Modeling in colorectal cancer screening: assessing external and predictive validity of MISCAN-Colon microsimulation model using NORCCAP trial results

Published as:

Buskermolen M, Gini A, Naber SK et al. Modeling in Colorectal Cancer Screening: Assessing External and Predictive Validity of MISCAN-Colon Microsimulation Model Using NORCCAP Trial Results. Med Decis Making. 2018 Nov;38(8):917-929.

ABSTRACT

Background: Microsimulation models are increasingly being used to inform colorectal cancer (CRC) screening recommendations. MISCAN-Colon is an example of such a model, used to inform the Dutch CRC screening program and United States Preventive Services Task Force guidelines. Assessing the validity of these models is essential to provide transparency regarding their performance. In this study we tested the external and predictive validity of MISCAN-Colon.

Methods: We validated MISCAN-Colon using the Norwegian Colorectal Cancer Prevention (NORCCAP) trial, a randomized controlled trial that examined the effectiveness of once-only flexible sigmoidoscopy (FS) screening. We simulated the study population and design of the NORCCAP trial in MISCAN-Colon and compared 10- to 12-year model-predicted hazard ratios (HRs) for overall and distal CRC incidence and mortality to those observed. In addition, we compared the numbers of screen-detected neoplasia. Finally, we predicted the trial's future results to allow for the assessment of predictive validity.

Results: MISCAN-Colon predicted an HR for overall CRC incidence (0.85), distal CRC incidence (0.82), overall CRC mortality (0.68) and distal CRC mortality (0.62). These were within the limits of the 95% confidence intervals of the NORCCAP trial results. Similar results were observed for the number of screen-detected cancers. The model significantly underestimated the number of screen-detected adenomas. Model-predicted HRs for CRC incidence and mortality up to 15- to 17-years follow-up were 0.84 and 0.72, respectively.

Conclusion: Although the underestimation of screen-detected adenomas requires further investigation, MISCAN-Colon is able to make a valid replication of the CRC incidence and mortality reduction of an FS screening trial, which suggests that it can be considered a useful tool to support decision making on CRC screening.

INTRODUCTION

Governments and health organizations aim to offer cancer screening programs that are effective, affordable and have a low burden for participants.[109, 110] Deciding which cancer screening strategy is most suitable for these programs is a complex task that involves making decisions in terms of type of screening test(s), frequency of testing and age range, type of follow-up test(s), and risk stratification for high-risk populations. The health benefits of cancer screening must be as large as possible and substantially exceed potential harms or patient burden. All these aspects can be incorporated into microsimulation models, which can predict the population-level impact of screening strategies in an affordable, timely, and ethical manner.[69, 111] The Microsimulation Screening Analysis-Colon (MISCAN-Colon) model is an example of a well-established microsimulation model that has been used to inform decision making on colorectal cancer (CRC) screening, including the design of the Dutch CRC screening program and the United States Preventive Services Task Force (USPSTF) guidelines on CRC screening.[69, 111] MISCAN-Colon simulates the development of adenomas, which may or may not progress to clinical CRC.[112] To simulate the sequence from adenoma formation to clinical CRC, the model incorporates parameter values that are derived from published data such as adenoma prevalence and lifetime CRC incidence.[113, 114] However, some other essential parameter values, which are crucial to simulating the adenoma carcinoma sequence and highly relevant to estimating the effectiveness of screening, are not available from existing evidence. Some characteristics of the sequence from adenoma to clinical cancer are difficult or impossible to observe in an ethically acceptable manner. For instance, the duration from adenoma formation to clinical CRC cannot be observed. Parameter estimates for these durations must therefore be inferred from data on adenoma prevalence and (interval) cancer incidence. In case of MISCAN-Colon, this inference is performed by calibration using results available from randomized controlled trials (RCTs) that investigate the effectiveness of CRC screening.[69] To ensure that model calibration is correct and model predictions are valid, regular assessment of model validity is essential. Model validation is an important process in model development. In literature several levels of model validity have been proposed: face, internal, cross, external, and predictive.[115, 116] The most robust levels of validity can be established through external and predictive validation; these validations entail comparing model results with real-world results and comparing model results with prospectively observed events. Although we found examples of other publications in which microsimulation models were externally validated, we did not find any examples of predictive validations. MISCAN-Colon has been validated externally previously, using the results of the UK Flexible Sigmoidoscopy Screening (UKFSS) trial, which involved once-only screening for CRC with flexible sigmoidoscopy (FS) with follow-up over a 10-year period. This validation was published by Rutter et al.[112] As a consequence, the MISCAN model was recalibrated using the UKFSS trial data, resulting in a longer average

duration of adenoma progression to cancer. Reassessing the performance of the recalibrated MISCAN model now requires revalidation.

In this study, we aimed to establish 2 types of validity of MISCAN-Colon. First, we aimed to reassess the external validity after the recalibration on the UKFSS trial. Second, we aimed to establish predictive validity of MISCAN-Colon. For these validations, we used the results of the Norwegian Colorectal Cancer Prevention (NORCCAP) trial, which involved once-only screening for CRC with FS.

METHODS

We used MISCAN-Colon to simulate NORCCAP trial outcomes and compared predictions with those observed. Primary validation targets were relative overall and distal CRC incidence reduction and mortality reduction observed by Holme et al.),[39] who described the 10- to 12-year follow-up results of the NORCCAP trial. These relative reductions were presented as the hazard ratio (HR) of an event in the intervention group relative to the same event in the control group. We calculated 4 HRs; overall and distal CRC incidence, as well as overall and distal CRC mortality. To simulate the NORCCAP trial, we adjusted MISCAN-Colon to the demography and screening behavior of the NORCCAP trial population.

NORCCAP trial

In the NORCCAP trial, individuals between the ages of 50 to 65 years from 2 Norwegian regions were randomly assigned to either a control group (n=78,220) or an intervention group that consisted of 2 arms (n=10,283 and n=10,289). Since there was no screening program in place in Norway during the study period, the control group did not receive routine CRC screening.[39] Baseline characteristics of the selected individuals are shown in **Supplementary Table 3.1**. In 1 intervention arm, individuals were offered a once-only FS (n=10,283). In the other intervention arm, individuals were offered an additional qualitative fecal occult blood test (FOBT) before FS (n=10,289), and 86.7% of the adherers to FS made use of this opportunity.[39, 117-119] A positive FS was defined as any polyp with a diameter of >10 mm or any histologically verified adenoma or carcinoma.[119] Individuals with a positive FS or FOBT were referred for follow-up colonoscopy.

The trial was carried out in 2 phases; individuals born from 1935 to 1945 were selected and randomized to undergo screening in 1999 and 2000 (i.e. 53-65 years old at time of screening), and individuals born from 1946 to 1950 were selected and randomized to undergo screening in 2001 (i.e. 49-54 years old at the time of screening). The latest NORCCAP trial publication covered all CRC-related events until December 31, 2011 (follow-up of 10 to 12 years).[39] In this latest publication, no distinction was made between the 2 different intervention arms regarding the results relevant for the current validation studies. Therefore, we compared model

outcomes with the overall results of the intervention arms. In the remainder of this article, we will use the term intervention group when referring to both intervention arms.

MISCAN-Colon

MISCAN-Colon is a microsimulation model for CRC developed at the Department of Public Health of the Erasmus University Medical Center (Rotterdam, the Netherlands). The model's structure, underlying assumptions, and calibration have been described in previous publications[120-122] and in General Appendix. Briefly, MISCAN-Colon simulates the life histories of a large population of individuals from birth until death. As each simulated person ages, 1 or more adenomas may develop. These adenomas can progress from small (≤5mm), to medium (6-9mm), to large (≥10mm). Some adenomas can develop into preclinical cancer, which may progress through stages I to IV. During each stage, CRC may be diagnosed because of symptoms. Survival after clinical diagnosis is determined by the stage at diagnosis, the localization of the cancer, and the person's age.[123]

Screening will alter some of the simulated life histories: some cancers will be prevented by the detection and removal of adenomas; other cancers will be detected in an earlier stage with a more favorable survival. However, screening can also result in serious complications, over-diagnosis and over-treatment (i.e. the detection and treatment of adenomas or cancers that would not have been diagnosed in the absence of screening). By comparing life histories with screening with the corresponding life histories without screening, MISCAN-Colon quantifies the effectiveness of screening, as well as the associated costs.

