Supplementary Online Content

Holme Ø, Løberg M, Kalager M, et al. Effect of flexible sigmoidoscopy screening on colorectal cancer incidence and mortality: a randomized clinical trial. *JAMA*. doi:10.1001/jama.2014.8266

This supplementary material has been provided by the authors to give readers additional information about their work.

Supplementary appendix:

1. **Estimation of the per protocol effect**

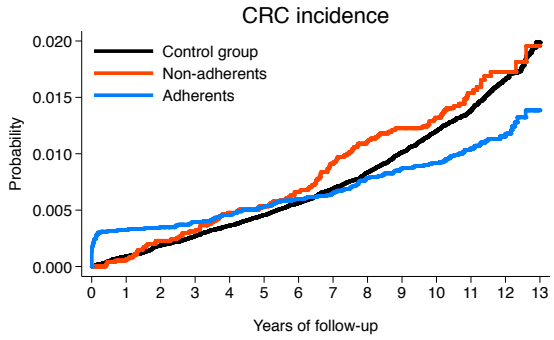
2. **Cost-effectiveness analysis**

Estimation of the per protocol effect

The intention to treat effect quantifies the effect of being assigned to treatment (in our study, screening) in the particular setting in which the randomized trial was conducted. The magnitude of the intention to treat effect is affected by the frequency and type of nonadherence with the assigned screening. The per protocol effect quantifies the effect of receiving the treatment (screening) as specified in the protocol in the particular setting in which the randomized trial was conducted. The per protocol effect does not depend on the observed level of adherence, and therefore it provides information beyond that provided by the intention of treat effect. Unfortunately, valid estimation of the per protocol effect relies on untestable conditions because the adherents are generally not a random sample of those assigned to treatment.

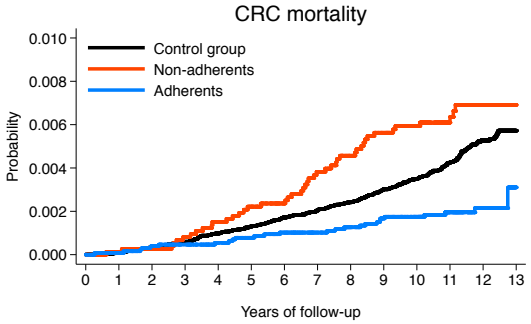
In our study, all 78,220 individuals in the control group are adherents with no screening because of the unavailability of screening outside of the trial. Of 20,572 individuals in the screening group, 12,955 underwent a sigmoidoscopy at baseline (adherents) and 7,617 did not (non-adherents). The CRC incidence rate (per 100,000 person-years) was 97 in the adherents and 144 in the non-adherents, the CRC mortality rate was 18 in the adherents and 58 in the non-adherents, and the overall mortality rate was 634 in the adherents and 1,600 in the non-adherents. eFigure 1 shows the cumulative probability of CRC incidence, CRC mortality, and all-cause mortality in controls, adherents, and non-adherents. eTable 1 shows the distribution of baseline characteristics in adherents and non-adherents.

eFigure 1: Cumulative CRC incidence, CRC mortality and all-cause mortality in the control, non-adherent, and adherent group.



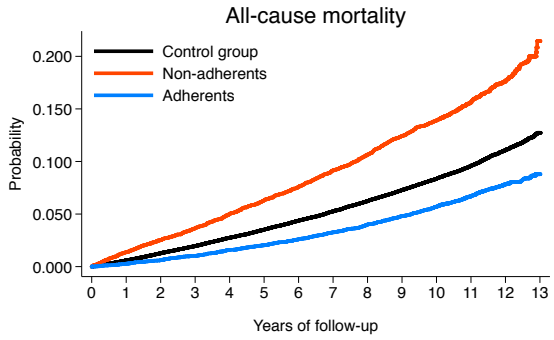
At risk

Adherents	12955	12795	12629	12454	12234	11990	3862
Non-adherents	7617	7346	7102	6852	6574	6308	1723
Control group	78220	76648	75059	73415	71598	69508	17280



