Original Article

Nonbleeding Adenomas: Evidence of Systematic False-Negative Fecal Immunochemical Test Results and Their Implications for Screening Effectiveness—A Modeling Study

Miriam P. van der Meulen, MD[1]; Iris Lansdorp-Vogelaar, PhD[1]; Else-Mariëtte B. van Heijningen, MSc[1]; Ernst J. Kuipers, MD, PhD[2,3]; and Marjolein van Ballegooijen, MD, PhD[1]

BACKGROUND: If some adenomas do not bleed over several years, they will cause systematic false-negative fecal immunochemical test (FIT) results. The long-term effectiveness of FIT screening has been estimated without accounting for such systematic false-negativity. There are now data with which to evaluate this issue. **METHODS:** The authors developed one microsimulation model (MISCAN [MIcrosimulation SCreening ANalysis]-Colon) without systematic false-negative FIT results and one model that allowed a percentage of adenomas to be systematically missed in successive FIT screening rounds. Both variants were adjusted to reproduce the first-round findings of the Dutch CORERO FIT screening trial. The authors then compared simulated detection rates in the second screening round with those observed, and adjusted the simulated percentage of systematically missed adenomas to those data. Finally, the authors calculated the impact of systematic false-negative FIT results on the effectiveness of repeated FIT screening. **RESULTS:** The model without systematic false-negativity simulated higher detection rates in the second screening round than observed. These observed rates could be reproduced when assuming that FIT systematically missed 26% of advanced and 73% of nonadvanced adenomas. To reduce the false-positive rate in the second round to the observed level, the authors also had to assume that 30% of false-positive findings were systematically false-positive. Systematic false-negative FIT testing limits the long-term reduction of biennial FIT screening in the incidence of colorectal cancer (35.6% vs 40.9%) and its mortality (55.2% vs 59.0%) in participants. **CONCLUSIONS:** The results of the current study provide convincing evidence based on the combination of real-life and modeling data that a percentage of adenomas are systematically missed by repeat FIT screening. This impairs the efficacy of FIT screening. *Cancer* **2016;122:1680-8.** © *2016 American Cancer Society.*

KEYWORDS: colorectal cancer, false-negative results, false-positive results, fecal occult blood test, mass screening.

INTRODUCTION

Colorectal cancer (CRC) is the second most common cause of cancer-related mortality in the Western world.[1] Screening can prevent part of these deaths through early detection and treatment. Repeated screening by means of the guaiac fecal occult blood test (gFOBT) reduces mortality by 15% to 33%, as shown in several trials.[2-4] Since these trials were performed, more sensitive FOBTs have been developed. One new version is the fecal immunochemical test (FIT), which is specific for the detection of human globin in stool. Several trials have demonstrated that FIT screening is associated with a higher diagnostic yield and higher adherence compared with gFOBT screening.[5-11] As a consequence, several countries, including Italy, Australia, Japan, and the Netherlands, have adopted population-based FIT screening,[12] whereas in the United States, for example, several local initiatives recently adopted FIT screening.[12,13]

New tests always raise the question of whether randomized controlled trials, which are expensive and take at least 10 years before results are produced, are necessary. To our knowledge, the long-term effectiveness of repeated FIT screening has not yet been studied empirically. Meanwhile, modeling studies have been used to extrapolate the higher diagnostic yield of FIT compared with gFOBT to determine the long-term effectiveness of FIT screening. These studies all assumed that FOBT results in consecutive screening rounds are independent of each other,[14] thereby suggesting that the increase in sensitivity when moving from gFOBT to FIT is retained over the rounds.

Corresponding author: Miriam P. van der Meulen, MD, Department of Public Health, Erasmus MC, University Medical Center Rotterdam, PO Box 2040, 3000 CA Rotterdam, the Netherlands; Fax: (011) 31-10-7038474; m.vandermeulen.1@erasmusmc.nl

[1]Department of Public Health, Erasmus Medical Center, Rotterdam, the Netherlands; [2]Department of Gastroenterology and Hepatology, Erasmus Medical Center, Rotterdam, the Netherlands; [3]Department of Internal Medicine, Erasmus Medical Center, Rotterdam, the Netherlands.

Additional supporting information may be found in the online version of this article.

