Published OnlineFirst May 17, 2017; DOI: 10.1158/1055-9965.EPI-16-0786

Research Article

Cancer Epidemiology, Biomarkers & Prevention

Do Men and Women Need to Be Screened Differently with Fecal Immunochemical Testing? A Cost-Effectiveness Analysis

Miriam P. van der Meulen[1], Atija Kapidzic[2], Monique E. van Leerdam[3], Alex van der Steen[1], Ernst J. Kuipers[2,4], Manon C.W. Spaander[2], Harry J. de Koning[1], Lieke Hol[2], and Iris Lansdorp-Vogelaar[1]

Abstract

Background: Several studies suggest that test characteristics for the fecal immunochemical test (FIT) differ by gender, triggering a debate on whether men and women should be screened differently. We used the microsimulation model MISCAN-Colon to evaluate whether screening stratified by gender is cost-effective.

Methods: We estimated gender-specific FIT characteristics based on first-round positivity and detection rates observed in a FIT screening pilot (CORERO-1). Subsequently, we used the model to estimate harms, benefits, and costs of 480 gender-specific FIT screening strategies and compared them with uniform screening.

Results: Biennial FIT screening from ages 50 to 75 was less effective in women than men [35.7 vs. 49.0 quality-adjusted life years (QALY) gained, respectively] at higher costs (€42,161 vs. −€5,471, respectively). However, the incremental QALYs gained and costs of annual screening compared with biennial screening were more similar for both genders (8.7 QALYs gained and €26,394 for women vs. 6.7 QALYs gained and €20,863 for men). Considering all evaluated screening strategies, optimal gender-based screening yielded at most 7% more QALYs gained than optimal uniform screening and even resulted in equal costs and QALYs gained from a willingness-to-pay threshold of €1,300.

Conclusions: FIT screening is less effective in women, but the incremental cost-effectiveness is similar in men and women. Consequently, screening stratified by gender is not more cost-effective than uniform FIT screening.

Impact: Our conclusions support the current policy of uniform FIT screening. *Cancer Epidemiol Biomarkers Prev; 26(8); 1328–36.* *©2017 AACR.*

Introduction

Colorectal cancer is the second most common cause of cancer-related mortality in the Western world (1). Screening can prevent part of these deaths by early detection and treatment of colorectal cancer and its precursor lesions. Consequently, several countries and local initiatives across the world have adopted population-based screening for colorectal cancer. The majority of these initiatives have opted for some form of fecal immunochemical testing (FIT; refs. 2, 3).

These screening programs use the same approach for both genders despite age and gender disparities in prevalence of colorectal neoplasia and a higher life expectancy in women. We previously showed that a uniform approach is cost-effective for primary colonoscopy screening, because the lower prevalence of advanced neoplasia in women is compensated by a higher life expectancy (4). However, we assumed that test characteristics for colonoscopy did not differ between genders, while there are strong indications that the test characteristics for fecal occult blood test (FOBT; including FIT) differ between men and women.

A Scottish guaiac-based FOBT (gFOBT) screening study reported more screen-detected colorectal cancers in men compared with women, whereas the number of interval colorectal cancers was similar in both groups, suggesting higher sensitivity in men (5). Two other studies likewise found a higher sensitivity of FIT for advanced neoplasia in men compared with women (6, 7), and also a higher positive predictive value (PPV), whereas FIT specificity in men was found to be significantly lower (6). The lower specificity in men was confirmed in the FIT screening trial CORERO-1 with a higher FPR in men compared with women (8).

These studies triggered a debate whether men and women should be screened differently with FIT. For instance, lowering the FIT cutoff in women will increase their sensitivity toward that of men, or in contrast, increasing the cutoff in women will increase their PPV toward that of men (9). Differences in test characteristics might affect the optimal cutoff for a positive FIT, but might also affect the optimal screening age range and interval. Microsimulation modeling can take these gender differences in test characteristics, but also in life expectancy and colorectal cancer incidence into account and estimate costs and quality-adjusted life years (QALY) gained of various screening strategies. In this study, the

[1]Department of Public Health, Erasmus Medical Centre, Rotterdam, the Netherlands. [2]Department of Gastroenterology and Hepatology, Erasmus Medical Centre, Rotterdam, the Netherlands. [3]Department of Medical Oncology, The Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital, Amsterdam, the Netherlands. [4]Department of Internal Medicine, Erasmus Medical Centre, Rotterdam, the Netherlands.

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (http://cebp.aacrjournals.org/).

Corresponding Author: Miriam P. van der Meulen, Department of Public Health, Erasmus MC, PO Box 2040, 3000 CA Rotterdam 3015 CE, the Netherlands. Phone: 311-0703-8459; Fax: 311-0703-8475; E-mail: m.vandermeulen.1@erasmusmc.nl

doi: 10.1158/1055-9965.EPI-16-0786

Published OnlineFirst May 17, 2017; DOI: 10.1158/1055-9965.EPI-16-0786

microsimulation model MISCAN-Colon was used to determine optimal screening strategies for men and women and to study whether screening stratified by gender is beneficial in terms of cost-effectiveness.

Materials and Methods

We developed two separate versions of the microsimulation model MISCAN-Colon for men and women. We estimated sensitivity and specificity of FIT based on the gender-specific positivity and detection rates observed in the CORERO-1 trial. We then simulated male and female populations screened with various FIT screening strategies to estimate the QALYs and costs of FIT screening by gender and determined efficient screening strategies for men and women. Thereafter, we compared costs and effects of screening stratified by gender with uniform screening.