At risk

Adherents	12955	12843	12691	12529	12325	12091	3903
Non-adherents	7617	7361	7125	6882	6620	6357	1743
Control group	78220	76777	75272	73722	72044	70127	17520



At risk

Adherents	12955	12843	12691	12529	12325	12091	3903
Non-adherents	7617	7361	7125	6882	6620	6357	1743
Control group	78220	76777	75272	73722	72044	70127	17520

eTable 1: **Summary of follow-up and baseline characteristics of adherents and non-adherents in the screening group.***

	Non-adherents n=7617	Adherents n=12,955	OR[a]	95% CI
Person-years				
For CRC incidence	78,329	143,100		
For CRC mortality	78,688	143,989		
CRC cases	111	142		
CRC deaths	45	26		
Gender				
Female	3644 (48)	6659 (51)	Ref	
Male	3973 (52)	6296 (49)	0.72	0.68-0.77
Age				
55-64y	4806 (35)	8846 (65)	Ref	
50-54y	2811 (41)	4109 (59)	0.70	0.65-0.75
Area of residency				
Oslo	4524 (44)	5734 (56)	Ref	
Telemark	3093 (30)	7221 (70)	1.79	1.68-1.91
Education				
< 9 years (elementary school)	2280(46)	2700 (54)	Ref	
9-12 years (High school)	3413 (34)	6549 (66)	1.49	1.38-1.61
>12 years (college, university)	1477 (30)	3467 (70)	1.88	1.71-2.06
Unknown	447	239		
Income				
Yearly income > NOK 200,000	3592 (31)	7970 (69)	Ref	
Yearly income < NOK 200,000	3953 (45)	4852 (55)	0.65	0.60-0.70
Unknown	72	133		
Marital status				
Unmarried	994 (52)	899 (48)	Ref	
Married	4249 (32)	9097 (68)	1.97	1.78-2.19
Divorced	1987 (46)	2323 (54)	1.25	1.12-1.40
Widow/-er	350 (40)	526 (60)	1.37	1.15-1.62
Unknown	37	110		
Ethnicity				
Norwegian	6243 (35)	11760 (65)	Ref	
Not Norwegian	1373 (55)	1109 (45)	0.66	0.60-0.73
Unknown	1	86		
Comorbidity[b]				
Cancer[c]	198 (39)	307 (61)	0.91	0.75-1.10
Digestive disease/symptoms	250 (38)	400 (62)	1.00	0.84-1.19
Chronic kidney disease	18 (75)	6 (25)	0.29	0.11-0.76
Coronary heart disease	256 (3)	223 (2)	0.61	0.50-0.74
Diabetes	112 (1)	118 (1)	0.87	0.66-1.16
Chronic obstructive pulmonary disease	111 (54)	94 (46)	0.74	0.55-1.00
Public benefits[d]				
Work assessment allowance	1997 (36)	3487 (64)	1.03	0.96-1.11
Sick leave benefit	1214 (33)	2474 (67)	1.09	1.00-1.18
Social assistance benefit	462 (74)	159 (26)	0.48	0.39-0.60
Disability pension	2260 (50)	2302 (50)	0.73	0.67-0.78
Retirement pension	28 (23)	93 (77)	1.38	0.89-2.14

*Numbers in parentheses are percentages

[a]Multivariate odds ratios (OR), adjusted for all other variables in the table.

[b]Comorbidity refers the number of participants in the screening group who were recorded with any of the listed diseases the year prior to the screening examination. The reference is not recorded with the specific disease.

[c]All cancers excluding colorectal or skin cancer

[d] Public benefits refer to the number of participants in the screening group who received any of the listed benefits from the Norwegian Labor and Welfare Administration the year prior to the screening examination. The reference is not receiving the specified benefit.