MISCAN-Colon was calibrated to the age-, stage-, and localization-specific incidence and survival of CRC as observed in Norway during the timeframe of the NORCCAP trial (1999-2011).[124] Data was provided by the Norwegian Cancer Registry. The age-specific prevalence and multiplicity distribution of adenomas was calibrated using the observations of autopsy studies.[114, 125-134] The preclinical duration of CRC and the adenoma dwell-time were calibrated to the rates of interval- and surveillance-detected cancers observed in RCTs evaluating screening using guaiac FOBTs and the once-only sigmoidoscopy UKFSS trial.[135-138]

Adjustment of MISCAN-Colon to the NORCCAP trial

We used MISCAN-Colon to simulate a population with an age distribution comparable to the NORCCAP trial (personal communication with research leader G. Hoff, 2016). Lifetables for 2005 (i.e. middle of the study period) were retrieved from Statistics Norway.[139] CRC incidence in the NORCCAP control group was 11% lower than incidence in the whole of Norway. We therefore adjusted the model accordingly by lowering the age-specific onset of adenomas by 11% for all ages. Comparing incidence rates observed in the NORCCAP trial, we assumed that nonadherers had a slightly higher age-specific onset of adenomas for all ages than individuals in the control group (relative risk of 1.05). In addition, age-specific onset in adherers was lowered for all ages to ensure that the overall CRC risk

in the intervention group did not differ from the CRC risk in the control group, taking participation rate into account (relative risk of 0.97). The control group was simulated for 18 years without intervention. Outcomes for CRC incidence and mortality were evaluated after 10 to 12 years (i.e. consistent with published results of the trial) and for every next year until 18 years of follow-up. Individuals with negative screening were followed for 18 years without further intervention, while for those with adenomas detected we simulated surveillance consistent with the Norwegian recommendations at the time of the trial.[140] We assumed age-specific participation rates for FS and FOBT as observed in the NORCCAP trial (**Supplementary Table 3.2**, personal communication with research leader G. Hoff, 2016).[39] Adherence at follow-up colonoscopy was derived from Holme et al.[39] Adherence for surveillance colonoscopies was not reported in trial publications; it was assumed to be 80%. Since FOBT characteristics can differ due to varying cutoff levels and manufacturers and the characteristics of the FOBT used in the NORCCAP trial are unknown, test sensitivity and specificity of the 1-sample FOBT could not be estimated from literature. We therefore fitted sensitivity and specificity to observed positivity rate and to detection rates of non-advanced adenomas, advanced adenomas, and carcinomas as observed in the NORCCAP trial (**Supplementary Table 3.3**, **3.4** and **3.5**). We assumed that test characteristics of FS and colonoscopy do not differ greatly between settings, and therefore, test sensitivity of FS and follow-up colonoscopy and specificity of follow-up colonoscopy were based on literature. [141] The test specificity of FS was adjusted based on the number of referrals after a negative test in the NORCCAP trial (**Table 3.4**). We simulated complete visualization of the rectosigmoid colon in 97% of individuals, of the descending colon in 23%, and of the cecum in less than 1% of individuals (personal communication with G. Hoff, 2010). We simulated that colonoscopy examinations completely visualized the sigmoid in more than 99% of the cases and completely visualized the entire colon in 89% of the cases (personal communication with G. Hoff, 2010). Using MISCAN-Colon, we simulated 4 different cohorts that each 10 million individuals differing by study arm (control group, intervention group) and age group (50-54 and 55-64) to rule out any distortion caused by the stochastic nature of the model. Model predictions were then rescaled to the size of the NORCCAP trial population.

Validation targets

Our primary validation targets were the overall and distal CRC incidence and mortality rate and HRs of overall and distal CRC incidence and mortality at 10- to 12-year follow-up (depending on the year of trial inclusion), in the intervention group relative to the control group. We defined the rectum, rectosigmoid, and sigmoid colon as distal locations, consistent with reported NORCCAP trial results. In addition, to enable the predictive validity of MISCAN-Colon after publication of the next NORCCAP trial results to be tested, we calculated the expected HRs of overall and distal CRC incidence and mortality up to 18 years of follow-up.

In addition, we considered several secondary validation targets. We computed the cumulative probability of overall and distal CRC incidence and mortality during the study period for intervention and control group. We computed the yearly risk ratios (RRs) of overall CRC incidence and mortality in the intervention group relative to the control group, as well as yearly RRs for CRC incidence and distal CRC incidence in the adherers relative to the control group during the follow-up of the trial. We also compared the number of screen-detected cancers and adenomas and the number of follow-up colonoscopies. In addition, we explored model-predicted stage distribution of all diagnosed cancers. Proportion per stage – localized (Dukes A and B) v. advanced (Dukes C and D) – were calculated and compared using a chi-squared test. Model outcomes were considered consistent when predicted within 95% confidence intervals (CIs) of the corresponding NORCCAP trial targets. Mathematical formulas for these outcomes are provided in **Supplementary Table 3.6**.

RESULTS

Overall and distal CRC incidence and mortality rates and hazard ratios

During the 10- to 12-year follow up of the NORCCAP trial, 141.0 CRC cases per 100,000 person-years occurred in the control group (95% CI, 132.8-149.7) and 112.6 in the intervention group (95% CI, 99.3-127.7), resulting in a lower risk of CRC incidence in those invited to FS screening (HR=0.80; 95% CI, 0.70-0.92).[39] Using MISCAN-Colon, we predicted an overall CRC incidence of 141.8 cases per 100,000 person-years in the control group and 120.1 cases in the intervention group (**Table 3.1A**, **Figure 3.1**).

Table 3.1. Hazard ratios: 10-12 years follow-up interventions effects of the NORCCAP trial including 95% confidence intervals for these effects and MISCAN- Colon predictions of these effects.

Outcome	Source	HR	Per 100,000 person-years	
			Control	Screened
A. CRC overall				
CRC Mortality	NORCCAP trial	0.73 (0.56, 0.94)	43.1 (38.7, 48.1)	31.4 (24.8, 39.7)
	MISCAN-Colon	0.68	40.5	27.8
CRC Incidence	NORCCAP trial	0.80 (0.70, 0.92)	141.0 (132.8, 149.7)	112.6 (99.3, 127.7)
	MISCAN-Colon	0.85	141.8	120.1
A. CRC distal				
Distal CRC Mortality	NORCCAP trial	0.79 (0.55, 1.11)	21.8 (18.7, 25.4)	17.2 (12.6,23.5)
	MISCAN-Colon	0.62	21.6	13.4
Distal CRC Incidence	NORCCAP-trial	0.76 (0.63, 0.92)	80.1 (74, 86.7)	60.9 (51.4, 72.2)
	MISCAN-Colon	0.82	78.2	64.0

Abbreviations: CRC, colorectal cancer; NORCCAP, Norwegian Colorectal Cancer Prevention Trial; MISCAN, Microsimulation Screening Analysis; HR, hazard ratio

These predicted CRC incidence rates were similar to the trial results and resulted in an HR for CRC incidence with once-only FS screening (with or without once-only FOBT) v. no screening of 0.85, consistent with the NORCCAP trial results. During the same simulated time frame in the NORCCAP trial, 43.1 CRC deaths per 100,000 person-years were reported in the control group (95% CI, 38.7-48.1) and 31.4 in the intervention group (95% CI, 24.8-39.7), showing in those invited to once-only FS screening a lower probability of dying of CRC (HR=0.73; 95% CI, 0.56-0.94).[39] MISCAN-Colon predicted 40.5 CRC deaths per 100,000 person-years in the control group and 27.8 in the intervention group, similar to the trial results. In addition, among those invited to FS screening, MISCAN-Colon predicted a lower probability of dying of CRC (HR=0.68), consistent with trial results (**Table 3.1A**, **Figure 3.1**). When considering only trial results on distal CRC incidence and mortality, MISCAN-Colon performances were similar: model-predicted distal CRC incidence and mortality rates, as well as HRs of distal CRC incidence and mortality, were all consistent with the observed trial results (**Table 3.1B** and **Figure 3.1**).