The CORERO trial

The CORERO-1 trial was a randomized controlled trial comparing attendance and detection rates of gFOBT, FIT, and sigmoidoscopy at first round screening. For the current study, we only used the data of FIT screening. Details from this trial have been described elsewhere (10, 11). In short, screening-naïve subjects ages 50 to 74 years, living in the southwest of the Netherlands were selected through municipal population registers. Screenees assigned in the FIT study arm received a kit with a single FIT (OC-Sensor). A cutoff of 10 μg hemoglobin/g feces (equivalent to 50 ng hemoglobin/mL) was used to indicate a positive test result. This was followed by the recommendation for a diagnostic colonoscopy. In total, 4,969 men and 5,039 women were invited, of which 59.8% of men and 64.6% of women returned the test. The positivity rate and detection rates were higher among men compared with women (Table 1). The PPV for advanced neoplasia did not differ significantly for men (42.1%) and women (37.0%; $P = 0.265$). Positivity rates, detection rates, and the PPV at higher cutoffs can be found elsewhere (8).

MISCAN-Colon

MISCAN-Colon is a well-established microsimulation model for colorectal cancer developed at Erasmus MC (Rotterdam, the Netherlands). The model has been extensively described previously (12, 13) and in the Supplementary Model Methods S1. In brief, MISCAN-Colon simulates life histories of a large group of individuals from birth to death. As each simulated person ages, one or more adenomas may develop. These adenomas can progress in size from small (≤5 mm) to medium (6–9 mm) to large size (≥10 mm). Some adenomas can develop into preclinical cancer, which may progress through stages I to IV.

At any time during the development of the disease, the process may be interrupted because a person dies of other causes. With screening, colorectal cancer may be prevented by the detection and removal of adenomas or detected at an earlier stage with a more favorable survival. In this way, colorectal cancer incidence and/or colorectal cancer-related mortality can be reduced. The life years gained by screening are calculated as the difference in model-predicted life years lived in the population with and without colorectal cancer screening.

Model input

Natural history. We developed two versions of the MISCAN-Colon model, one for each gender. The two versions were separately calibrated to gender-specific prescreening data on the age-specific incidence of colorectal cancer as observed in the Netherlands before the introduction of screening (between 1999 and 2003; ref. 14) and the gender and age-specific prevalence and multiplicity distribution of adenomas as observed in autopsy studies (15–22). The size distribution of adenomas was calibrated to the size distribution of adenomas detected in a colonoscopy trial (23). Survival after clinical diagnosis is based on 1989 to 2003 survival data obtained from the Dutch Comprehensive Cancer Centre (14). The preclinical duration of colorectal cancer and the adenoma dwell-time were calibrated to the rates of interval and surveillance-detected cancers observed in randomized controlled trials evaluating screening using gFOBTs and a once-only sigmoidoscopy (24). The model outcomes showed good concordance with trial results (Supplementary Model Methods).

Study population. We modeled the age distribution and life expectancy separately for the male and female version of the model. We modeled the Dutch population aged 25 to 85 years in 2015 (25), and all individuals were followed until death. Life expectancy was based on gender-specific life-tables from 2011 obtained from Statistics Netherlands (25).

Screening strategies. FIT screening was simulated in the population starting in year 2015. Individuals were offered screening according to different FIT screening schedules varying by:

- Age to start screening: 40, 45, 50, 55, 60, and 65 years
- Age to stop screening: 70, 75, 80, and 85 years
- Screening interval: 1, 1.5, 2, and 3 years

Table 1. Positivity rates and detection rates FIT with a cutoff of 10 μg Hb/g feces as simulated with equal and gender-specific test characteristics and as observed in CORERO-1

	Positivity rate	Detection rate of nonadvanced adenomas	Detection rate of advanced neoplasia[a]
Men			
Observed ($n = 2,857$)	10.71%	2.56%	4.38%
Simulated with equal FIT characteristics	8.60%	1.80%	3.48%
Simulated with gender-specific FIT characteristics	10.75%	2.55%	4.38%
Women			
Observed ($n = 3,129$)	6.30%	0.86%	2.17%
Simulated with equal FIT characteristics	7.89%	1.50%	2.72%
Simulated with gender-specific FIT characteristics	6.29%	0.86%	2.17%

NOTE: The model simulated with equal FIT characteristics had a GOF of 56.3, compared with a GOF of 0.0008 in the model with gender-specific FIT characteristics, a difference of 56.3.

[a]An advanced adenoma was defined as an adenoma of 10 mm or greater in size, and/or with 25% or greater villous component and/or high-grade dysplasia.

Published OnlineFirst May 17, 2017; DOI: 10.1158/1055-9965.EPI-16-0786

The cut-off level for a positive FIT result varied between 10 (FIT^{10}), 15, 20, 30, and 40 μg hemoglobin/g feces (equivalent to 50–200 ng Hb/mL). This resulted in a total of 480 different screening strategies per gender.

If adenomas were detected, individuals entered a surveillance program according to the Dutch guidelines for follow-up after polypectomy (26). We assumed that surveillance colonoscopies would be performed until at least 75 years of age, or until the stop age for screening, whichever was latest. If no adenomas were found at diagnostic colonoscopy, the individuals were assumed to be at low-risk for colorectal cancer and did not return to the screening program until after 10 years.

Attendance. To identify the optimal screening strategies for participants, we analyzed the strategies with full attendance (100%). In the sensitivity analysis, we looked at alternative gender-specific attendance levels based on the CORERO-1 trial (see Supplementary Table S1).

Costs. The analysis was performed from a third-party payer perspective. All costs are presented in Supplementary Table S1. We adjusted all costs to reflect the 2012 level, using the Dutch Consumer Price Index (27).