NOK: Norwegian kroner. 200,000 NOK = 33,000 US dollars.

The table shows that some baseline characteristics (e.g. gender, area of residency, ethnic background, income, education and marital status) are strong predictors of adherence. Valid estimation of the per protocol effect requires adjustment for these and any other prognostic factors that predict adherence. As a result, the per protocol effect cannot be generally estimated via a naïve per protocol analysis that directly compares the outcomes of the adherents and the controls.

Exploiting the exclusion restriction

The condition that all joint predictors of adherence and outcome have been measured and appropriately adjusted for cannot be empirically verified. However, this condition can be tested in our study (because all individuals in the control group are adherents by design) under the plausible assumption that randomization to screening can only affect the outcome through actual screening. This assumption is known as the exclusion restriction.

Therefore we can test whether adherence occurs at random (with respect to the outcomes of interest) in the screening group: were adherence random, then the CRC incidence and the CRC mortality rates among the non-adherents in the screening group would be equal to those in the control group. To test for CRC incidence, we fit a Cox model restricted to non-adherents in the screening group (n=7,617) and all individuals in the control group (n=78,220). The HR (95% CI) of CRC incidence for non-adherents versus controls was 1.02 (0.84-1.24) before and 1.02 (0.83-1.24) after adjustment for the variables in the eTable 1. These estimates suggest that non-adherents and controls (and therefore adherents) have a similar risk of CRC incidence, and therefore that a direct comparison of adherents and non-adherents (or controls) may be appropriate to estimate the per protocol effect.

We fit a similar model for CRC mortality. The HR (95% CI) of CRC mortality for non-adherents versus controls was 1.35 (0.99-1.85) before and 1.30 (0.95-1.80) after adjustment for the variables in eTable 1. These estimates indicate that non-adherents and controls (and therefore adherents) have a different risk of CRC mortality, and that direct adjustment for the measured baseline characteristics is insufficient to eliminate the differences.

Controls and non-adherents had similar CRC incidence, but non-adherents had a higher CRC mortality compared with controls. One reason may be more advanced stages of cancer at diagnosis (Table 4 in the main manuscript). In a Cox model for all-cause mortality, the HR (95% CI) was 1.62 (1.52-1.72) before and 1.35 (1.26-1.43) after adjustment for the covariates in the eTable 1.

The results from the CRC mortality analysis indicate that per protocol effect estimates will be biased when using any method that requires measurement and adjustment for all joint predictors of adherence and outcome. These methods include conventional regression, standardization, inverse probability weighting, propensity score matching, etc. To unbiasedly estimate the per protocol effect of screening on mortality, we need a method that does not require adjustment for joint predictors of adherence and outcome: instrumental variable (IV) estimation.

Instrumental variable estimation

We used the randomization indicator as an IV. The randomization indicator is likely to meet the three instrumental assumptions: it is strongly associated with attendance to screening, we expect no effect of randomization on the outcome except through screening (the exclusion restriction), and we expect the intention to treat effect to unbiasedly estimate the effect of assignment (no confounding for the effect of the instrument). We used the standard IV estimator to estimate the 10-year risk difference that would have been observed between the screening and control groups if individuals in the screening group had actually

undergone sigmoidoscopy. The 10-year cut off was chosen because all individuals were followed for at least 10 years.

We used two-stage least squares (implemented in the ivregress command in STATA) with and without adjustment for the variables in eTable 1. All analyses were adjusted for age group (55-64 years vs 50-54 years). Conservative 95% confidence intervals were calculated using robust variances. For comparison purposes, we also calculated the intention to treat risk difference and the risk difference between the control group and the adherents in the screening group.

Because individuals in the control group were unaware of their status as controls and had essentially no access to screening, the monotonicity condition is expected to hold (no individual could have always done the opposite of what he/she was randomized to, regardless of his/her actual randomized group). Under monotonicity, our approach estimates the "effect in the treated", e.g., the effect among adherents in the screening group. That is, our per protocol effect estimate can be interpreted as the 10-year risk difference that would have been observed if the trial had been restricted to the adherents in the screening group (had we been able to identify them as adherents before randomization). Therefore we need to be cautious when comparing the intention-to-treat and per protocol estimates, as the former apply to the entire trial population and the latter to the adherents only.