DOI: 10.1002/cncr.29952, **Received:** December 14, 2015; **Accepted:** January 7, 2016, **Published online** April 8, 2016 in Wiley Online Library (wileyonlinelibrary.com)

However, to the best of our knowledge, little is known regarding when adenomas or carcinomas start to bleed and how often they bleed.[15,16] The gFOBT trials already gave an indirect measure of bleeding patterns in carcinomas,[17] but data concerning adenomas were scarce. If adenomas of a similar size do not have an equal chance to bleed, they also do not have an equal chance of being detected by FIT. This would suggest that FIT results over subsequent rounds are dependent on each other: previous false-negative results increase the probability of another such result. As long as adenomas do not bleed at all, they will cause persistent, so-called systematic false-negative FIT results. Without accounting for this phenomenon, it is possible to overestimate the effectiveness of a screening program with FIT.[18]

Data from repeat screening have demonstrated that the diagnostic yield of adenomas decreases with consecutive screening rounds.[19-21] This is expected due to the depletion of the prevalence pool of adenomas. However, if systematic false-negative test results occur, the diagnostic yield will decrease even further than explained by that depletion.

We used the MISCAN (MIcrosimulation SCreening ANalysis)-Colon microsimulation model to compare simulated detection rates of repeated FIT screening with those observed in a Dutch population-based screening study. We examined to what extent incorporating systematic false-negative FIT results is necessary to explain those observed rates. Finally, we calculated the impact of the estimated amount of systematic false-negative FIT results on the reduction in incidence and mortality of a FIT screening program.

MATERIALS AND METHODS

Overview

We used the MISCAN-Colon model to reproduce the design and first-round findings of the Dutch CORERO FIT screening trial.[7,20,22] We then compared detection rates in the second screening round simulated by the model without systematic false-negative FIT results with those observed in the trial. Subsequently, we developed a variant of the model with systematic false-negative FIT results and estimated how many systematic false-negative FIT results were needed to optimally fit the adenoma detection rates. We also compared the observed second-round false-positive rate with the simulated rate to account for the possible effect of systematic false-positive FIT results. The combination of systematic false-positive and false-negative FIT results was defined as "systematic FIT failure." We then validated the estimates for systematic FIT failure with the third round of the CORERO trial and with first-round observations of the CORERO trial in the group that performed a 2-sample FIT. Finally, we calculated the impact of systematic FIT failure on the effectiveness of FIT screening by running screening programs in the models with and without systematic FIT failure.

The CORERO Trial

The CORERO phase 1 trial was a randomized controlled trial comparing attendance to and yield of gFOBT with those of FIT and sigmoidoscopy at first-round screening.[7] Subsequently, the phase 2 CORERO trial examined the attendance and detection rates of repeat FIT screening at different intervals.[20] Details from these trial phases have been described elsewhere.[7,20,22,23] In short, screening-naive subjects aged 50 to 74 years who were living in the southwest of the Netherlands were selected through municipal population registers. In the first round, these individuals were sent a kit with a single FIT test (OC-Sensor; Eiken Chemical Company, Tokyo, Japan). A cutoff value of 10 μg of hemoglobin per gram (hemoglobin/g) of feces (equivalent to 50 ng of hemoglobin/mL) was used to indicate a positive test result, after which a colonoscopy was offered by experienced endoscopists. A second screening round took place after an interval of 1, 2, or 3 years (group I, II, and III, respectively).[20] A third screening round took place after an interval of 2 years for all 3 groups.[23] In the first screening round, 4523 individuals (62.6%) returned the test kit. Of these participants, 3427 (90.6%) responded in the second screening round. Because in the current analysis we were interested in systematic FIT failure, we only considered those who participated in both screening rounds for the main analysis and those who participated in all 3 rounds for the validation to the third round.

The phase 1 CORERO trial also contained a group of individuals who received a kit with 2 FIT samples (group IV).[22] They were instructed to conduct the 2 tests on subsequent days. Of this group, 1876 invitees (61.2%) returned the test kit.

MISCAN-Colon

MISCAN-Colon is a microsimulation model for CRC that was developed at the Department of Public Health of the Erasmus University Medical Center (Rotterdam, the Netherlands). The model's structure, underlying assumptions, and calibration are described in the online Supporting Information and previous publications.[24,25] In brief,

the MISCAN-Colon model simulates the relevant life histories of a large population of individuals from birth to death. CRC arises in this population according to the adenoma-carcinoma sequence.[26,27] Greater than 1 adenoma can occur in an individual and each adenoma can independently develop into a CRC. Adenomas may progress in size from small (≤5 mm) to medium (6-9 mm) to large (≥10 mm). Although the majority of adenomas will never turn into cancer, some will eventually become malignant. Cancer starts as a symptomless process and can progress from localized cancer of stage I to metastasized cancer of stage IV. In every stage of disease, there is a probability of the CRC being diagnosed due to the development of symptoms versus a symptomless progression into the next stage. At any time during the development of the disease, the process may be interrupted because an individual dies of other causes. With FIT screening, lesions can be detected before clinical diagnosis; a screened individual with a positive test result will be referred for a colonoscopy for the detection and removal of adenomas and cancers, hopefully at an earlier stage of disease. In this way, CRC incidence and/or CRC-related mortality can be reduced.