FIT costs were assumed to be €21.90 based on an internal study (including a single KIT, packing material, material and personnel costs of the analysis, postage costs, and organizational costs). The assumed costs of a colonoscopy were based on estimates in the COCOS trial: €192 for a negative colonoscopy and €329 for a colonoscopy with polypectomy (28). Because of the recent discussion on colonoscopy costs in the United States (29), we considered costs that were twice and four times as high in a sensitivity analysis. Costs for colonoscopy complications were based on Diagnosis Treatment Combination (DTC) rates, derived from the Dutch Health Care Authority (30). Costs for treatment of colorectal cancer were divided into three clinically relevant phases of care: initial treatment, continuous care, and terminal care. The initial care phase was defined as the first 12 months after diagnosis, the terminal care phase was defined as the final 12 months of life, and the continuing care phase was defined as all months in between. For patients surviving less than 24 months, the final 12 months were allocated to the terminal care phase, and the remaining months were allocated to the initial care phase. Initial treatment costs were based on DTC rates, except for oxaliplatin. The costs for oxaliplatin were derived from the Dutch Health Care Insurance Board (31). We assumed that during the continuous care phase, individuals would follow the Dutch colorectal cancer treatment guidelines (32), and costs for periodic control were based on DTC rates. Terminal care costs were based on a Dutch last-year-of-life-cost-analysis (33). We assumed that these costs increased with stage at diagnosis, at a rate observed for U.S. patients (34, 35). Dutch terminal care costs for individuals who died from colorectal cancer were approximately 40% of the U.S. costs. We further assumed that terminal care costs of colorectal cancer patients who die from other causes were also 40% of the U.S. costs.

Utility losses. We assumed no utility loss for a FIT, a utility loss equal to 2 days per colonoscopy (0.0055 QALYs), and two weeks of life per complication (0.0384 QALYs). We also assigned a utility loss to each life-year with colorectal cancer care (Supplementary Table S1; ref. 36).

Analysis

Estimating FIT sensitivity and specificity. The sensitivity and specificity of FIT were fitted to the gender-specific positivity and detection rates observed in the first round of the CORERO-1 trial. FIT sensitivity and specificity were estimated by minimizing the difference between observed and expected (i.e., model simulated) trial outcomes. Trial outcomes used for estimation were (i) positivity rate (PR) and detection rate (DR) of (ii) colorectal cancer, (iii) advanced adenomas, and (iv) nonadvanced adenomas for both men and women, for a total of 8 trial outcomes. The observed detection rate of advanced adenomas was fitted to the detection rate of large (i.e., ≥10 mm) adenomas in the model, as the model does not incorporate histology.

We estimated FIT characteristics twice: once assuming equal and once assuming gender-specific sensitivity and specificity. If the goodness-of-fit (GOF) of the model with gender-specific FIT characteristics was significantly better than the model with equal characteristics, we assumed that FIT characteristics indeed differed between men and women. The GOF was calculated as the sum of deviances between observed and simulated outcomes using the following formula:

$$2*\left[obs*\left(ln\left(\frac{obs}{n}\right)-ln\left(\frac{sim}{m}\right)\right)\right]+2*\left[(n-obs)*\left(ln\left(\frac{n-obs}{n}\right)-ln\left(\frac{m-sim}{m}\right)\right)\right]$$

FIT characteristics differed significantly between men and women if the difference in GOF of the model with gender-specific FIT characteristics and the model assuming equal FIT characteristics exceeded 7.815 (χ^2 distributed, three degrees of freedom).

Costs and benefits of uniform and gender-based screening. We used the MISCAN model to calculate costs and benefits of all 480 different screening strategies by gender, including no screening. Costs and QALYs gained were discounted by 3% per year (37). For reference, we first compared outcomes of no screening, biennial screening from age 50 to 75 and annual screening from age 50 to 75 for men and women.

Subsequently, we used incremental cost-effectiveness analysis to determine the cost-effective screening strategies among all 480 evaluated screening by gender. To obtain these strategies, we ruled out strategies that were more costly and less effective than other strategies (simple dominance) or combinations of other strategies (extended dominance). The remaining strategies are known as cost-effective or "efficient." On a plot of QALYs gained versus costs, the line connecting efficient strategies is called the efficient frontier. We calculated the incremental cost-effectiveness ratio (ICER) of each efficient strategy by comparing its costs and effects with those of the next less costly and less effective efficient strategy.

Finally, to determine the benefit of screening stratified by gender on a population level, we combined the efficient strategies of men with the efficient strategies of women, thereby creating the gender-based screening strategies. The costs and QALYs gained for men and women were summed, based on the distribution of men and women in the population. Then, the efficient gender-based screening strategies were determined and compared with the efficient strategies of uniform screening in the total population. We considered a difference in benefit between gender-based and uniform screening of ≥10% significant (38).

Published OnlineFirst May 17, 2017; DOI: 10.1158/1055-9965.EPI-16-0786

Table 2. Outcomes of an annual and biennial screening program with FIT with a cutoff of 10 μg Hb/g feces and gender-specific FIT characteristics, screening from age of 50–75 years per 1,000 participants (100% attendance)

	CRC incidence	CRC deaths	QALYs gained[a]	Total screening costs (€)[a,b]	Treatment costs (€)[a]	Total costs (€)[a]
Men						
No screen	37.9	20.7		0	499,783	499,783
Biennial screening	23.3	9.6	49	136,267	358,045	494,312
Annual screening	20.1	8.1	56	198,137	317,038	515,175
Women						
No screen	31.9	18.5		0	420,600	420,600
Biennial screening	23.4	10.6	36	114,881	347,880	462,761
Annual screening	19.8	8.6	44	182,660	306,495	489,155

[a]3% discounted.
[b]Including primary test, diagnostic colonoscopy, surveillance, and complications.

Sensitivity analysis

We performed seven sensitivity analyses on different test characteristics of FIT: (i) we assumed only specificity differed between men and women; (ii) we assumed only sensitivity differed; (iii) we assumed no difference in sensitivity and specificity; (iv) we assumed a difference in sensitivity of colorectal cancer similar to the difference in sensitivity of advanced adenomas; (v) we assumed that sensitivity of FIT in women is primarily lower for progressive adenomas and to a lesser extent for nonprogressive adenomas; (vi) we assumed that a percentage of adenomas do not bleed and can therefore never be detected by FIT, unless they grow; and (vii) a similar analysis where we assumed that this percentage was higher in women than in men (Supplementary Table S1).

We also performed sensitivity analyses on differential attendance for men and women, colonoscopy costs, treatment costs, discounting rates, and including societal costs (Supplementary Table S1).