Per protocol effect estimates

There were 208 CRC cases in the screening group (n=20,572) and 890 in the control group (n=78,220) after 10 years of follow-up (eTable 2). The intention-to-treat risk difference was -0.22% (95% CI -0.38% to -0.06%). The per protocol risk difference was -0.42% (95% CI -0.69 to -0.15). There were 63 CRC deaths in the screening group and 258 in the control group after 10 years. The intention-to-treat risk difference was -0.06% (95% CI -0.14% to 0.03%) and the per protocol risk difference was -0.10 (95% CI -0.26% to 0.05%).

Thus the per protocol analysis suggests a larger benefit in the adherents regarding both CRC and CRC death. Because the per protocol analysis is restricted to the first 10 years of follow-up for simplicity (all subjects were followed for at least 10 years), the 95% confidence intervals are wider and the point estimates closer to the null than they would be if the entire follow-up of 13 years had been included. Figure 2 in the main manuscript suggests that the effect on CRC mortality is not noticeable until after 9 years of follow up, and that the effect seems to increase with additional follow-up.

There were 1,772 deaths from all causes in the screening group and 6,454 in the control group after 10 years. The intention-to-treat risk difference was -0.22% (95% CI -0.65% to 0.22%) and the per protocol risk difference was -0.08 (95% -0.79% to 0.64%), which does not support an effect of screening on all-cause mortality even in the adherents (eTable 2).

eTable 2: Intention to treat and per protocol risk differences after 10 years of follow-up.

	Screening group (n=20,572)		Control group (n=78,220)	Intention to treat Risk difference, % (95% CI)	Per protocol Risk difference, % (95% CI)		
	Cases Adherents (n=12,955)	Cases Non-adherents (n=7617)	Cases		Unadjusted	IV Age-adjusted	IV Multivariate-adjusted[a]
CRC incidence	116	92	890	-0.22 (-0.38, -0.06)	-0.34 (-0.53, -0.16)	-0.35 (-0.60, -0.10)	-0.42 (-0.69, -0.15)
CRC mortality	22	41	258	-0.06 (-0.14, 0.03)	-0.20 (-0.28, -0.11)	-0.09 (-0.22, 0.05)	-0.10 (-0.25, 0.05)
All-cause mortality	733	1,039	6,454	-0.22 (-0.65, 0.22)	-3.25 (-3.71, -2.80)	-0.35 (-1.03, 0.34)	-0.08 (-0.79, 0.64)

[a] Adjusted for all covariates in eTable 1.

Cost-effectiveness analysis

Health outcome

The health outcomes were measured by means of CRC cases and CRC deaths. Hence, the incremental health effects were prevented CRC cases and CRC deaths, and were estimated by subtracting the expected numbers of CRC cases or deaths in the screening group from the observed numbers in the screening group. The expected numbers of CRC cases and CRC deaths were estimated by multiplying the incidence or mortality rate, respectively, in the control group with the corresponding number of person-years in the screening group. (eTable 3).

eTable 3: Number of prevented CRC cases and CRC deaths in the screening group by age. The rate in the control group is per 100,000 person-years.

	Rate control group	Person-years screening group	Expected number screening group	Observed number screening group	Number of cases prevented
CRC incidence					
50-54 years	83.2	66,660	55	37	18
55-64 years	154.2	128,185	198	171	27
50-64 years	130.3[a]	194,845	254	208	46
CRC mortality					
50-54 years	21.8	66,849	15	12	3
55-64 years	46.6	128,955	60	51	9
50-64 years	38.3[a]	195,804	75	63	12

[a]Age-standardized rates.
Some numbers do not add up due to rounding.