MISCAN-Colon was calibrated to the age-specific, stage-specific, and localization-specific incidence of CRC as observed in the Netherlands before the introduction of screening (ie, between 1999-2003)[28] and the age-specific prevalence and multiplicity distribution of adenomas as observed in autopsy studies.[29-37] The size distribution of adenomas was calibrated to the size distribution detected with colonoscopy in the COlonography for Screening (COCOS) trial.[38] The preclinical duration of CRC 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[2-4] and a once-only sigmoidoscopy[39] and demonstrated good concordance with the observed reduction in mortality (see online Supporting Information).

FIT Characteristics

Test characteristics of the 1-sample FIT were fitted to the positivity and detection rates of nonadvanced and advanced adenomas and carcinomas as observed in the first screening round of the CORERO trial and its counterpart in 2 other regions of the Netherlands.[8,10]

Sensitivity and systematic false-negative FIT results

An advanced adenoma was defined as an adenoma measuring ≥10 mm, with a villous component of ≥25% and/or high-grade dysplasia. Because the model does not incorporate histology, the observed detection rate of nonadvanced and of advanced adenomas was fitted to the detection rate of small to medium and of large adenomas in the model, respectively. We modeled sensitivity by giving each lesion a probability to cause a positive FIT result. The fitted per-lesion probabilities were 0% for adenomas measuring 1 to 5 mm, 11.4% for adenomas measuring 6 to 9 mm, 34.4% for adenomas measuring ≥ 10 mm, 50.3% for carcinomas long before clinical diagnosis, and 82.5% for carcinomas shortly before clinical diagnosis.

We then included systematic false-negative FIT results by simulating the following concept. When an adenoma starts to develop, it will not bleed at first. During that phase, it will not have any chance to cause a positive FOBT. Once it starts to bleed, it has a random chance, based on the sensitivity, to cause a positive FOBT. We simulated this process in the model by discerning adenomas that already were bleeding with a random sensitivity, and adenomas that did not bleed yet with a sensitivity of zero, a systematic false-negativity. The random sensitivity was corrected upward, such that the overall sensitivity for adenomas in the first screening round was not affected.

The probability of a systematic false-negative FIT result was estimated separately for advanced and nonadvanced adenomas.

It is interesting to note that all types of sensitivity are set per lesion. If an individual had a second adenoma or a CRC that generated a positive test result, or had a positive test result due to the lack of specificity (chance detection), an adenoma that was "missed" (it did not by itself generate a positive test) could still be detected through diagnostic colonoscopy.

Specificity and systematic false-positive FIT results

In previous modeling studies, we modeled a lack of specificity as a per-person probability of having a positive test, independent of whether the individual had neoplasia. The fitted probability in this study was 95.7%.

The same concept as systematic false-negative FIT results can in principle also occur in false-positivity. This would occur for example if individuals have a constitutional or chronic condition causing fecal bleeding over several years. In this case, individuals who had a negative colonoscopy after a positive FIT are put on hold for several FIT screening rounds (as was done in the CORERO trial); this results in a depletion of the prevalence pool of such individuals and thus in lower than expected false-positive rates in later rounds.

We modeled systematic false-positive test results by assuming that a percentage of individuals would always test false-positive. We did this by assigning these individuals with a (0% specificity) systematic (false-)positivity, as opposed to individuals with a random specificity. The random specificity was increased so that the introduction of systematic false-positivity did not affect the overall number of false-positive test results in a first screening round.

Colonoscopy Characteristics

The sensitivity of diagnostic colonoscopies was assumed to be 75% for adenomas measuring ≤ 5 mm, 85% for adenomas measuring 6 to 9 mm, and 95% for adenomas measuring ≥ 10 mm and CRC.[40]

Data Analysis

Calibration of systematic FIT failure

At first, a population was simulated with birth years that matched those of the invitees to the first screening round of the CORERO trial. This was done separately for each group (group I through group IV). The FIT screening strategies (1-day interval for the 2-sample group and a 1-year, 2-year, or 3-year interval for the 1-sample groups), the attendance over the 2 screening rounds, and the compliance to diagnostic colonoscopy after a positive FIT result also were matched to the observed data. For both the observed and simulated data, we presented aggregated data for the 3 groups in the current study.

We then simulated 2 consecutive screening rounds with a 1-sample FIT at a cutoff value of 10 μg of hemoglobin/g of feces in 10 million individuals with the model without systematic FIT failure. The positivity and detection rates and the positive predictive value of the second round were determined and compared with the observed rates and their 95% confidence intervals (95% CIs) in the CORERO trial. The positivity and detection rates were determined by dividing the number of events (individuals with a positive FIT result, individuals with a detected adenoma) by the number of individuals screened.