Results

FIT characteristics

Assuming equal FIT characteristics for both sexes, the simulated PR and DR at a cutoff of 10 μg hemoglobin/g feces (50 ng Hb/mL) were higher in men than in women, due to a higher prevalence of colorectal neoplasia in men. Under this assumption, the simulated PR and DR in men were lower and in women higher than the observed rates (Table 1). Allowing FIT characteristics to vary by sex significantly improved the FIT of the model to the observed CORERO-1 trial rates (Table 1). FIT specificity needed to be lower and sensitivity for (non)advanced adenomas needed to be higher in men than in women to replicate the observed FIT positivity and detection rates by gender, whereas the sensitivity for colorectal cancer was similar in both sexes (Supplementary Table S2).

Screening outcomes by gender

Using the model with gender-specific test characteristics, MISCAN predicted that biennial FIT[10] screening between 50 and 75 years led to more profound reduction in colorectal cancer incidence and mortality compared with no screening in men than in women (Table 2). Women had less life years and QALYs gained than men per 1,000 participants (35.7 vs. 49.0 QALYs gained, respectively), at higher costs (€42,161 vs. −€5,471, respectively). Annual screening also yielded fewer QALYs gained (44.4 vs. 55.7) and higher costs (€68,555 vs. €15,391) in women than men when compared with no screening. However, the incremental QALYs gained and costs for annual screening compared with biennial screening were more similar between both sexes with 8.7 QALYs gained and €26,394 for women versus 6.7 QALYs gained and €20,863 for men.

When all strategies were considered (also varying screening age range and interval), costs remained higher and QALYs gained lower in women compared with men for all strategies (Fig. 1). There was considerable overlap in which strategies were efficient between men and women, as six efficient screening strategies were identical (Table 3).

Benefit of gender-based screening

Supplementary Table S3 shows all efficient screening strategies stratified by gender and efficient uniform screening strategies. Six of these strategies included an identical screening strategy for men and women. Table 4 shows an example of uniform screening strategies and screening strategies stratified by gender at different willingness-to-pay thresholds. At a willingness-to-pay threshold of €0, screening stratified by gender consisted of screening men only. At a willingness-to-pay threshold of €20,000, the most effective strategy was equal in men and women; thus, there was no difference between gender-stratified and uniform screening. The costs and QALYs gained of all efficient strategies are shown in Fig. 2. For screening strategies with few screening rounds, screening stratified by gender dominated uniform screening, albeit the difference was small. The widest gap in QALYs gained between uniform screening and screening stratified by gender was at savings of €16,867: screening both men and women ages 60 to 70 years triennially gained less QALYs (54) than screening men ages 60 to 70 years biennially and screening women ages 60 to 70 years triennially (58; Supplementary Table S3), a difference of 7%. From willingness-to-pay thresholds of €1,300 or higher, there was no difference between screening stratified by gender and uniform screening.

Sensitivity analysis

The performed sensitivity analyses resulted in different strategies to be on the efficient frontier. However, in all sensitivity analyses, the added value of screening stratified by gender compared with uniform screening was marginal (Supplementary Fig. S1). At a willingness-to-pay threshold of €20,000 per QALY gained, there was no difference between uniform and stratified screening when assuming differential attendance levels for men and women. At this threshold level, the difference in QALY gained