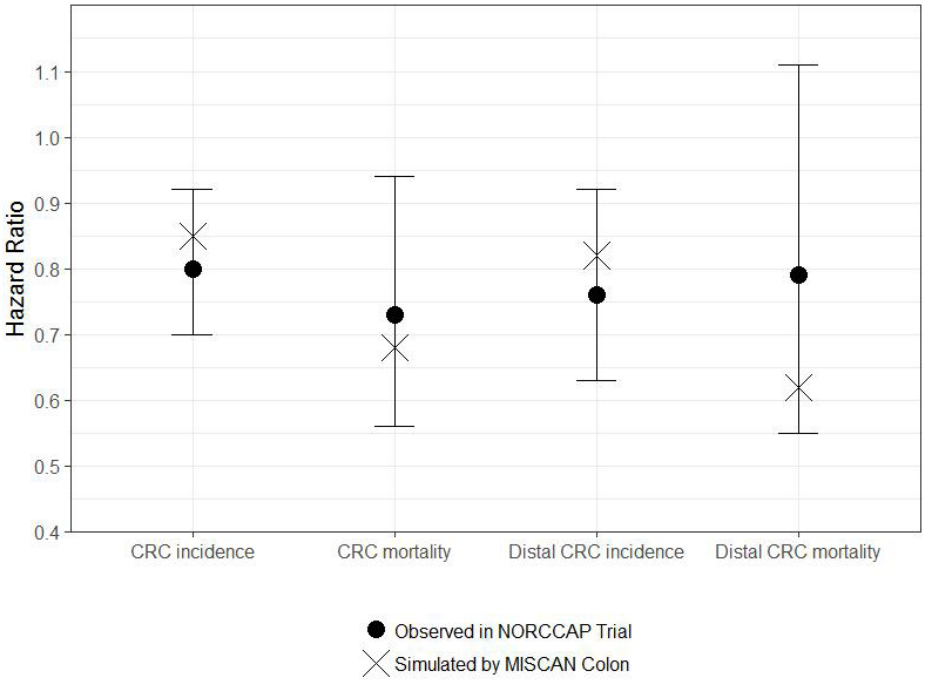


Figure 3.1. Hazard ratios: 10- to 12-year follow-up intervention effects of NORCCAP trial including 95% confidence intervals for these effects and MISCAN-Colon predictions of these effects. Abbreviations: CRC, colorectal cancer; NORCCAP, Norwegian Colorectal Cancer Prevention Trial; MISCAN, Microsimulation Screening Analysis.

Cumulative probability of overall and distal CRC incidence and mortality

Most MISCAN-Colon predictions of the cumulative probability of overall CRC incidence and mortality and distal CRC mortality in the control and intervention groups were consistent with the NORCCAP trial results (**Figure 3.2**). MISCAN-Colon underestimated some of the cumulative probabilities in the first half of the trial follow-up. In the last years of follow-up, the predicted cumulative probability of overall CRC incidence in the control and intervention groups increased more than expected based on trial results, leading to a small but significant difference in the final years.

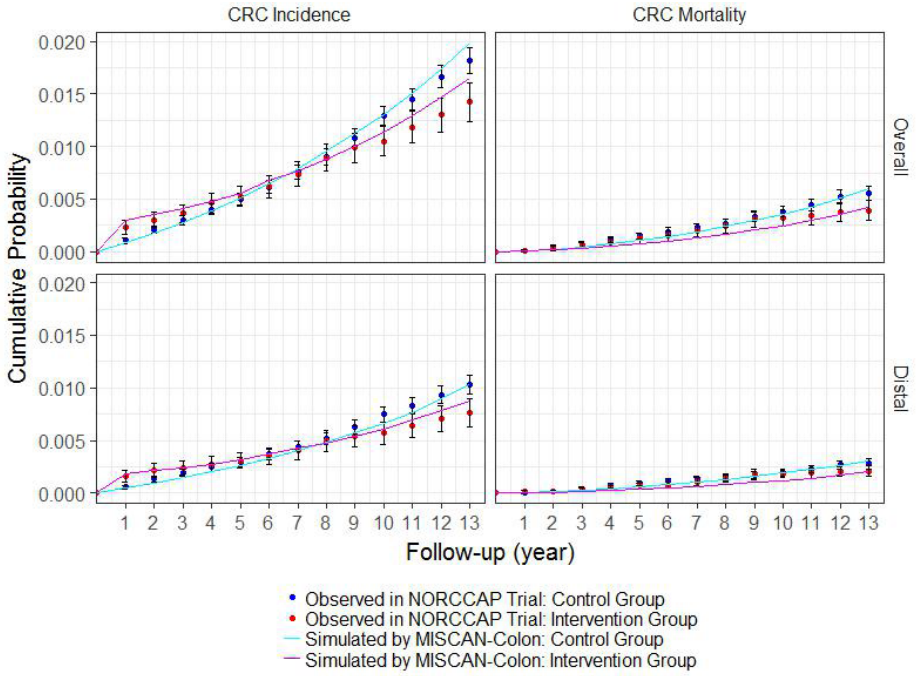


Figure 3.2. Cumulative probability of overall CRC incidence, overall CRC mortality, distal CRC incidence and distal CRC mortality: 10- to 12-year follow-up intervention effects of NORCCAP trial including 95% confidence intervals for these effects and MISCAN-Colon predictions of these effects. This figure is a replication of Figure 2 as published in Holme et al. (2014). Abbreviations: CRC, colorectal cancer; NORCCAP, Norwegian Colorectal Cancer Prevention; MISCAN, Microsimulation Screening Analysis.

Yearly risk ratios of CRC incidence and mortality

Most of the MISCAN-Colon predictions were consistent with the NORCCAP trial results regarding yearly RRs for overall CRC incidence of the intervention group relative to the control group (**Figure 3.3**). MISCAN-Colon significantly overestimated relative overall CRC incidence risk in the intervention group in year 1 and 10 and significantly underestimated this risk in year 8. Regarding yearly RRs for overall CRC mortality of the intervention group relative to the control group, the MISCAN-Colon predictions were consistent with

the NORCCAP trial results. Similar patterns were observed when comparing the RRs for overall CRC incidence and distal CRC incidence of the adherers to screening relative to the control group (**Figure 3.4**).

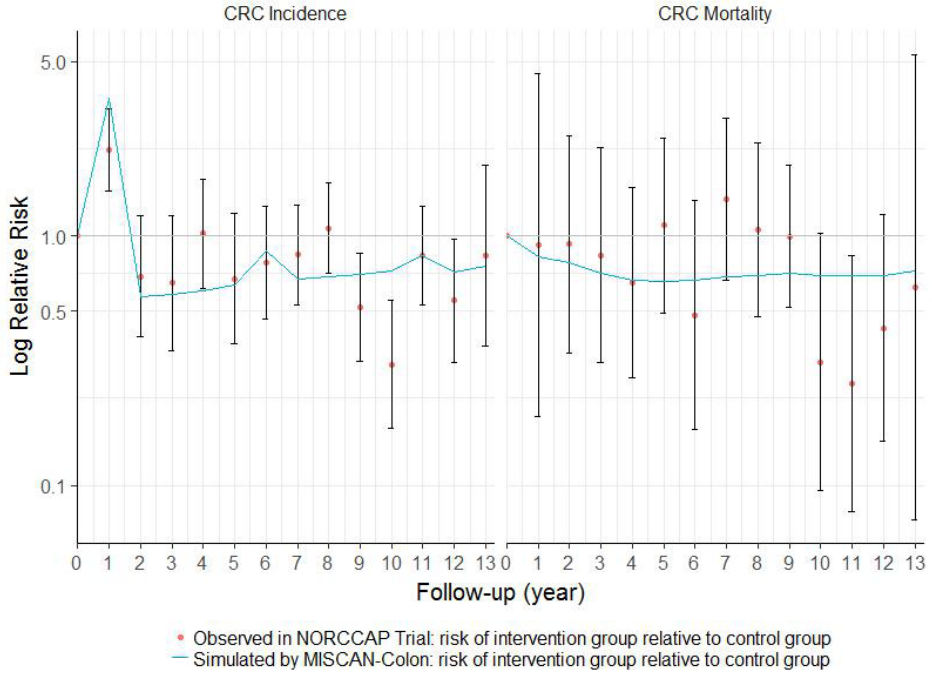


Figure 3.3. Yearly risk ratios for colorectal cancer incidence and mortality in screening group relative to the control group: 10- to 12-year follow-up intervention effects of NORCCAP trial including 95% confidence intervals for these effects and MISCAN-Colon predictions of these effects. This figure is a replication of Figure 3 as published in Holme et al. (2014)
Abbreviations: CRC, colorectal cancer; NORCCAP, Norwegian Colorectal Cancer Prevention Trial; MISCAN, Microsimulation Screening Analysis.

MISCAN-Colon overestimated relative overall CRC incidence risk in year 1 and underestimated this risk in year 8. The predictions of relative distal CRC incidence risk in adherers were all within the confidence intervals of the NORCCAP trial results.