Subsequently, the size of the systematic component of specificity and sensitivity was estimated by minimizing the difference between the observed and simulated rates in the second screening round. We first estimated the size of the systematic component of specificity on the observed difference in the false-positive rate between the first and second screenings. We then estimated the probability of systematic false-negative test results for nonadvanced adenomas and for advanced adenomas on the observed second round detection rate of nonadvanced adenomas and advanced adenomas.

Validation of systematic FIT failure

To validate the systematic model, we used both the version of the MISCAN-Colon model with and that without systematic FIT failure to project positivity and detection rates in individuals undergoing a third screening round and compared these with the observed third screening round of the CORERO trial. We also simulated 2-sample FIT results with and without systematic FIT failure and compared these with the 2-sample observed data from group IV in the CORERO trial to validate the model in a different data set.

Effectiveness

We simulated a Dutch population born in 1955 until death who all attended the screening protocol as was introduced in the Netherlands in 2014 (starting age of 55 years, stopping age of 75 years, and a 2-year interval). We considered full attendance to explore the effect on individuals who comply with the complete program. First, we assumed no systematic FIT failure and we then introduced systematic false-negativity, systematic false-positivity, and finally both. We compared the mortality reduction, incidence reduction, lifetime number of colonoscopies per person, number needed to scope to prevent 1 death (NNScope), life-years gained, and quality-adjusted life-years gained for all 4 scenarios. The life-years and quality-adjusted life-years gained were undiscounted and alternatively discounted with 3%.

In addition, to demonstrate the maximum possible effect of systematic FIT failure, we repeated these runs assuming that all first-round false-negative adenomas were nonbleeding adenomas.

RESULTS

The CORERO Trial

In the first round of the CORERO trial, 8.4% of the participants had a positive FIT, whereas in the second round, the positivity rate decreased to 5.8%.[20] This decline was caused by a decline in both false-positive test results (false-positive rate) and true-positive test results. The decline in true-positive test results was reflected in the detection rate of nonadvanced adenomas, advanced adenomas, and CRCs (Table 1).

Simulated Results of the First Round

The model successfully reproduced the observed positivity and detection rates of FIT in the first round. As explained

TABLE 1. Observed and Simulated Positivity Rates, Detection Rates, and the PPV of the First and Second Rounds: Data From the 1-Sample FIT With a Cutoff Value of 10 μg Hb/g of Feces in All Groups of the CORERO Trial and Model Without and With Systematic FIT Failure

	First Round			Second Round		
	Observed (95% CI) N=4523	Simulated Without Systematic FIT Failure	Simulated With Systematic FIT Failure	Observed (95% CI) N=3427	Simulated Without Systematic FIT Failure	Simulated With Systematic FIT Failure
Positivity rate, %[a]	8.4 (7.6-9.2)	8.3	8.3	5.8 (5.1-6.6)	7.3	5.8
False-positive rate, %[a]	3.5 (2.9-4.0)	3.4	3.4	2.9 (2.3-3.4)	3.5	2.7
Detection rate of nonadvanced adenomas, %[a]	1.7 (1.3-2.0)	1.7	1.7	1.2 (0.9-1.6)	1.7	1.3
Detection rate of advanced adenomas, %[a]	2.8 (2.3-3.3)	2.7	2.7	1.6 (1.2-2.0)	1.9	1.6
Detection rate of CRC, %[a]	0.49 (0.28-0.69)	0.49	0.50	0.18 (0.04-0.32)	0.28	0.27
Detection rate of advanced neoplasia, %[a]	3.3 (2.8-3.8)	3.2	3.2	1.8 (1.3-2.2)	2.1	1.8
PPV, %[a]	38.9 (34.0-43.9)	38.5	38.6	30 (23.6-36.4)	29.4	31.7

Abbreviations: 95% CI, 95% confidence interval; CRC, colorectal cancer; FIT, fecal immunochemical test; Hb, hemoglobin; PPV, positive predictive value.
[a] Number of cases divided by the total number of individuals screened.

earlier, the introduction of systematic FIT failure did not appear to affect the simulated positivity and detection rates of the first screening round (Table 1).

Simulated Results of the Second Round

In the second round, the model without systematic FIT failure was found to have a considerably higher positivity rate than observed. Although the simulated positivity rate declined by 1.0 percentage point between the first and second rounds (from 8.3% to 7.3%), the observed decline was 2.6 percentage points (from 8.4% to 5.8%). This smaller simulated decline resulted both from a smaller decline in the false-positive rate and a smaller decline in simulated detection rates. The false-positive rate simulated by the model without systematic FIT failure increased by 0.1 of a percentage point from the first to the second screening round (from 3.4% to 3.5%), whereas in real life the false-positive rate decreased by 0.6 of a percentage point (from 3.5% to 2.9%). The simulated rate for the second round was outside the 95% CI of the observed rate. In the model with systematic FIT failure, the observed false-positive rate was fitted best when assuming 1.3% of the participants had systematic false-positivity, which corresponds to 30% of the individuals with a false-positive result in the first round.