Costs

Costs were measured in a health care perspective. The main source for inputs on unit costs were derived from a report from Tappenden et al[1] who based their estimations on the British National Health Service (NHS) reference costs for 2003. In addition, some cost information was based on expert opinion.[2] The distribution of CRC treatment costs by years after diagnosis were based on Norwegian estimates.[3] The direct cost related to screening included costs of invitation, costs of the screening intervention (FOBT and sigmoidoscopy), follow-up procedures (colonoscopy) and screening complications, see eTable 4 for details.[1]

eTable 4: Direct cost components related to the screening.

Cost category	Cost per unit (US$)[a]	Source
Invitation/reminder letter	3.0	Expert opinion[2]
FOBT in invitation letter	4.4	Expert opinion[2]
FOBT, examination cost	7.4	Expert opinion[2]
Pathology cost (polyps)	64.0	Tappenden 2004[1]
Flexible sigmoidoscopy (with and without therapy)	110.0	Tappenden 2004[1]
Colonoscopy (with and without therapy)	401.8	Tappenden 2004[1]
Admittance for postpolypectomy bleeding	533.6	Tappenden 2004[1]
Treatment for bowel perforation (major surgery)	11,534.6	Tappenden 2004[1]

[a]Adjusted to 2013 level.

We used available lifetime cost for CRC treatment including diagnostics and follow-up by stage of CRC (eTable 5, adapted from Table 22 and 23 in Tappenden et al).[1] Stage of CRC was classified as Dukes A, B, C, or D. We had 68 cases of CRC with unknown Dukes stage. The mortality of these cases was in between the mortality of Dukes C and D (20.7% for Dukes C, 57.4% for Dukes D, and 36.1% for unknown Dukes stage), thus, we assumed that half of these cancers were Dukes C and half Dukes D.

For Dukes A, the treatment costs for asymptomatic (screen-detected) CRC was estimated to be lower than for symptomatic CRC because some Dukes A cancers may be treated endoscopically, thus reducing surgery costs.[1]

eTable 5: Lifetime costs for treatment of CRC according to Dukes' stage.[1]. Proportion of colorectal cancers according to stage and mode of detection (%).

Stage	Lifetime cost (US$)[a]	Proportion control	Proportion 50-54 years age group	Proportion 55-64 years age group
Screen-detected Dukes A	15,465	-	13.5	11.1
Screen-detected Dukes B	26,537	-	5.4	1.8
Screen-detected Dukes C	40,694	-	2.7	4.7
Screen-detected Dukes D	25,481	-	2.7	1.2
Symptomatic Dukes A	18,342	20.1	10.8	14.6
Symptomatic Dukes B	26,537	22.7	10.8	15.2
Symptomatic Dukes C	40,694	29.8	18.9	25.1
Symptomatic Dukes D	25,481	27.4	35.1	26.3

[a]Adjusted to 2013 level.

Our analysis applies a 10 years perspective, thus the lifetime costs were split into annual costs from the time of diagnosis. We used the distribution of CRC treatment costs over years after diagnosis that was observed in a previous economic report of the NORCCAP trial.[3] In this report, it was shown that the majority of costs could be attributed to the first years after diagnosis. The cumulative proportion of costs of treatment in the first five years of diagnosis is shown in eTable5. Total costs were calculated per person in both the screening group and control group by adding all cost components.

eTable 6: Proportion of costs according to years after diagnosis for treatment of CRC according to Dukes' stage.[3]

Year after diagnosis		1	2	3	4	5
Dukes Stage	A	0.82	0.86	0.98	0.99	1
	B	0.77	0.90	0.96	1	1
	C	0.65	0.82	0.91	1	1
	D	0.85	0.99	1	1	1