Disease detection at screening

MISCAN-Colon predicted 41 screen-detected CRCs, which was consistent with the NORCCAP trial results. For the number of follow-up colonoscopies and screen-detected adenomas, the MISCAN-Colon predictions were significantly lower than what was actually observed. While in the NORCCAP trial, 2524 (95% CI, 2432-2616) colonoscopies were performed and 2210 (95% CI, 2123-2297) adenomas were detected, MISCAN-Colon predicted 2408 colonoscopies performed and 2105 adenomas detected (**Table 3.2**).

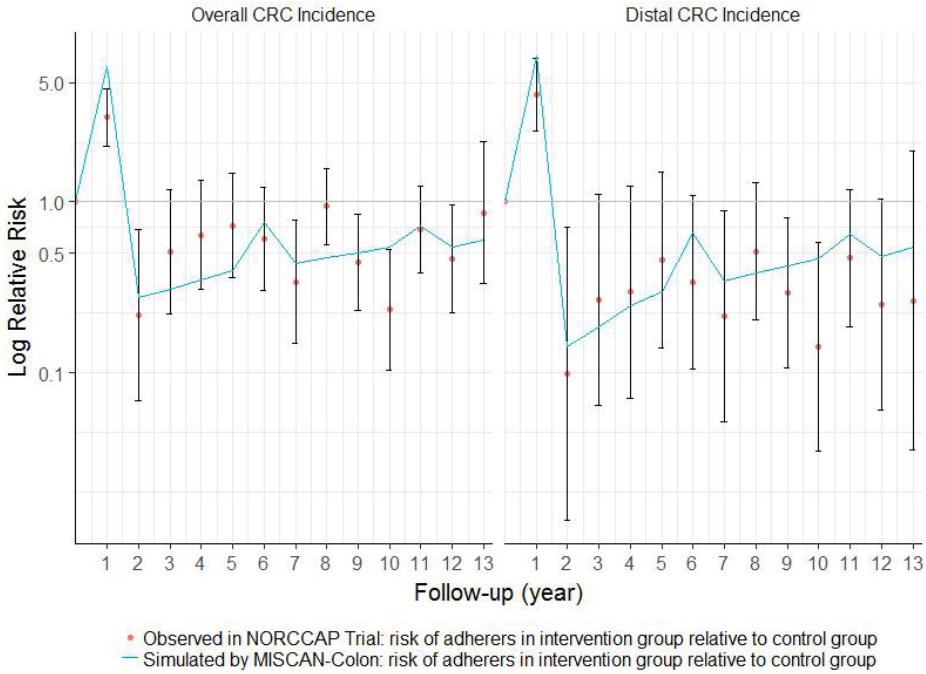


Figure 3.4. Yearly risk ratios for overall and distal colorectal cancer incidence in screening adherers relative to control group: 10- to 12-year follow-up intervention effects of NORCCAP trial including 95% confidence intervals for these effects and MISCAN-Colon predictions of these effects. This figure is a replication of Figure 4 as published in Holme et al. (2014)

Abbreviations: CRC, colorectal cancer; NORCCAP, Norwegian Colorectal Cancer Prevention Trial; MISCAN, Microsimulation Screening Analysis.

Table 3.2. Outcomes at screening: NORCCAP trial results and MISCAN-Colon predictions of these results. Numbers of individuals are reported with 95% confidence intervals for NORCCAP trial results.

Outcome	Source	Number	95% interval
Diagnostic colonoscopies	NORCCAP trial	2524	(2432, 2616)
	MISCAN-Colon	2408	
CRC detected at screening	NORCCAP trial	41	(28, 54)
	MISCAN-Colon	52	
Adenomas detected at colonoscopy			
Total	NORCCAP trial	2210	(2123, 2297)
	MISCAN-Colon	2105	
Advanced adenomas	NORCCAP trial	582	(535, 629)
	MISCAN-Colon	519	
Non-advanced adenomas	NORCCAP trial	1628	(1552, 1704)
	MISCAN-Colon	1586	

Abbreviations: CRC, colorectal cancer; NORCCAP, Norwegian Colorectal Cancer Prevention Trial; MISCAN, Microsimulation Screening Analysis; HR, hazard ratio

Stage distribution

For stage distribution, MISCAN-Colon predictions of both intervention and control group were similar to those observed in the trial (**Table 3.3**).

Table 3.3. Stage distribution of diagnosed colorectal cancers during the 10-12 year follow-up of the NORCCAP trial compared to MISCAN-Colon predictions

	NORCCAP trial*		MISCAN-Colon		
Control Group	No.	(%)	No.	(%)	P value
Localized CRC	470	(45.5%)	489	(47.3%)	
Advanced CRC	562	(54.5%)	545	(52.7%)	0.45
Intervention group					
Localized CRC	117	(49.4%)	155	(52.6%)	
Advanced CRC	120	(50.6%)	139	(47.4%)	0.50

Abbreviations: CRC, colorectal cancer; NORCCAP, Norwegian Colorectal Cancer Prevention Trial; MISCAN, Microsimulation Screening Analysis; *Unclassified cancers in the control group (N=16) and in the intervention group (N=4) were excluded from this table.

Prediction of future follow-up results

For the 15- to 17- year follow-up of the NORCCAP trial, MISCAN-Colon predicted an HR of 0.84 for overall CRC incidence, an HR of 0.72 for overall CRC mortality, an HR of 0.81 for distal CRC incidence, and an HR of 0.66 for distal CRC mortality (**Table 3.4**). NORCCAP trial results for these years are not yet available.

Table 3.4. Hazard ratios: MISCAN-Colon predictions for future follow-up results NORCCAP trial.

Follow up years*	End of data retrieval**	Overall CRC mortality			Overall CRC incidence			Distal CRC mortality			Distal CRC incidence		
		HR	Control group	Screen group	HR	Control group	Screen group	HR	Control group	Screen group	HR	Control group	Screen group
10-12	2011	0.68	40.5	27.8	0.85	142	120	0.62	21.6	13.4	0.82	78.2	64.0
11-13	2012	0.69	43.2	29.8	0.84	147	123	0.63	23.0	14.4	0.81	81.4	65.8
12-14	2013	0.70	45.8	31.9	0.83	153	127	0.63	24.4	15.5	0.80	84.5	67.8
13-15	2014	0.70	48.4	34.1	0.84	159	133	0.64	25.8	16.6	0.80	87.7	70.5
14-16	2015	0.71	51.1	36.4	0.84	164	138	0.65	27.1	17.7	0.81	90.8	73.2
15-17	2016	0.72	53.8	38.7	0.84	170	143	0.66	28.6	18.9	0.81	93.9	75.8
16-18	2017	0.73	56.5	41.0	0.84	175	147	0.67	30.1	20.1	0.81	96.9	78.1

Abbreviations: CRC, colorectal cancer; NORCCAP, Norwegian Colorectal Cancer Prevention; MISCAN, Microsimulation Screening Analysis; HR, hazard ratio

Numbers under control group and screen group presented per 100 000 person years.

*The screening intervention was performed in 1999, 2000 and 2001. Since the closure date for data retrieval is the same for all participants, the number of follow-up years differs among the participants.

**Last day of the year

DISCUSSION

In this study, we tested the validity of the MISCAN-Colon model using data from the NORCCAP trial. Regarding our primary validation targets, we showed that MISCAN-Colon can accurately estimate the impact of an once-only FS screening trial on CRC incidence and mortality. In addition, we expect the follow-up results of the NORCCAP trial to be published in the near future and then we will be able to compare our predictions to these results to test predictive validity of MISCAN-Colon. Regarding our secondary validation targets, MISCAN predictions for cumulative probabilities of incidence and mortality and yearly RRs of CRC incidence and mortality in the intervention group relative to the control group were in line with the NORCCAP trial results as well. The predicted number of screen-detected CRCs was also similar to the number found in the NORCCAP trial, but the model significantly underestimated the number of screen-detected adenomas.

It is essential that microsimulation models, such as MISCAN-Colon, are validated regularly to provide transparency regarding their performance. External validation requires data of (large) clinical trials with sufficient follow-up.[115, 116] Previously, MISCAN-Colon was validated externally using the results of another once-only FS screening trial – namely, the UKFSS trial.[112] MISCAN-Colon then underestimated CRC incidence reduction due to screening and overestimated the screen-detection of adenomas and cancers in the intervention arm. These outcomes suggested that the assumed values for the duration of adenoma formation to symptomatic CRC were too short in MISCAN-Colon. As a consequence of these validation findings, MISCAN-Colon was re-calibrated using the UKFSS trial data, resulting in a longer average duration of adenoma progression to symptomatic CRC. In the current validation study, MISCAN-Colon predictions were highly similar to the NORCCAP trial results, suggesting that the recalibrated MISCAN-Colon allows for accurate predictions of CRC incidence and mortality reduction of FS screening.