The detection rate for nonadvanced adenomas simulated by the model without systematic FIT failure did not change between the first and second round (1.7% both rounds), whereas the observed detection rate declined from 1.7% to 1.2%. Here also, the simulated rate in the second round was outside the 95% CI of the observed rate. When assuming that 73% of nonadvanced adenomas systematically tested false-negative, the model with systematic FIT failure fitted the observed second-round detection rate best.

For advanced adenomas, the second-round detection rate simulated without systematic FIT failure was inside the 95% CI of the observed rate, although again, the observed rate demonstrated a larger decrease between the rounds (from 2.8% to 1.9% [simulated] vs 1.6% [observed]). In the model with systematic FIT failure, the observed detection rate was fitted best when assuming 26% of advanced adenomas systematically tested false-negative.

For CRC, the simulated detection rate declined by 0.21 of a percentage point from the first to the second round (from 0.49% to 0.28%), whereas the observed decline was larger at 0.31 of a percentage point (from 0.49% to 0.18%). However, the cancer detection rate in the second round was within the wide 95% CI of the observed rate.

TABLE 2. Positivity Rates, Detection Rates, and the PPV of the Third Round With 1-Sample FIT and the First Round With 2-Sample FIT With a Cutoff Value of 10 μg Hb/g of Feces as Observed in Group IV of the COR-ERO Trial and in the Model Without and With Systematic FIT Failure

	First Round 2-Sample FIT			Third Round 1-Sample FIT		
	Observed (95% CI) N=1876	Simulated Without Systematic FIT Failure	Simulated With Systematic FIT Failure	Observed (95% CI) N=2907	Simulated Without Systematic FIT Failure	Simulated With Systematic FIT Failure
Positivity rate, %[a]	12.7 (11.2-14.3)	15.1	13.8	4.6 (3.9-5.4)	6.6	5.0
False-positive rate, %[a]	6.1 (5.0-7.2)	6.4	5.8	2.3 (1.8-2.9)	3.4	2.6
Detection rate of nonadvanced adenomas, %[a]	2.6 (1.8-3.3)	3.3	2.9	1.3 (0.9-1.7)	1.6	1.1
Detection rate of advanced adenomas, %[a]	3.4 (2.6-4.2)	4.6	4.3	0.9 (0.6-1.3)	1.4	1.1
Detection rate of CRC, %[a]	0.69 (0.32-1.07)	0.75	0.75	0.10 (0.0-0.3)	0.21	0.22

Abbreviations: 95% CI, 95% confidence interval; CRC, colorectal cancer; FIT, fecal immunochemical test; Hb, hemoglobin; PPV, positive predictive value.
[a] Number of cases divided by the total number of individuals screened.

Third Round and 2-Sample Group

The observed positivity and detection rates of the third round of 1-sample screening and the first round of 2-sample screening are shown in Table 2. Without systematic FIT failure, the simulated positivity and detection rates all were higher than observed and the majority fell outside of the 95% CIs of the observed rates in the COR-ERO trial. With systematic FIT failure, all the simulated positivity and detection rates decreased and fell within the 95% CI of the observed rates, although most remained higher than the observed rates.

Effectiveness

The introduction of systematic false-positive test results decreased the number of screened individuals who needed to undergo a colonoscopy to prevent 1 death; the NNScope was reduced by 7.4% (from 41.1 to 38.0), while maintaining a rate of nearly 99% of life-years gained (Table 3). It reduced the average number of diagnostic and total colonoscopies performed per person who initiated the screening program by 24% and 19%, respectively. With the introduction of systematic false-negative test results, incidence reduction, mortality reduction, and life-years gained from screening declined by 9.4%, 4.5%, and 3.8%, respectively, compared with no systematic false-negativity, whereas the NNScope of the program also decreased. Together, both elements of systematic FIT failure resulted in a decline of 13.0% in incidence reduction (from 40.9% to 35.6%), a decline of 6.4% in mortality reduction (from 59.0% to 55.2%), and a decline of 5.2% in life-years gained (from 245 life-years gained to 232 life-years gained per 1000 participants). If the life-years gained were 3% discounted, the decline was 4.8% (from 115 life-years gained to 109 life-years gained). The NNscope and the average number of diagnostic and total colonoscopies per person also declined. When we assumed that all first-round false-negative lesions were systematically false-negative, the incidence reduction was found to decline by 34.8% (from 40.9% to 26.7%), the mortality reduction declined by 16.4% (from 59.0% to 49.3%), and the life-years gained declined by 14.0% (from 245 life-years gained to 211 life-years gained per 1000 participants).