Published OnlineFirst May 17, 2017; DOI: 10.1158/1055-9965.EPI-16-0786

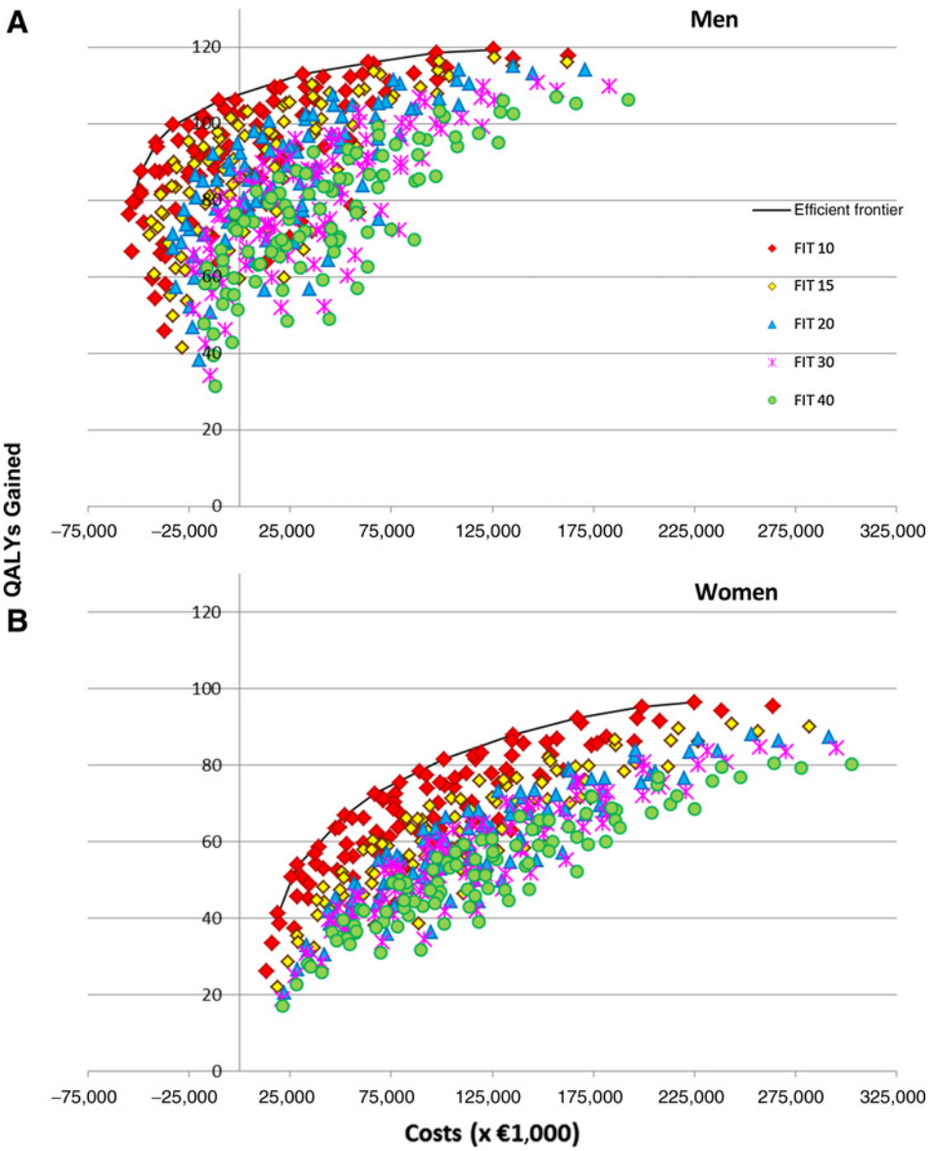

Figure 1.
Costs and QALYs gained per 1,000 participants of FIT screening with five different cutoffs and with different start and stop age and screening interval, 3% discounted.

between uniform screening and screening stratified by gender was highest when assuming quadruple colonoscopy costs, but did not exceed 3 QALY per 1,000 participants (approximately 3%; see Supplementary Figure S1, Panel B).

Discussion

Our study demonstrates that FIT screening is more (cost-) effective in men than in women due to a higher prevalence of colorectal neoplasia and a better test sensitivity for (advanced) adenomas in men. Nevertheless, screening women remained highly cost-effective compared with no screening. Despite the difference in cost-effectiveness compared with no screening, the ICER of different screening strategies did not differ substantially between men and women, and the optimal screening strategies for men and women were either the same or very similar. As a result, FIT screening stratified by gender dominated uniform screening with less intensive screening (maximum difference 58 vs. 54 QALYs gained, respectively), but resulted in equal costs and QALYs gained from a willingness-to-pay threshold of €1,300. Thus, FIT screening stratified by gender was not more cost-effective than uniform FIT screening.

Given the differences in costs and QALYs gained compared with no screening between men and women, it may come as a surprise that the efficient strategies and incremental cost-effectiveness ratios are quite similar between sexes. Cost-effectiveness of intensifying screening is however determined by the yield of the additional screening rounds. The prevalence of (advanced) neoplasia in the screened population decreases each screening round, depending on FIT sensitivity. At the first screening round, men

Published OnlineFirst May 17, 2017; DOI: 10.1158/1055-9965.EPI-16-0786

Table 3. Screening strategies on the cost efficiency frontier in (i) men and (ii) women, 3% discounted

	Cutoff	Start age	Stop age	Interval	# Screens	Costs[a]	QALY gained[a]	Costs (€)/QALY gained	ICER
Men									
1	FIT10	60	70	2	6	−54,815	76	−719	−719
2	FIT10	55	70	2	8	−48,971	87	−560	524
3	FIT10	55	70	1.5	11	−41,287	95	−434	998
4	FIT10	55	75	1.5	14	−32,956	100	−331	1,859
5	FIT10	50	75	1.5	17	−10,543	106	−99	3,525
6	FIT10	50	75	1	26	31,175	113	276	6,006
7	FIT10	50	80	1	31	63,867	116	551	10,809
8	FIT10	45	80	1	36	97,580	118	824	13,285
9	FIT10	40	80	1	41	125,815	119	1,054	33,234
Women									
1	FIT10	60	70	3	4	18,948	41	462	462
2	FIT10	60	70	2	6	25,890	51	512	730
3	FIT10	60	70	1.5	7	28,696	54	533	860
4	FIT10	55	70	1.5	11	52,302	67	783	1,818
5	FIT10	55	75	1.5	14	66,794	72	921	2,549
6	FIT10	55	75	1	21	101,147	81	1,241	3,823
7	FIT10	50	75	1	26	135,405	88	1,543	5,449
8	FIT10	50	80	1	31	167,127	92	1,812	7,131
9	FIT10	45	80	1	36	199,055	95	2,092	10,818
10	FIT10	40	80	1	41	225,426	96	2,341	23,267

[a]Per 1,000 participants.

have a higher prevalence of (advanced) neoplasia than women, but the prevalence of advanced neoplasia in men after one screening round will become lower than in unscreened women. As a consequence, the yield of initiating screening in women is higher than the yield of intensifying screening in men. This effect is demonstrated by a lower ICER of the first efficient strategy of women than the ICER of the second efficient strategy of men. The yield of further intensifying FIT screening depends on the residual number of nondetected neoplasia. The lower sensitivity of FIT in women compared with men necessitates more frequent screening in women than men, while the lower initial prevalence of neoplasia compensates this, leading to similar efficient strategies with a similar ICER.

We have modeled the differential performance of FIT between men and women as a difference in sensitivity for (advanced) adenomas. This does not necessarily mean that FIT is less accurate in women. Rather, adenomas in women are less likely to give blood in stool, and therefore, FIT is not able to detect these adenomas, resulting in a lower sensitivity of the test for adenomas. One explanation for the differential performance of FIT is the fact that a greater proportion of adenomas in men are generally located in the left hemicolon. Because this could influence results if the (missed) right-sided lesions progress more rapidly, we added a sensitivity analysis in which the difference in FIT performance between men and women primarily existed for progressive (i.e., faster growing) adenomas rather than for nonprogressive (slow growing) adenomas. The conclusion of this sensitivity analysis was in line with the base case analysis. Potential other reasons for the differential performance of FIT are gender differences in hemoglobin concentration of blood, fecal volume, and a lower colonic transit in women than men (5).