Since age-specific CRC incidence, stage distribution, and survival differ per country, region, and timeframe, we adjusted the model to the Norwegian population during the NORCCAP trial period. The adjustments specific for Norway were independent from the trial data. In 2 instances, we decided to adjust the model inputs based on the control group of the NORCCAP trial, which makes this external validation partially dependent (as described in Eddy et al.[116]). First of all, we noticed that CRC incidence in the NORCCAP trial control group was 11% lower than in the Norwegian CRC registry, which may be attributed to regional differences in CRC incidence. We therefore lowered the age-specific adenoma onset for all ages by 11%. This adjustment may affect some absolute outcomes such as CRC incidence and mortality rate in intervention and control group but not the relative impact of the screening intervention. Second, since CRC incidence in nonadherers may be higher due to 'healthy screenee bias,' we compared the control group CRC incidence to the CRC incidence in nonadherers in the intervention group. Consequently, we raised the age-specific

adenoma onset for all ages in the nonadherers with 5% and lowered this multiplier for adherers to screening such that the modelled overall CRC risk in the intervention group was equal to the CRC risk in the control group. This adjustment had no substantial impact on the validation targets. In supplementary methods (**Supplementary Tables 3.7-3.10** and **Figures 3.1-3.4**), the results without these corrections based on CRC incidence in the control group and in the nonadherers are shown, and indeed, some absolute outcomes are different but the relative outcomes are largely the same. These types of adjustments are needed to ensure appropriate external validation of any screening simulation model. Importantly, we did not use any information on screening participants of the NORCCAP trial in our model adjustments. Therefore, we consider this analysis to be an external validation of screening effectiveness and unobservable parameter values in MISCAN-Colon.

Despite the encouraging findings of well-predicted HRs, some secondary outcomes were not consistent with the NORCCAP trial. We observed 3 discrepancies between the simulated and observed data. First, we observed that the number of screen-detected adenomas predicted by MISCAN-Colon was lower than the actual number of screen-detected adenomas in the NORCCAP trial, while incidence (reduction) was correctly predicted. We have 3 possible explanations for these seemingly conflicting outcomes. First, the outcomes may not be as conflicting as they seem. Having too few adenomas in the model implies that we may have overestimated progression of distal adenomas to match distal cancer incidence, which is consistent with the slightly underestimated distal CRC incidence reduction as simulated by MISCAN-Colon compared to that observed. Second, we lowered the CRC risk in the model to reflect the lower incidence in the control group compared to the Norwegian incidence rate. However, we do not know whether the lower risk holds for all ages or just for those ages included in the trial. Consequently, we may have underestimated the CRC risk at older ages and thus the prevalence of adenomas in the ages before that (i.e. in the ages being screened). Finally, the NORCCAP trial is just 1 trial with data on adenoma detection rates. In previous validations to other studies, model-predicted adenoma detection rates have been close to those observed. The NORCCAP data may be an outlier in this respect, as, for instance, the distal adenoma detection rate was 12.1% in the UKFSS trial, and 17.4% in the NORCCAP trial.[39, 142] MISCAN-Colon predicted a distal adenoma detection rate of 15.4%.

The second discrepancy we found was that not all predicted yearly RRs of overall and distal CRC incidence and mortality of the intervention group relative to the control group were within the confidence intervals of the NORCCAP trial results. We suggest 2 explanations for this. First, performance dates of surveillance colonoscopies were not registered in the study. We suspect that incorrect predictions of yearly RRs in the intervention group relative to the control group (as shown in **Figure 3.3**) are related to the adherence to surveillance after a positive colonoscopy. MISCAN-Colon simulated surveillance at exactly 5 and 10 years after initial screening, which is consistent with Norwegian screening guidelines.[140] However, the RRs of CRC incidence in the intervention group of the NORCCAP trial

showed peaks 1 to 3 years earlier than we would expect if surveillance would have been performed at 5 and 10 years after a positive colonoscopy. It seems plausible that some of the participants might have undergone surveillance 1 to 3 years earlier. Second, the observed yearly RRs of mortality in the NORCCAP trial fluctuate. Therefore, rather than an underestimation of MISCAN-Colon of the RR of mortality in the intervention group compared to the control group in year 7 (as shown in **Figure 3.3**), the high mortality in the intervention group in that year may have been the result of chance.

Last, despite that all the MISCAN-Colon predictions are within the 95% confidence intervals, the prediction for distal CRC mortality deviates considerably from that observed. Although this deviation could be interpreted as a lack of fit, one should be careful with such an assessment. Confidence intervals reflect the level of plausibility of each estimation. It means that if a certain number of trials similar to NORCCAP were performed, in 95% of these trials, distal CRC mortality reduction would have been reported in the 95% CI of the NORCCAP trial. From an inference point of view, the 95% CI represents the interval for which we are 95% confident that the true value falls within its limits. Since the numbers of distal CRC deaths occurring in both intervention and control group in the NORCCAP trial are very low (substantially less than the number of overall and distal CRC cases and the number of overall CRC deaths), this wide confidence interval reflects the uncertainty of the results. Therefore, it is too early to conclude whether our MISCAN-Colon predictions are correct or incorrect. Validation against the pooled results of several sigmoidoscopy trials such as been published in Holme et al.[143] is an obvious next step to assess model fit against distal CRC mortality.

Despite the increased use of simulation models to inform cancer screening programs, very few of those models have been extensively validated. We searched for other publications regarding external validation of microsimulation models used for predicting cancer screening effectiveness. We found that publications explicitly demonstrating external validation of cancer screening microsimulation models are scarce and, to the best of our knowledge, publications regarding predictive validation were nonexistent. In a systematic review of Koleva et al.,[144] it was concluded that none of the models used for breast cancer screening were externally validated. However this finding may be nuanced by arguing that external validation is sometimes performed without publishing the results.[145] In addition to the review of Koleva et al., we found 2 external validation studies of models on ovarian cancer screening[146, 147] and several on lung cancer screening.[148-151] Although these models are designed to predict the impact of interventions, only one of these was validated for important screening effectiveness outcomes such as incidence and mortality reduction. In the current study, we validated, besides mortality and incidence reduction outcomes, a variety of other intermediate outcomes. These intermediate outcomes are also highly relevant for the validity of a model predicting cancer screening effectiveness, as they may lead to very different predictions with respect to the cost-effectiveness of screening. In our opinion, this

elaborate validation is an important strength of the current work. In addition, assessing the predictive validity of the model is an additional novel feature in the validation of cancer screening simulation models.

Irrespective of these strengths, 2 limitations are noteworthy. First, we did not vary the sensitivity of screening and follow-up tests by location of adenomas. Although studies indicate that the sensitivity of FOBT and follow-up colonoscopy for right-sided premalignant lesions in the colon may differ from the sensitivity for left-sided premalignant lesions, there is not yet consensus on this topic.[152-154] Second, although this study offers promising evidence of the validity of our model, it does not directly imply that MISCAN-Colon predictions are also valid for other settings, such as other screening trials with different screening tests. Validation of MISCAN-Colon is a continuous process that will be frequently repeated whenever new important results regarding CRC screening RCTs are published. In this continuous process, we have already validated our model using 5 of 9 RCTs included in the Cochrane Library on the benefits of CRC screening:[155] 3 out of 4 guaiac FOBT trials;[156] and, including this study, 2 out of 5 FS trials.[112, 157] Model validation using 2 of the remaining FS trials may not be feasible or useful considering that the interpretation of one of the trials may be affected by the frequent occurrence of opportunistic screening among the trial population,[158] and the other includes only a small number of participants.[155] We are currently in the process of validating the model against the remaining FS trial[157] as well as data from the Italian fecal immunochemical test (FIT) screening program performed in Florence during 1993-2008 (mean follow-up, 11 years). The first results are promising, further indicating the validity of MISCAN-Colon for FS and FIT screening effectiveness.

In conclusion, this study demonstrates that the MISCAN-Colon model can accurately estimate the main outcomes of a trial that measures the effectiveness of once-only FS CRC screening. These findings, in combination with our other validation results, suggest that MISCAN-Colon is a useful decision-making tool for public health organizations and governments involved in CRC screening. Furthermore, we made predictive validation possible by presenting our model outcomes before publication of trial results. Finally, by publishing the results of this validation study, we can provide more transparency regarding the performance of modelling in general, which is crucial for the role of modelling in public health decision making.