DISCUSSION

The results of the current analysis demonstrate that the lower detection rate of adenomas observed in the second round of FIT screening is not only caused by depletion of the adenoma prevalence pool. We needed to assume that 73% of nonadvanced adenomas and 26% of advanced adenomas are systematically missed by FIT to simulate the observed decrease in detection rates between the first and the second rounds of FIT screening. We also needed to assume that 30% of false-positive results are systematically false-positive to simulate the observed decrease in the positivity rate. Compared with the projections without systematic FIT failure, the projections with systematic FIT failure resulted in a 5.2% decrease in life-years gained by biennial FIT screening from age 55 to 75 years, whereas the reduction in incidence and mortality decreased by 13.0% and 6.4%, respectively.

Systematic false-negative test results are a very plausible and even expected phenomenon, with nonbleeding adenomas being the most obvious explanation. Other explanations include a longer time for hemoglobin decay

TABLE 3. Simulated Changes in Lifelong Outcomes of the Screening Program (Aged 55-75 Years, 2-Year Interval, FIT With a Cutoff Value of 10 μg Hb/g Feces) in Participants: 4 FIT Failure Model Versions With No Discounting

	Model Without Systematic FIT Failure	Model With Systematic False-Negative Test Results	Model With Systematic False-Positive Test Results	Model With Systematic FIT Failure (Both Types)
Incidence reduction	40.9%	37.1%	39.8%	35.6%
Mortality reduction	59.0%	56.4%	58.1%	55.2%
Average no. of colonoscopies per participant (diagnostic plus surveillance)	0.86	0.79	0.78	0.7
Average no. of diagnostic colonoscopies per participant	0.49	0.39	0.44	0.37
NNscope to prevent 1 death	41.1	39.7	38	36.4
Life-years gained[a]	245	235	242	232
QALYs gained[a]	257	243	253	238

Abbreviations: FIT, fecal immunochemical test; Hb, hemoglobin; NNScope, number needed to scope; QALYs, quality-adjusted life-years.
[a] Per 1000 persons starting the screening program.

caused by a proximal location of adenomas or longer colonic transit in an individual, and different sampling techniques. It is expected that systematic false-negative FIT results are associated with the decreased effectiveness of a FIT screening program because adenomas that are systematically missed in consecutive screening rounds decrease the sensitivity of a multiround screening program. This impairs screening effectiveness in terms of incidence and mortality reduction.[18] Systematic false-positive test results also are not unexpected. Obviously, there are individuals with a constitutional or chronic condition causing fecal bleeding. If those individuals are selected out at earlier rounds, it explains why the number of false-positive test results in subsequent rounds decreases. Accounting for this latter mechanism had a positive influence by reducing the NNScope and the average number of colonoscopies per participant. The NNscope to prevent 1 death also was reduced when accounting for systematic false-negative test results. This may seem surprising at first, but is fully plausible. As shown by the calibrated percentages, systematic false-negative test results especially play a role in nonadvanced adenoma, in which they reduced the number of colonoscopies relatively more than that they reduced the effectiveness of adenoma removal.

In this analysis, we focused on the systematic false-negativity of FIT for the detection of adenomas. Based on an earlier analysis of the gFOBT Hemoccult trials, we already identified a lower sensitivity for carcinomas long before clinical diagnosis compared with shortly before clinical diagnosis, which we have since incorporated into our model.[17] Due to the low number of carcinomas in the CORERO trial, it was not possible to reliably validate these earlier assumptions regarding cancers.

A strength of the current study is that we validated the estimated systematic FIT failure to the third round of screening as well as in a different study population using a 2-sample FIT. Here also the model with systematic FIT failure better matched real-life observations compared with the model without systematic FIT failure, both for nonadvanced adenomas and advanced adenomas. This further strengthens the case for the occurrence of systematic false-negative FIT results.

Two limitations of the current study are noteworthy. First, we estimated systematic FIT failure for a cutoff level of 10 μg of hemoglobin/g of feces. Based on the available data, we could not estimate it for higher cutoff values. However, if systematic false-negative FIT results are primarily caused by nonbleeding adenomas, one could argue that the same number of adenomas will not bleed when testing with a higher FIT cutoff value. The differences in detection rates between cutoff values will then be fully explained by the sensitivity for those adenomas that do bleed. Second, we assumed that bleeding and nonbleeding adenomas have the same probability of progressing to cancer. If nonbleeding adenomas have a lower probability to progress than bleeding adenomas, this will attenuate the impact on the effectiveness of FIT screening.