Estimating cost-effectiveness

We calculated the incremental cost-effectiveness ratio (ICER) in a ten-year perspective for the 50-54 and 55-65 year age-groups separately, and for the whole cohort (50-64 years), with once only flexible sigmoidoscopy screening and 50% of participants also invited for FOBT. The incremental health effect was measured by prevented CRC-cases or CRC-deaths, while the incremental costs included costs related to screening, treatment, and follow-up. The comparison was no screening (the control group) for the corresponding age-group. According to current guidelines, all costs were discounted at 3.5% per year.[4] All costs were converted to US dollars by using the average exchange rate for 2003 and adjusted to 2013 level using the UK Consumer Price Index.[5]

Overall, the incremental cost of preventing one CRC in the 50-64 year age group was $58,448, and $226,002 for preventing one death from CRC (eTable 7).

eTable 7: Incremental cost per prevented CRC and death from CRC in the NORCCAP trial by age group.

Group	Prevented CRC cases	Prevented CRC deaths	Incremental costs*	ICER: CRC – incidence	ICER: CRC death
50-54 years	18	3	826,164	44,727	321,784
55-64 years	27	9	1,884,236	67,532	199,912
50-64 years	46	12	2,710,400	58,448	226,002

** Incremental cost is the average difference in cost per person (standardized by sex and age-group) between the screening group and the control group multiplied by the number of persons in the screening group. Costs are in 2013 US$. ICER: Incremental cost-effectiveness ratio. Some numbers do not add up due to rounding.*

We found that the cost per prevented CRC-case was lower in the 50-54 year age group compared with the 55-64 year age group. This is mainly explained by the larger reduction in CRC incidence in the 50-54 year age group, HR 0.67 (95% CI 0.47-0.94), than in the 55-64 year age group, HR 0.86 (95% CI 0.73-1.03). Additionally, the screening costs per person were higher in the 55-64 year age group due to more findings of polyps at screening flexible sigmoidoscopy with subsequent colonoscopy. The result from this subgroup analysis should however be interpreted with caution as there is no statistically significant difference in the risk of CRC between the two age groups.

As our results are restricted to 10 years of follow-up, the ICER estimates may change with further follow-up. As shown in Figure 3A of the paper, the yearly risk of CRC in the screening group is reduced beyond 10 years of follow-up; implying that costs of treatment will continue to favor the screening group also after 10 years; at least for the 13 years observed in our trial.

Further, these data are based on treatments given in the years around 2003. During the last years, more expensive treatment is offered, especially with regard to chemotherapy, with only minor effect on life expectancy. These changes in treatment would only reduce the ICERs presented in eTable 7, making screening more favorable.

eReferences

1. Tappenden P, Eggington S, Nixon R, et al. Colorectal cancer screening options appraisal: cost-effectiveness, cost-utility and resource impact of alternative screening options for colorectal cancer. Report to the English Bowel Cancer Screening Working Group, September 2004. Available at: http://www.cancerscreening.nhs.uk/bowel/scharr.pdf. Accessed June 2nd, 2014.
2. Aas E. Cost-effectiveness of screening for colorectal cancer with once-only flexible sigmoidoscopy and faecal occult blood test. University of Oslo. 2008; Available at: http://www.med.uio.no/helsam/forskning/nettverk/hero/publikasjoner/skriftserie/2008/2008_6.pdf. Accessed June 2nd, 2014.
3. Aas E. The effect of screening on treatment cost: The case of colorectal cancer. University of Oslo. 2008; Available at: http://www.med.uio.no/helsam/forskning/nettverk/hero/publikasjoner/skriftserie/2008/2008_5.pdf. Accessed June 2nd, 2014.
4. National Institute for Health and Clinical Excellence. Available at: http://www.nice.org.uk/media/68D/29/The_guidelines_manual_2009_-_Chapter_7_Assessing_cost_effectiveness.pdf. Accessed June 2nd, 2014.
5. Consumer Price Index UK. Available at: http://www.rateinflation.com/consumer-price-index/uk-historical-cpi?start-year=2003&end-year=2014. Accessed June 9th, 2014.