Acknowledgments:

We thank the leaders of the NORCCAP trial for providing data regarding essential model inputs: Ø. Holme; M. Løberg; M. Bretthauer; G. Hoff

In addition, we thank the Norwegian Cancer registry for providing data regarding CRC incidence and mortality in Norway.

SUPPLEMENTARY METHODS

Supplementary Table 3.1. Baseline characteristics of study participants of the Norwegian Colorectal Cancer Prevention Trial. *

	No(%)			
	Group			
	Control (n=78 220)		Screening (n=20 572)	
Age, mean (SD), y*	56.1	(3.8)	56.9	(3.8)
Sex				
Men	38922	(49.8)	10269	(49.9)
Women	39298	(50.2)	10303	(20.1)
Age group, y				
50-54	37131	(47.5)	6920	(33.6)
55-64	41089	(52.5)	13652	(66.4)
Area of residence				
Telemark County	15176	(19.4)	10314	(50.1)
City of Oslo	63044	(80.6)	10258	(49.9)

*Data retrieved from Table 1, Holme et al. (2014). There was no data available regarding personal or family history for any of the individuals selected for the trial.

Supplementary Table 3.2. Comparison of incidence rate in NORCCAP trial control group and in intervention group: adherers and non-adherers.

	No. Individuals*	**CRC incidence***	**Incidence rate**	**Baseline incidence**
Control group	78220	1086	0.0139	1
Non-Adherers	7617	111	0.0146	1.05
Adherers	12955	142	0.0110	0.971**

*Data retrieved from Table 4, Holme et al. (2014)

**Computation of incidence rate for adherers:

incidence ratio non adherers relative to control group = (111/7617)/(1086/78220)=1.050

incidence ratio adherers relative to control group = $(1\text{-}1.05^{*}(1\text{-}0.63^{\#}))/0.63^{\#}$

#adherence in intervention group'

Supplementary Table 3.3. Observed adherence rates in the NORCCAP trial compared to simulated adherence rates in MISCAN-Colon

Characteristic	NORCCAP trial	MISCAN-Colon
Adherence flexible sigmoidoscopy	63.0%*	63.0%
Adherence FOBT in adherers to flexible sigmoidoscopy (involving only one intervention arm) **	86.7%**	86.7%
Adherence diagnostic colonoscopy	95.6%*	95.6%
Adherence surveillance colonoscopy	*unknown*	80%
Reach sigmoid sigmoidoscopy	97% ***	97%
Reach coecum colonoscopy	89%***	89%

Abbreviation: FOBT, fecal occult blood test.
*Data retrieved from Gondal et al. (2003)
** FOBT was performed in adherers to flexible sigmoidoscopy, before flexible sigmoidoscopy was performed
***Provided by research leader G. Hoff

Supplementary Table 3.4. Screening test characteristics used in MISCAN-Colon to simulate the NORCCAP trial

Test characteristic	Sigmoidoscopy	Colonoscopy	FOBT
Sensitivity small adenomas (<5 mm)	75%	75%	0%
Sensitivity medium adenomas (6-9 mm)	85%	85%	7.6%
Sensitivity large adenomas (>10 mm)	95%	95%	17.6%
Sensitivity CRC long before clinical detection*	95%	95%	35.2%
Sensitivity CRC short before clinical detection*	95%	95%	71.6%
Specificity	97.6%	100 %	96.3%

Abbreviation: CRC, colorectal cancer.
*Sensitivity of FOBT is higher in the stage in which the cancer would have been diagnosed in the absence of screening than in earlier stages.[138] FOBT: fecal occult blood test.

Supplementary Table 3.5. Observed positivity rate* and positive predictive values** of screening tests of the NORCCAP trial compared to those simulated in MISCAN-Colon

	FOBT		Sigmoidoscopy	
	NORCCAP trial	MISCAN-Colon	NORCCAP trial	MISCAN-Colon
Positivity rate	5.6%	5.5%	18.8%	16.5%
PPV	44.1%	45.3%	96.0%	94.7%
PPV adenomas	17.8%	18.7%	69.7%	70.6%
PPV Advanced adenomas	22.0%	22.3%	24.8%	22.8%
PPV cancer	4.3%	4.3%	1.6%	1.4%

Abbreviations: FOBT, fecal occult blood test; PPV, positive predictive value
*positivity rate: the number of individuals with a positive FOBT or flexible sigmoidoscopy (either false or true positive) divided by the total number of individuals adhering to FOBT or flexible sigmoidoscopy.
** positive predictive value: the number of individuals with adenomas, advanced adenomas or CRC detected at follow-up colonoscopy divided by the number of individuals adhering to follow-up colonoscopy after a positive FOBT or sigmoidoscopy.

Supplementary Table 3.6. Mathematical formulations of the study outcomes

Primary outcome targets	Formula	Confidence Interval	Comments
Incidence and mortality rates	$incidence\ (or\ mortality)\ rate = \frac{d_{(i)}}{n_{(i)}}$ $n_{(i)}$= number of person years until year i $d_{(i)}$ = number of events until year i	We computed corresponding 95% Confidence Intervals assuming a Poisson distribution as follows: Lower bound: $L = \frac{\chi^2_{2d_{(i)},0.025}}{2n_{(i)}}$ Upper bound: $U = \frac{\chi^2_{2(d_{(i)}+1),0.975}}{2n_{(i)}}$	For the NORCCAP Trial Results these were directly computed from the results published in Holme et al. 2014 (as authors did not report 95% confidence intervals for incidence and mortality rates)
Hazard ratio	In the NORCCAP trial, authors computed age-adjusted Hazard ratios (HRs) using a Cox models including age as covariate in the models. In our analysis, using Cox model was not a feasible way. Thus, we adjusted the model to simulate age-specific population with and without screening and we computed, therefore, "*age-adjusted*" rate ratios as: $\frac{incidence\ (or\ mortality)\ rate\ intervention\ group}{incidence\ (or\ mortality)\ rate\ control\ group}$ Assuming that these measures were for definition age-adjusted, we assumed those consistent estimations of trial's HR. Furthermore, we assumed those measures as an accurate estimation of relative risk $Rate\ Ratios \approx HR \approx RR$	Confidence intervals for HRs were computed in Holme et al. (2014) using Cox models.	For the NORCCAP trial results these were directly obtained from the results as published in Holme et al. (2014)
Secondary outcome targets			

Supplementary Table 3.6. Mathematical formulations of the study outcomes (*continued*)

Primary outcome targets	Formula	Confidence Interval	Comments
Cumulative probability	$Cumulative\ probability = 1 - \hat{S}(t) = 1 - \prod_{t_{(i)} \le t} \frac{n_{(i-)} - d_{(i)}}{n_{(i-)}}$ With $\hat{S}(t) = 1\ for\ t < t_{(i)}$ t = time in years $n_{(i-)}$ = number of person years in year i $d_{(i)}$ = number of events in year i	95% Confidence Interval may be computed as follow: Lower bound: $^{L} = 1 - (\hat{S}(t) + 1.96\sqrt{\hat{V}\{\hat{S}(t)\}})$ Upper bound: $^{U} = 1 - (\hat{S}(t) - 1.96\sqrt{\hat{V}\{\hat{S}(t)\}})$ With variance V computed using Greenwood's formula: $\hat{V}\{\hat{S}(t)\} = \hat{S}(t)^2 \sum_{t_{(i)} \le t} \frac{d_{(i)}}{n_{(i-)}(n_{(i-)} - d_{(i)})}$	For the NORCCAP Trial Results and the MISCAN-Colon predictions the same formula was used. Required inputs from the NORCCAP trial were provided by the trial leaders.
Yearly risk ratio	$\widehat{RR}_i = \frac{a_i/(a_i + b_i)}{c_i/(c_i + d_i)}$ a_i = number of persons with events in year i in intervention group b_i = number of persons without events in year i in intervention group c_i = number of persons with events in year i in control group d_i = number of persons without events in year i in control group	95% Confidence Interval are computed as follows: Lower bound: $l = e^{\ln \widehat{RR}_i - 1.96\widehat{SD}[\ln \widehat{RR}_i]}$ Upper bound: $U = e^{\ln \widehat{RR}_i + 1.96\widehat{SD}[\ln \widehat{RR}_i]}$ Where: $\widehat{SD}[\ln \widehat{RR}_i] = \sqrt{\frac{b_i/a_i}{a_i + b_i} + \frac{d_i/c_i}{c_i + d_i}}$	For the NORCCAP Trial Results and the MISCAN-Colon predictions the same formula was used. Required inputs from the NORCCAP trial were provided by the trial leaders.