To our knowledge, there is no other CRC screening model to date that has addressed the issue of systematic false-negative FIT results with which to compare our estimates. With regard to observed data from consecutive screening rounds, an earlier FIT study in Italy reported an

even larger difference in (advanced) adenoma yield between the first and the second rounds after 2 years than observed from the CORERO trial data we used.[21] Conversely, a Dutch FIT study from Amsterdam reported a less pronounced difference between the screening rounds.[19] Given these data and the 95% CIs of the rates in the CORERO trial, there remains uncertainty regarding the exact amount of systematic false-negative FIT results. Therefore, it is interesting to note that under the extreme assumption that all adenomas missed during the first round will be missed at subsequent rounds, the mortality reduction decreased by 16.4% (from 59.0% to 49.3%) instead of 6.4% (from 59.0% to 55.2%). There also is some uncertainty regarding the exact mechanism. We assumed that adenomas that have never bled can only start to bleed when progressing from a nonadvanced to an advanced stage. In real life, this will happen on a more continuous basis (eg, when an adenoma grows from 6 to 8 mm). We also assumed that adenomas either have no sensitivity or full sensitivity, whereas in real life, presumably, some adenomas will bleed more often or may shed more blood when bleeding than others, leading to varying sensitivity levels.

The results of the current study provide convincing evidence that a percentage of adenomas are systematically missed by repeated FIT screening, presumably due to nonbleeding adenomas. This phenomenon lowers the impact of FIT screening on mortality reduction by an estimated 6.4%.

FUNDING SUPPORT

Funded by ZonMw (the Netherlands Organisation for Health Research and Development) as project number 120720011.

CONFLICT OF INTEREST DISCLOSURES

The authors made no disclosures.

AUTHOR CONTRIBUTIONS

Iris Lansdorp-Vogelaar and Marjolein van Ballegooijen conceived the idea for the study and supervised the model simulations and analysis. All authors contributed to the analysis and interpretation of the data. Miriam P. van der Meulen drafted the article. All authors have contributed to the critical revision of the article. All authors read and approved the final article.

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. Hardcastle JD, Armitage NC, Chamberlain J, Amar SS, James PD, Balfour TW. Fecal occult blood screening for colorectal cancer in the general population. Results of a controlled trial. *Cancer*. 1986; 58:397-403.
3. Kronborg O, Fenger C, Olsen J, Jorgensen OD, Sondergaard O. Randomised study of screening for colorectal cancer with faecal-occult-blood test. *Lancet*. 1996;348:1467-1471.
4. Mandel JS, Church TR, Ederer F, Bond JH. Colorectal cancer mortality: effectiveness of biennial screening for fecal occult blood. *J Natl Cancer Inst*. 1999;91:434-437.
5. Allison JE, Sakoda LC, Levin TR, et al. Screening for colorectal neoplasms with new fecal occult blood tests: update on performance characteristics. *J Natl Cancer Inst*. 2007;99:1462-1470.
6. Guittet L, Bouvier V, Mariotte N, et al. Comparison of a guaiac based and an immunochemical faecal occult blood test in screening for colorectal cancer in a general average risk population. *Gut*. 2007; 56:210-214.
7. Hol L, van Leerdam ME, van Ballegooijen M, et al. Screening for colorectal cancer: randomised trial comparing guaiac-based and immunochemical faecal occult blood testing and flexible sigmoidoscopy. *Gut*. 2010;59:62-68.
8. Hol L, Wilschut JA, van Ballegooijen M, 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;100:1103-1110.
9. Park DI, Ryu S, Kim YH, et al. Comparison of guaiac-based and quantitative immunochemical fecal occult blood testing in a population at average risk undergoing colorectal cancer screening. *Am J Gastroenterol*. 2010;105:2017-2025.
10. van Rossum LG, van Rijn AF, Laheij RJ, et al. Random comparison of guaiac and immunochemical fecal occult blood tests for colorectal cancer in a screening population. *Gastroenterology*. 2008;135:82-90.
11. Whitlock EP, Lin JS, Liles E, Beil TL, Fu R. Screening for colorectal cancer: a targeted, updated systematic review for the U.S. Preventive Services Task Force. *Ann Intern Med*. 2008;149:638-658.
12. Schreuders EH, Ruco A, Rabeneck L, et al. Colorectal cancer screening: a global overview of existing programmes. *Gut*. 2015;64:1637-1649.
13. Gupta S, Sussman DA, Doubeni CA, et al. Challenges and possible solutions to colorectal cancer screening for the underserved. *J Natl Cancer Inst*. 2014;106:dju032.
14. Lansdorp-Vogelaar I, Knudsen AB, Brenner H. Cost-effectiveness of colorectal cancer screening. *Epidemiol Rev*. 2011;33:88-100.
15. Lang CA, Ransohoff DF. On the sensitivity of fecal occult blood test screening for colorectal cancer. *J Natl Cancer Inst*. 1997;89: 1392-1393.
16. Ransohoff DF, Lang CA, Young GP. Colorectal cancer screening: clinical guidelines and rationale. *Gastroenterology*. 1997;113:1423-1424.
17. 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;115:2410-2419.
18. Loeve F, Boer R, van Oortmarssen GJ, van Ballegooijen M, Habbema JD. Impact of systematic false-negative test results on the performance of faecal occult blood screening. *Eur J Cancer*. 2001;37: 912-917.
19. Denters MJ, Deutekom M, Bossuyt PM, Stroobants AK, Fockens P, Dekker E. Lower risk of advanced neoplasia among patients with a previous negative result from a fecal test for colorectal cancer. *Gastroenterology*. 2012;142:497-504.
20. van Roon AH, Goede SL, van Ballegooijen M, et al. Random comparison of repeated faecal immunochemical testing at different intervals for population-based colorectal cancer screening. *Gut*. 2013;62:409-415.
21. Zorzi M, Barca A, Falcini F, et al. Screening for colorectal cancer in Italy: 2005 survey. *Epidemiol Prev*. 2007;31(2-3 suppl 2):49-60.
22. van Roon AH, Wilschut JA, Hol L, et al. Diagnostic yield improves with collection of 2 samples in fecal immunochemical test screening without affecting attendance. *Clin Gastroenterol Hepatol*. 2011;9:333-339.
23. Kapidzic A, Grobbee EJ, Hol L, et al. Attendance and yield over three rounds of population-based fecal immunochemical test screening. *Am J Gastroenterol*. 2014;109:1257-1264.
24. 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.

