision making and service planning: the case of colorectal cancer screening in Ireland. *BMC Health Service Research*, *13*, 105.
- Solli, O. (2013). Diabetes in Norway: Costs, Health-related Quality of Life and Cost-effectiveness of Life-style Interventions. *Faculty of Medicine, University of Oslo*.
- Statistics Norway. (2016). Smoking habits. Available at: <https://www.ssb.no/en/helse/statistikker/royk/aar/2015-02-03>.
- Statistics Norway. (2018a). National life tables 2017. Available from: <https://www.ssb.no/en/dode>
- Statistics Norway. (2018b). Population count 2017. Available at: <https://www.ssb.no/en/befolkning?de=Population+count>.
- Van Hees, F., Saini, S. D., Lansdorp-Vogelaar, I., Vijan, S., Meester, R. G., de Koning, H. J., . . . van Ballegooijen, M. (2015). Personalizing colonoscopy screening for elderly individuals based on screening history, cancer risk, and comorbidity status could increase cost effectiveness. *Gastroenterology*, *149*(6), 1425-1437.
- Vogl, M., Wenig, C. M., Leidl, R., & Pokhrel, S. (2012). Smoking and health-related quality of life in English general population: implications for economic evaluations. *BMC Public Health*, *12*, 203.
- Waatevik, M., Skorge, T. D., Omenaas, E., Bakke, P. S., Gulsvik, A., & Johannessen, A. (2013). Increased prevalence of chronic obstructive pulmonary disease in a general population. *Respiratory Medicine*, *107*(7), 1037-1045.
- Warren, J. L. (2009). Adverse Events After Outpatient Colonoscopy in the Medicare Population. *Annals of Internal Medicine*, *150*(12), 849.
- Wolfs, C. A., Dirksen, C. D., Kessels, A., Willems, D. C., Verhey, F. R., & Severens, J. L. (2007). Performance of the EQ-5D and the EQ-5D+C in elderly patients with cognitive impairments. *Health and Quality of Life Outcomes*, *5*, 33.