Our finding of lower test sensitivity of FIT in women is in concordance with two other studies (6, 7), but in contrast with one other study (39). Even though our sensitivity estimates are based on a single study, we are confident that this does not influence our results. We performed extensive sensitivity analyses on test characteristics and found our results to be robust for these assumptions. Also, a German study found a per-person sensitivity

Table 4. Example of the most effective uniform screening strategies and screening strategies stratified by gender at different willingness-to-pay thresholds, 3% discounted

Willingness to pay		Cutoff	Start age		Stop age		Interval		# Screens		Costs[a]	QALY gained[a]
			M	W	M	W	M	W	M	W		
€0	By gender	FIT10	60	X	70	X	2	X	6	X	−27,062	38
	Uniform	FIT10	60		70		3		4		−16,867	54
€2,000	By gender	FIT10	55		75	70	1.5		14	11	10,209	83
	Uniform	FIT10	55		70		1.5		11		6,096	81
€5,000	By gender	FIT10	50	55	75	75	1.5	1	17	21	46,005	94
	Uniform	FIT10	50		75		1.5		17		40,020	92
€10,000	By gender	FIT10	50		75	80	1		26	31	100,007	102
	Uniform	FIT10	50		80		1		31		116,147	104
€20,000	By gender	FIT10	45		80		1		36		148,956	107
	Uniform	FIT10	45		80		1		36		148,956	107

[a]Per 1,000 participants.

Published OnlineFirst May 17, 2017; DOI: 10.1158/1055-9965.EPI-16-0786

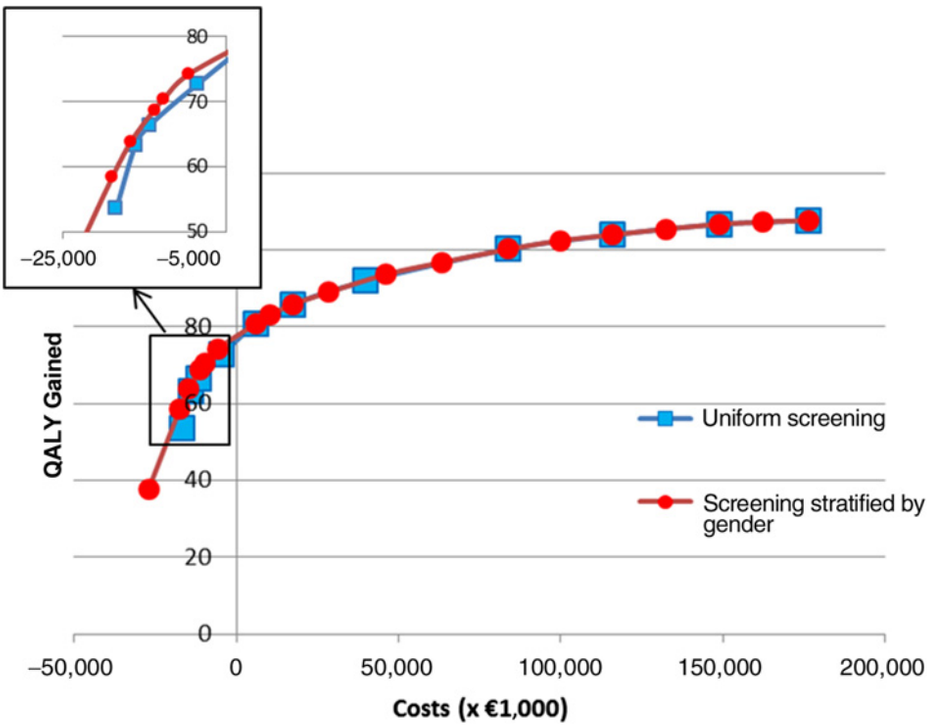

Figure 2.
Costs and QALY gained of strategies on the cost-efficiency frontier per 1,000 participants with uniform screening and screening stratified by gender, 3% discounted.

for advanced neoplasia of 30.7% for women compared with 47.7% for men (6). Our sensitivity estimates concern a per-lesion instead of per-person sensitivity. Because of multiple lesions and a probability for a positive FIT for other reasons than colorectal neoplasia (e.g., hemorrhoids), the sensitivity on a person-level is higher than the per-lesion sensitivity. Our per-person sensitivity as calculated from the model output is quite similar to the German study (32.5% for women vs. 55.4% for men).

Our results are obtained by assuming perfect (100%) attendance, because assuming imperfect adherence could result in overly aggressive screening in hope that, on average, screening is performed at the desired frequency. This would lead to overscreening in those who adhere with recommendations, with the potential for unnecessary risks. However, we showed in the sensitivity analyses that assuming gender-specific realistic attendance did not influence our conclusions.

To our knowledge, this is the first cost-effectiveness analyses to determine the optimal FIT screening strategy by gender. Two limitations are noteworthy. First, we assumed that all differences in the prevalence of adenomas and colorectal cancer incidence between men and women were caused by a difference in adenoma onset and probability to progress to colorectal cancer. We did not assume any differences in dwelling time of adenomas. However, as the relative risk for men and women of nonadvanced adenomas in a German study is similar to the relative risks of colorectal cancer in the Netherlands in the corresponding age group (RR, 1.5; ref. 6), we believe it is likely that the dwelling time of adenomas does not differ significantly between men and women. Second, we introduced a sensitivity analysis in which a proportion of adenomas are systematically missed, but assumed this proportion was equal for men and women. If this proportion does differ, it might influence the preferred screening ages and interval, in theory making screening stratified by gender more beneficial. There are not enough data yet to study this phenomenon for men and women separately, but we did include a sensitivity analysis with a hypothetical difference in the proportion, showing the same conclusion in this sensitivity analysis as the base case analysis.

Various investigators have argued that colorectal cancer screening should be stratified by gender because of the difference in prevalence of (advanced) neoplasia (40, 41) and the gender-related differences in FIT accuracy (5, 9). Our study shows that the added value of gender-based screening is at most marginal. Furthermore, screening stratified by gender may also have disadvantages: some men and women may be confused by the differential recommendations to the point that they no longer attend screening. A slight impact of stratified screening recommendations on attendance will easily offset its marginal benefit. On the other hand, screening stratified by gender may increase attendance because participants feel that the recommendations are better tailored to their risk. Therefore, future research is needed in this area.