25. Vogelaar I, van Ballegooijen M, Zauber AG. Modeler Profiler of the MISCAN-Colon Microsimulation Model For Colorectal Cancer. 2012. Available from: https://cisnet.flexkb.net/mp/pub/cisnet_colorectal_sk_erasmus_profile.pdf.
26. Morson B. President's address. The polyp-cancer sequence in the large bowel. *Proc R Soc Med*. 1974;67(6 pt 1):451-457.
27. Muto T, Bussey HJ, Morson BC. The evolution of cancer of the colon and rectum. *Cancer*. 1975;36:2251-2270.
28. Netherlands Comprehensive Cancer Organisation. Dutch cancer figures. http://cijfersoverkanker.nl/. Accessed March 14, 2012.
29. Arminski TC, McLean DW. Incidence and distribution of adenomatous polyps of the colon and rectum based on 1,000 autopsy examinations. *Dis Colon Rectum*. 1964;7:249-261.
30. Bombi JA. Polyps of the colon in Barcelona, Spain. An autopsy study. *Cancer*. 1988;61:1472-1476.
31. Chapman I. Adenomatous polypi of large intestine: incidence and distribution. *Ann Surg*. 1963;157:223-226.
32. Clark JC, Collan Y, Eide TJ, et al. Prevalence of polyps in an autopsy series from areas with varying incidence of large-bowel cancer. *Int J Cancer*. 1985;36:179-186.
33. Jass JR, Young PJ, Robinson EM. Predictors of presence, multiplicity, size and dysplasia of colorectal adenomas. A necropsy study in New Zealand. *Gut*. 1992;33:1508-1514.
34. Johannsen LG, Momsen O, Jacobsen NO. Polyps of the large intestine in Aarhus, Denmark. An autopsy study. *Scand J Gastroenterol*. 1989;24:799-806.
35. Rickert RR, Auerbach O, Garfinkel L, Hammond EC, Frasca JM. Adenomatous lesions of the large bowel: an autopsy survey. *Cancer*. 1979;43:1847-1857.
36. Vatn MH, Stalsberg H. The prevalence of polyps of the large intestine in Oslo: an autopsy study. *Cancer*. 1982;49:819-825.
37. Williams AR, Balasooriya BA, Day DW. Polyps and cancer of the large bowel: a necropsy study in Liverpool. *Gut*. 1982;23:835-842.
38. Stoop EM, de Haan MC, de Wijkerslooth TR, 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.
39. Atkin WS, Edwards R, Kralj-Hans I, et al; UK Flexible Sigmoidoscopy Trial Investigators. Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. *Lancet*. 2010;375:1624-1633.
40. van Rijn JC, Reitsma JB, Stoker J, Bossuyt PM, van Deventer SJ, Dekker E. Polyp miss rate determined by tandem colonoscopy: a systematic review. *Am J Gastroenterol*. 2006;101:343-350.