Supplementary Table 3.6. Mathematical formulations of the study outcomes (*continued*)

Primary outcome targets	Formula	Confidence Interval	Comments
Colonoscopy attendance Screen-detected adenomas and cancers Stage distribution	These values were computed as proportions. Thus, ratio between events e and population under study η, as follows: $proportion = \hat{p} = \frac{e}{n}$ In the following table we reported the values used in computing each fraction: Proportion | e | n Colonoscopy attendance | Individuals that performed colonoscopy | Invited individuals Screen-detected adenomas/cancer | Screen-detected adenomas (cancers) | Invited individuals Localized cancer | Localized cancer diagnosed | All cancer diagnosed Advanced cancer | Advanced cancer diagnosed | All cancer diagnosed	95% Confidence Interval are computed as follows: Lower bound: $L = \hat{p} - 1.96\widehat{SD}$ Upper bound: $U = \hat{p} + 1.96\widehat{SD}$ *Where:* $\widehat{SD} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ These proportions and corresponding 95% confidence intervals were reported in relation to n. Thus, $\hat{P}$, L , and U were multiplied for n (total individuals invited or total number of cancer diagnosed in NORCAAP trial) before reporting those values in study's tables. In addition, we compared localized (advanced) cancer proportions using Pearson χ^2test (Rothman KJ, Greenland S. Modern Epidemiology (2nd edition). Philadelphia: Lippincott-Raven 1998).	For the NORCCAP Trial Results these were directly obtained from the values as published in Holme et al. (2014). Confidence intervals were computed using the formulas reported in this table.

Supplementary Table 3.7. Hazard ratios: 10-12 years follow-up interventions effects of the NORCCAP trial including 95% confidence intervals for these effects and MISCAN- Colon predictions of these effects.

Outcome	Source	HR	Per 100,000 person-years	
			Control	Screened
A. CRC overall				
CRC Mortality	NORCCAP trial	0.73 (0.56, 0.94)	43.1 (38.7, 48.1)	31.4 (24.8, 39.7)
	MISCAN-Colon	0.68	45.5	30.8
CRC Incidence	NORCCAP trial	0.80 (0.70, 0.92)	141 (132.8, 149.7)	112.6 (99.3, 127.7)
	MISCAN-Colon	0.83	159.9	133.3
A. CRC distal				
Distal CRC Mortality	NORCCAP trial	0.79 (0.55, 1.11)	21.8 (18.7, 25.4)	17.2 (12.6,23.5)
	MISCAN-Colon	0.62	24.2	15.0
Distal CRC Incidence	NORCCAP-trial	0.76 (0.63, 0.92)	80.1 (74, 86.7)	60.9 (51.4, 72.2)
	MISCAN-Colon	0.82	88.3	71.1

Abbreviations: CRC, colorectal cancer; NORCCAP, Norwegian Colorectal Cancer Prevention Trial; MISCAN, Microsimulation Screening Analysis; HR, hazard ratio

Supplementary Table 3.8. Outcomes at screening: NORCCAP trial results and MISCAN Colon predictions of these results. Numbers of individuals are reported with 95% confidence intervals for NORCCAP trial results.

Outcome	Source	Number	95% interval
Diagnostic colonoscopies	NORCCAP trial	2524	(2432, 2616)
	MISCAN-Colon	2732	
CRC detected at screening	NORCCAP trial	41	(28, 54)
	MISCAN-Colon	59	
Adenomas detected at colonoscopy			
Total	NORCCAP trial	2210	(2123, 2297)
	MISCAN-Colon	2432	
Advanced adenomas	NORCCAP trial	582	(535, 629)
	MISCAN-Colon	595	
Non-advanced adenomas	NORCCAP trial	1628	(1552, 1704)
	MISCAN-Colon	1838	

Abbreviations: CRC, colorectal cancer; NORCCAP, Norwegian Colorectal Cancer Prevention Trial; MISCAN, Microsimulation Screening Analysis; HR, hazard ratio

Supplementary Table 3.9. Stage distribution of diagnosed colorectal cancers during the 10-12 year follow-up of the NORCCAP trial compared to MISCAN Colon predictions

	NORCCAP trial*		MISCAN-Colon		
Control Group	No.	(%)	No.	(%)	P value
Localized CRC	470	(45.5%)	538	(47.3%)	
Advanced CRC	562	(54.5%)	571	(52.7%)	0.45
Intervention group					
Localized CRC	117	(49.4%)	173	(52.6%)	
Advanced CRC	120	(50.6%)	154	(47.4%)	0.50

Abbreviations: CRC, colorectal cancer; NORCCAP, Norwegian Colorectal Cancer Prevention Trial; MISCAN, Microsimulation Screening Analysis; *Unclassified cancers in the control group (N=16) and in the intervention group (N=4) were excluded from this table.

Supplementary Table 3.10. Hazard ratios: MISCAN Colon predictions for future follow up results NORCCAP trial.

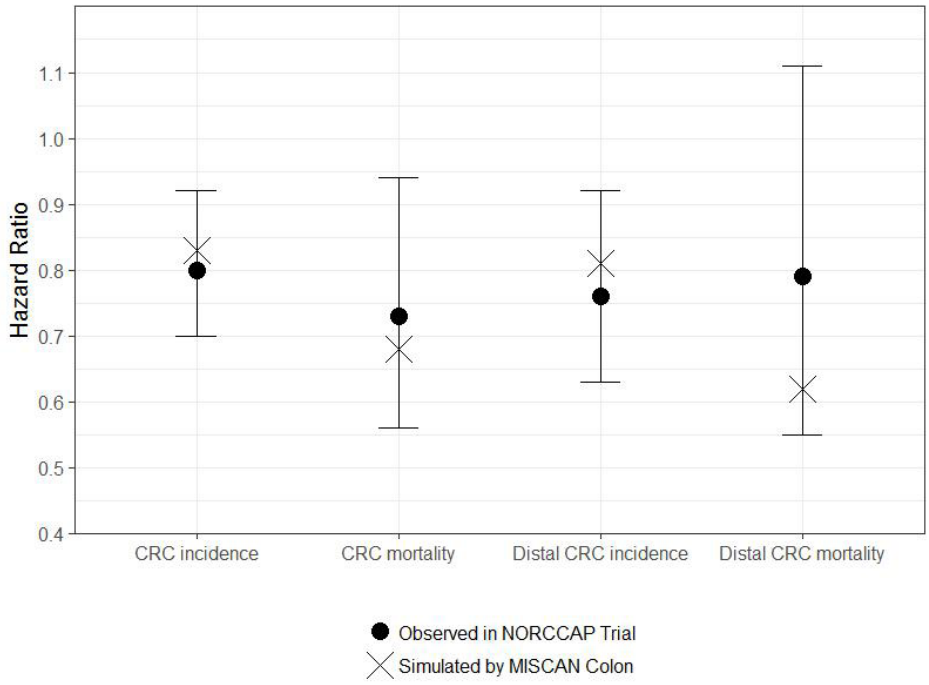
Follow up years*	End of data retrieval **	Overall CRC mortality			Overall CRC incidence			Distal CRC mortality			Distal CRC incidence		
		HR	Control group	Screen group	HR	Control group	Screen group	HR	Control group	Screen group	HR	Control group	Screen group
10-12	2011	0.68	45.5	30.8	0.83	160	133	0.62	24.2	15.0	0.81	88.3	71.1
11-13	2012	0.68	48.5	33.0	0.82	166	137	0.63	25.8	16.1	0.80	91.8	73.0
12-14	2013	0.69	51.4	35.4	0.82	173	142	0.63	27.4	17.3	0.79	95.4	75.3
13-15	2014	0.69	54.5	37.7	0.82	179	148	0.64	29.0	18.5	0.79	99.0	78.3
14-16	2015	0.70	57.5	40.2	0.83	186	153	0.65	30.6	19.8	0.79	102.5	81.3
15-17	2016	0.71	60.4	42.7	0.83	192	159	0.66	32.2	21.1	0.79	106.1	84.1
16-18	2017	0.71	63.6	45.2	0.82	198	163	0.66	33.9	22.4	0.79	109.6	86.6

Abbreviations: CRC, colorectal cancer; NORCCAP, Norwegian Colorectal Cancer Prevention; MISCAN, Microsimulation Screening Analysis; HR, hazard ratio