Another area for future research is to evaluate the comparative effectiveness of FIT screening and other screening modalities in men and women separately. Earlier studies showed not much difference in cost-effectiveness between a FIT screening program and colonoscopy screening for the population as a whole (42). However, as sensitivity of FIT is lower in women than men, the additional sensitivity of colonoscopy compared with FIT is also higher in women leading to lower comparative effectiveness of FIT with colonoscopy. If the lower sensitivity of FIT in women does not apply to other stool-based tests, the comparative effectiveness of newer tests, such as stool-DNA tests, could also be different than in men.

In conclusion, this study shows that the (cost-)effectiveness of FIT screening is higher in men than in women due to a higher FIT

sensitivity and a higher prevalence of neoplasia in men. However, optimal screening strategies were similar in men and women with respect to interval, age range, and FIT cutoff. Screening stratified by gender does not improve cost-effectiveness, and therefore, our findings support uniform screening of men and women as currently applied in FIT screening programs, like in the Netherlands.

Disclosure of Potential Conflicts of Interest

H.J. de Koning took part in a one-day advisory meeting on biomarkers organized by M.D. Anderson/Health Sciences during the 16th World Conference on Lung Cancer. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: M.P. van der Meulen, M.E. van Leerdam, E.J. Kuipers, H.J. de Koning, I. Lansdorp-Vogelaar
Development of methodology: M.P. van der Meulen, H.J. de Koning, I. Lansdorp-Vogelaar
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): M.E. van Leerdam
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): M.P. van der Meulen, M.E. van Leerdam, A. van der Steen, H.J. de Koning, I. Lansdorp-Vogelaar
Writing, review, and/or revision of the manuscript: M.P. van der Meulen, A. Kapidzic, M.E. van Leerdam, A. van der Steen, E. Kuipers, M.C.W. Spaander, H.J. de Koning, L. Hol, I. Lansdorp-Vogelaar
Study supervision: M.E. van Leerdam, E.J. Kuipers, H.J. de Koning, L. Hol, I. Lansdorp-Vogelaar

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received October 4, 2016; revised February 24, 2017; accepted May 4, 2017; published OnlineFirst May 17, 2017.

References

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA Cancer J Clin 2011;61:69–90.
2. Gupta S, Sussman DA, Doubeni CA, Anderson DS, Day L, Deshpande AR, et al. Challenges and possible solutions to colorectal cancer screening for the underserved. J Natl Cancer Inst 2014;106:dju032.
3. Schreuders EH, Ruco A, Rabeneck L, Schoen RE, Sung JJ, Young GP, et al. Colorectal cancer screening: a global overview of existing programmes. Gut 2015;64:1637–49.
4. Lansdorp-Vogelaar I, van Ballegooijen M, Zauber AG, Boer R, Wilschut J, Winawer SJ, et al. Individualizing colonoscopy screening by sex and race. Gastrointest Endosc 2009;70:96–108.
5. Steele RJ, McClements P, Watling C, Libby G, Weller D, Brewster DH, et al. Interval cancers in a FOBT-based colorectal cancer population screening programme: implications for stage, gender and tumour site. Gut 2012;61:576–81.
6. Brenner H, Haug U, Hundt S. Sex differences in performance of fecal occult blood testing. Am J Gastroenterol 2010;105:2457–64.
7. Stegeman I, de Wijkerslooth TR, Stoop EM, van Leerdam M, van Ballegooijen M, Kraaijenhagen RA, et al. Risk factors for false positive and for false negative test results in screening with fecal occult blood testing. Int J Cancer 2013;133:2408–14.
8. Kapidzic A, van der Meulen MP, Hol L, van Roon AH, Looman CW, Lansdorp-Vogelaar I, et al. Gender differences in fecal immunochemical test performance for early detection of colorectal neoplasia. Clin Gastroenterol Hepatol 2015;13:1464–71.
9. Isabel Portillo CF. Should FIT cut-offs be gender-specific? Chicago USA: World Endoscopy Organization; 2014. Available from: http://www.worldendo.org/wp-content/uploads/2016/08/4_3_isabel_portillo_the_basque_country_ddw2014.pdf
10. Hol L, van Leerdam ME, van Ballegooijen M, van Vuuren AJ, van Dekken H, Reijerink JC, et al. Screening for colorectal cancer: randomised trial comparing guaiac-based and immunochemical faecal occult blood testing and flexible sigmoidoscopy. Gut 2010;59:62–8.
11. Hol L, Wilschut JA, van Ballegooijen M, van Vuuren AJ, van der Valk H, Reijerink JC, et al. Screening for colorectal cancer: random comparison of guaiac and immunochemical faecal occult blood testing at different cut-off levels. Br J Cancer 2009;100:1103–10.
12. Lansdorp-Vogelaar I, van Ballegooijen M, Boer R, Zauber A, Habbema JD. A novel hypothesis on the sensitivity of the fecal occult blood test: results of a joint analysis of 3 randomized controlled trials. Cancer 2009;115:2410–9.
13. Loeve F, Boer R, van Oortmarssen GJ, van Ballegooijen M, Habbema JD. The MISCAN-COLON simulation model for the evaluation of colorectal cancer screening. Comput Biomed Res 1999;32:13–33.
14. Netherlands Comprehensive Cancer Centre. Dutch cancer. [cited 2012 Mar 14]. Available from: http://cijfersoverkanker.nl/.
15. Bombi JA. Polyps of the colon in Barcelona, Spain. An autopsy study. Cancer 1988;61:1472–6.
16. Chapman I. Adenomatous polypi of large intestine: incidence and distribution. Ann Surg 1963;157:223–6.
17. Clark JC, Collan Y, Eide TJ, Esteve J, Ewen S, Gibbs NM, et al. Prevalence of polyps in an autopsy series from areas with varying incidence of large-bowel cancer. Int J Cancer 1985;36:179–86.
18. Jass JR, Young PJ, Robinson EM. Predictors of presence, multiplicity, size and dysplasia of colorectal adenomas. A necropsy study in New Zealand. Gut 1992;33:1508–14.
19. Johannsen LG, Momsen O, Jacobsen NO. Polyps of the large intestine in Aarhus, Denmark. An autopsy study. Scand J Gastroenterol 1989; 24:799–806.
20. Rickert RR, Auerbach O, Garfinkel L, Hammond EC, Frasca JM. Adenomatous lesions of the large bowel: an autopsy survey. Cancer 1979;43: 1847–57.
21. Vatn MH, Stalsberg H. The prevalence of polyps of the large intestine in Oslo: an autopsy study. Cancer 1982;49:819–25.
22. Williams AR, Balasooriya BA, Day DW. Polyps and cancer of the large bowel: a necropsy study in Liverpool. Gut 1982;23:835–42.
23. Stoop EM, de Haan MC, de Wijkerslooth TR, Bossuyt PM, van Ballegooijen M, Nio CY, et al. Participation and yield of colonoscopy versus non-cathartic CT colonography in population-based screening for colorectal cancer: a randomised controlled trial. Lancet Oncol 2012;13:55–64.
24. Atkin WS, Edwards R, Kralj-Hans I, Wooldrage K, Hart AR, Northover JM, et al. Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. Lancet 2010; 375:1624–33.
25. Statistics Netherlands. [cited 2012 Mar 14]; Available from: http://statline.cbs.nl/StatWeb/default.aspx.
26. CBO guideline. [cited 2014 Sep 25]. Available from: http://www.mdl.nl/uploads/240/1308/Richtlijn_Coloscopie_Surveillance_definitief_2013.pdf.
27. Statistics Netherlands. [cited 2012 Mar 14]. Available from: http://statline.cbs.nl/StatWeb/publication/?DM=SLNL&PA=70936ned&D1=a&D2=584,597,610,623,636,649&STB=T,G1&VW=T.
28. Stoop EM. Population-based colorectal cancer screening by colonoscopy or CT-colonography [thesis]. Rotterdam, the Netherlands: Erasmus University Rotterdam; 2013.
29. Rosenthal E. Colonoscopies explain why U.S. leads the world in health expenditures. The New York Times. 2013 Jun 1 [cited 2014 Oct 9]. Available from: http://www.nytimes.com/2013/06/02/health/colonoscopies-explain-why-us-leads-the-world-in-health-expenditures.html?pagewanted=all&_r=1&.
30. Dutch Health Care Authority. Prices and performance. Utrecht, the Netherlands: Dutch Health Care Authority; [cited 2012 Mar 14]. Available from: http://www.nza.nl/regelgeving/tarieven/.
31. College voor zorgverzekeringen (CVZ). Farmacotherapeutisch Kompas. Diemen, the Netherlands: College voor zorgverzekeringen; 2014.
32. Netherlands Comprehensive Cancer Centre. Guidelines for Oncologic Treatments. [cited 2012 Mar 14]. Available from: http://www.oncoline.nl.