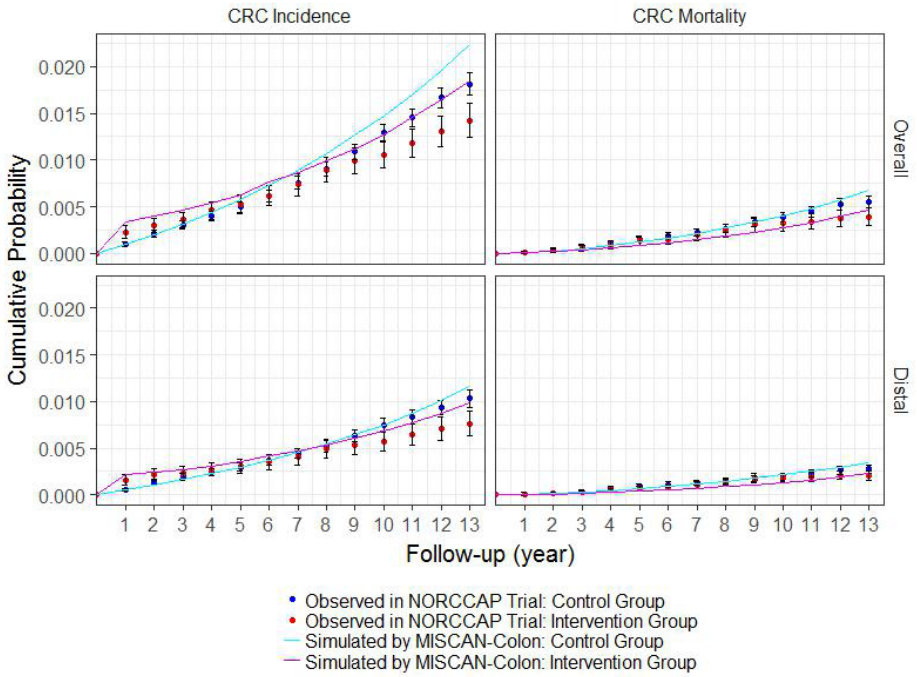
Numbers under control group and screen group presented per 100 000 person years.

*The screening intervention was performed in 1999, 2000 and 2001. Since the closure date for data retrieval is the same for all participants, the number of follow-up years differs among the participants.

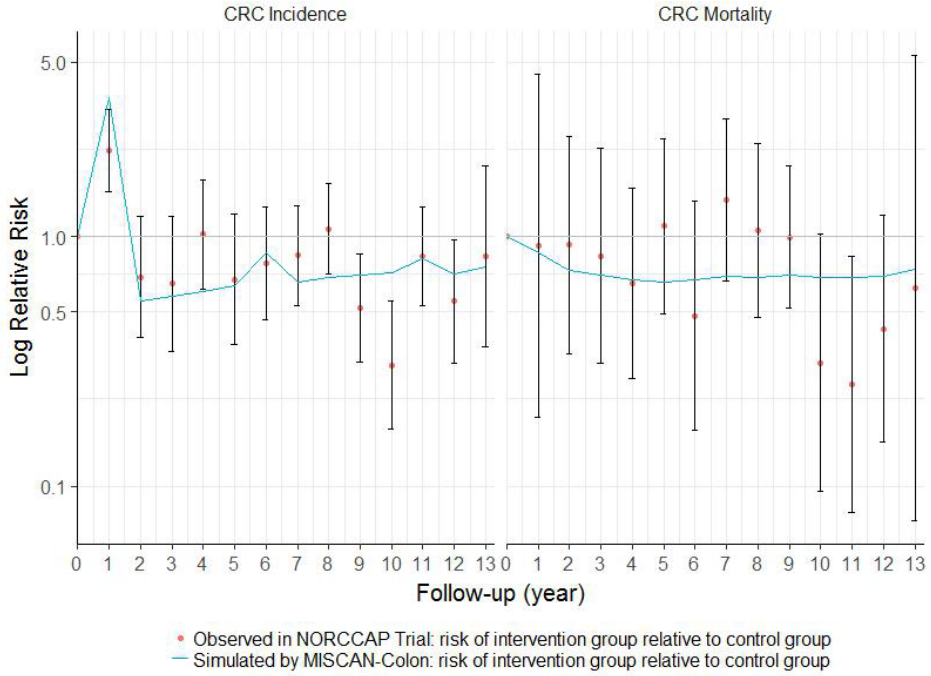
**Last day of the year



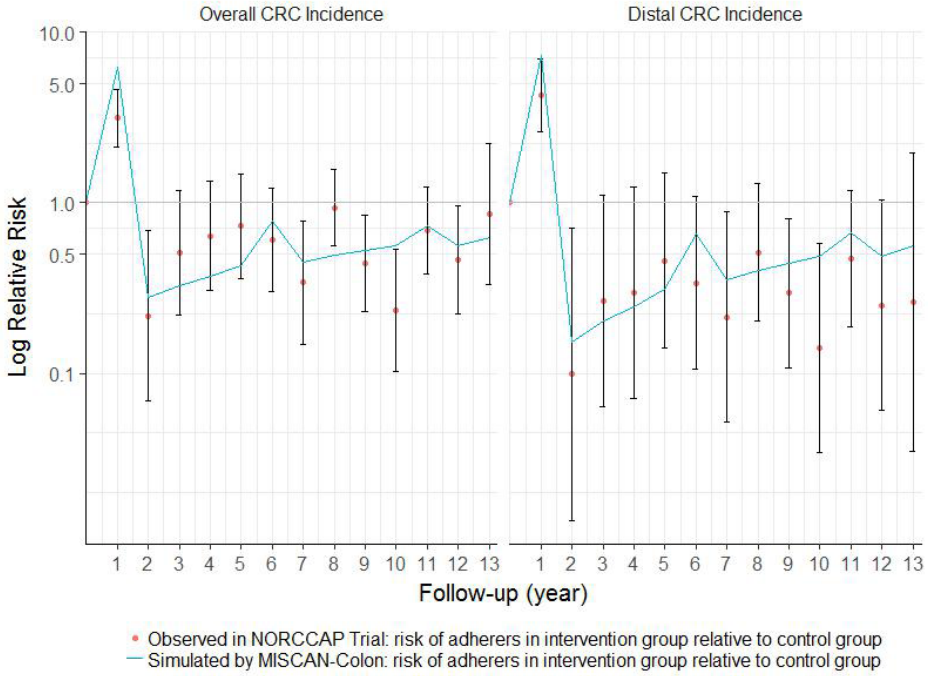
Supplementary Figure 3.1. Hazard ratios: 10- to 12-year follow-up intervention effects of NORCCAP trial including 95% confidence intervals for these effects and MISCAN-Colon predictions of these effects.
Abbreviations: CRC, colorectal cancer; NORCCAP, Norwegian Colorectal Cancer Prevention Trial; MISCAN, Microsimulation Screening Analysis.



Supplementary Figure 3.2. Cumulative probability of overall CRC incidence, overall CRC mortality, distal CRC incidence and distal CRC mortality: 10- to 12-year follow-up intervention effects of NORCCAP trial including 95% confidence intervals for these effects and MISCAN-Colon predictions of these effects. This figure is a replication of Figure 2 as published in Holme et al. (2014).
Abbreviations: CRC, colorectal cancer; NORCCAP, Norwegian Colorectal Cancer Prevention; MISCAN, Microsimulation Screening Analysis.



Supplementary Figure 3.3. Yearly risk ratios for colorectal cancer incidence and mortality in screening group relative to the control group: 10- to 12-year follow-up intervention effects of NORCCAP trial including 95% confidence intervals for these effects and MISCAN-Colon predictions of these effects. This figure is a replication of Figure 3 as published in Holme et al. (2014)
Abbreviations: CRC, colorectal cancer; NORCCAP, Norwegian Colorectal Cancer Prevention Trial; MISCAN, Microsimulation Screening Analysis.



Supplementary Figure 3.4. Yearly risk ratios for overall and distal colorectal cancer incidence in screening adherers relative to control group: 10- to 12-year follow-up intervention effects of NORCCAP trial including 95% confidence intervals for these effects and MISCAN-Colon predictions of these effects. This figure is a replication of Figure 4 as published in Holme et al. (2014)

Abbreviations: CRC, colorectal cancer; NORCCAP, Norwegian Colorectal Cancer Prevention Trial; MISCAN, Microsimulation Screening Analysis.