Published OnlineFirst May 17, 2017; DOI: 10.1158/1055-9965.EPI-16-0786

33. de Kok IM, Polder JJ, Habbema JD, Berkers LM, Meerding WJ, Rebolj M, et al. The impact of healthcare costs in the last year of life and in all life years gained on the cost-effectiveness of cancer screening. Br J Cancer 2009; 100:1240–4.
34. Lansdorp-Vogelaar I, van Ballegooijen M, Zauber AG, Habbema JD, Kuipers EJ. Effect of rising chemotherapy costs on the cost savings of colorectal cancer screening. J Natl Cancer Inst 2009;101:1412–22.
35. Yabroff KR, Lamont EB, Mariotto A, Warren JL, Topor M, Meekins A, et al. Cost of care for elderly cancer patients in the United States. J Natl Cancer Inst 2008;100:630–41.
36. Ness RM, Holmes AM, Klein R, Dittus R. Utility valuations for outcome states of colorectal cancer. Am J Gastroenterol 1999;94:1650–7.
37. Siegel JE, Torrance GW, Russell LB, Luce BR, Weinstein MC, Gold MR. Guidelines for pharmacoeconomic studies. Recommendations from the panel on cost effectiveness in health and medicine. Panel on cost effectiveness in health and medicine. Pharmacoeconomics 1997;11: 159–68.
38. Donken R, de Melker HE, Rots NY, Berbers G, Knol MJ. Comparing vaccines: a systematic review of the use of the non-inferiority margin in vaccine trials. Vaccine 2015;33:1426–32.
39. Wong MC, Ching JY, Chan VC, Lam TY, Luk AK, Ng SS, et al. Factors associated with false-positive and false-negative fecal immunochemical test results for colorectal cancer screening. Gastrointest Endosc 2015; 81:596–607.
40. Lieberman D. Race, gender, and colorectal cancer screening. Am J Gastroenterol 2005;100:2756–8.
41. Regula J, Kaminski MF. Targeting risk groups for screening. Best Pract Res Clin Gastroenterol 2010;24:407–16.
42. Lansdorp-Vogelaar I, Knudsen AB, Brenner H. Cost-effectiveness of colorectal cancer screening. Epidemiol Rev 2011;33:88–100.

Published OnlineFirst May 17, 2017; DOI: 10.1158/1055-9965.EPI-16-0786

Cancer Epidemiology, Biomarkers & Prevention

AACR American Association for Cancer Research

Do Men and Women Need to Be Screened Differently with Fecal Immunochemical Testing? A Cost-Effectiveness Analysis

Miriam P. van der Meulen, Atija Kapidzic, Monique E. van Leerdam, et al.

Cancer Epidemiol Biomarkers Prev 2017;26:1328-1336. Published OnlineFirst May 17, 2017.