Supplementary Online Content


eAppendix

eMethods, p.2

eFigure 1. The stages of disease in the semi-Markov model, p.5

eFigure 2. Simulated versus observed adenoma prevalence in selected autopsy studies (with 95% confidence intervals), p.6

eFigure 3. Simulated versus observed colorectal cancer incidence in 1975-1979 Surveillance Epidemiology and End Results program data, p.7

eFigure 4. Simulated versus observed distal colorectal cancer incidence in the intervention group of the UK Flexible Sigmoidoscopy Trial, p.9

eFigure 5. Integrating MISCAN modules for two example patients, p.10

eFigure 6. Bootstrap analysis for average cancer incidence and adenoma detection rates at Kaiser Permanente Northern California, p.12

eFigure 7. Simulated versus observed average cancer incidence and adenoma detection rates at Kaiser Permanente Northern California, p.13

eFigure 8. Simulated versus observed cancer incidence rates (cumulative) in the Kaiser Permanente Northern California data, p.15

eTable 1. Sensitivity analysis results: The adenoma detection rate-outcome relationship for various modeling scenarios, p.17

eTable 2. Modeled results: Effectiveness of screening colonoscopy according to quintile of adenoma detection rate (0% discounted), p.19

eTable 3. Modeled results: Resources and complications for colonoscopy screening according to quintile of adenoma detection rate (0% discounted), p.20

eReferences. p.21

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

© 2015 American Medical Association. All rights reserved.
Microsimulation model structure

MISCAN-colon is a stochastic, semi-Markov, microsimulation model for colorectal cancer (CRC) programmed in Delphi (Borland Software Corporation, Scotts Valley, California, United States). It can be used to explain and predict trends in CRC incidence and mortality and to quantify the effects and costs of primary prevention of CRC, screening for CRC, and surveillance after polypectomy.

The term ‘microsimulation’ implies that the individuals are moved through the model one at a time, rather than as proportions of a cohort. The term ‘semi-Markov’ implies that MISCAN-colon, unlike traditional Markov models, does not assume annual state transitions; instead it generates durations in states, allowing future state transitions to depend on past transitions, and thereby increases model flexibility and computational performance. The term ‘stochastic’ implies that the model determines the states and corresponding durations by drawing from probability distributions, rather than using fixed values. Hence, the results of the model are subject to random variation.

MISCAN-colon consists of three modules: a demography module, natural history module, and screening module.

Demography module

Using birth- and life-tables representative for the population under consideration, MISCAN-colon draws a date of birth and a date of non-CRC death for each individual simulated. The maximum age an individual can achieve is assumed to be 100 years.

Natural history module

Transitions

As each simulated person ages, one or more adenomas may develop (eFigure 1). These adenomas can be either progressive or non-progressive. Both progressive and non-progressive adenomas can grow in size from small (≤5mm), to medium (6-9mm), to large (≥10mm); however, only progressive adenomas can develop into preclinical cancer. A preclinical cancer may progress through stages I to IV without symptoms, or be diagnosed during each stage CRC because of symptoms. After clinical diagnosis, CRC survival is simulated using age-, stage-, and localization-specific survival estimates for clinically diagnosed CRC as obtained by Rutter and colleagues. For individuals with synchronous CRCs at time of diagnosis, the survival of the most advanced cancer is used. The date of death for individuals with CRC is set to the earliest simulated death due either to CRC or another cause (‘Demography module’).

Transition rates and durations

An individual’s risk of developing adenomas depends on the individual’s age and a personal Gamma-distributed risk index (non-homogeneous Poisson process). As a result of the latter most individuals develop no adenomas, whilst some develop many. We assumed that the distribution of adenomas over the colon and rectum equals the distribution of cancers as observed in SEER before the introduction of screening. The age-specific onset of adenomas and the dispersion of the personal risk index were calibrated to data on the prevalence and multiplicity distribution of adenomas as observed in autopsy studies (eFigure 2). The age-specific probability of adenoma-progressivity and the age- and localization-specific transition
probabilities between preclinical cancer stages and between preclinical and clinical cancer stages were simultaneously calibrated to SEER data on the age-, stage-, and localization-specific incidence of CRC as observed before the introduction of screening (eFigure 3).3 The average durations between the preclinical cancer stages were calibrated to the rates of screen-detected and interval cancers observed in randomized controlled trials evaluating screening using guaiac fecal occult blood tests.14-16 This exercise has been described extensively elsewhere.17 The average duration from the emergence of an adenoma until progression into preclinical cancer (i.e. the adenoma dwell-time) was calibrated to the rates of interval cancers (including surveillance detected cancers) observed in a randomized controlled trial evaluating once-only sigmoidoscopy screening (eFigure 4).17 We assumed an equal overall dwell-time for adenomas developing into CRC from a medium size (30% of all CRCs) and from a large size (70% of all CRCs). All durations in the adenoma and preclinical cancer phase were drawn from Exponential distributions. Durations of the disease stages within the adenoma and preclinical cancer phase, respectively, were assumed to be perfectly correlated (i.e. if a small adenoma grows into a medium-sized adenoma rapidly, it will also grow into a large adenoma or develop into CRC rapidly). However, durations in the adenoma phase were assumed to be uncorrelated with durations in the preclinical cancer phase (i.e. a rapidly growing adenoma does not necessarily develop into a rapidly progressing cancer). The proportion of medium sized, non-progressive adenomas growing large and the average duration in the medium size, non-progressive adenoma state were calibrated to size-specific adenoma detection rates observed in a Dutch randomized controlled trial on colonoscopy screening (not shown). All calibrations were performed using the Nelder-Mead search algorithm to minimize deiances from observed values based on log-likelihood functions (Poisson likelihood for incidence, Binomial likelihood for adenoma prevalence, and Multinomial likelihood for cancer stages).

**Screening module**

Screening will alter some of the simulated life histories: some cancers will be prevented by the detection and removal of adenomas, while other cancers will be detected in an earlier stage with a more favorable survival. As the stage-specific survival of screen-detected CRC as observed in randomized controlled trials on guaiac fecal occult blood testing was substantially more favorable than that of clinically detected CRC, even after correcting for lead-time bias,17 we assigned those screen-detected cancers that would have been clinically detected in the same stage the survival corresponding to a one stage less progressive cancer. Hence, a cancer screen-detected in stage II, that would also have been clinically diagnosed in stage II, is assigned the survival of a clinically diagnosed stage I cancer. The only exceptions were screen-detected stage IV cancers. These cancers were always assigned the survival of a clinically diagnosed stage IV cancer.

Besides positive health effects of screening, the model also allows for the evaluation of colonoscopy-related complications and over-diagnosis and over-treatment of CRC (i.e. the detection and treatment of cancers that would not have been diagnosed without screening).

**Integrating modules**

The demography module generates a date of birth and a date of non-CRC death for each individual simulated, creating a life-history without adenomas or CRC. Then, in Patient A in eFigure 5, the natural history module generates an adenoma. This adenoma progresses into preclinical cancer, which, in the absence of screening, is diagnosed because of symptoms in stage II and results in CRC death before non-CRC death would have occurred. In the screening module a screening examination is simulated, indicated by the black arrow. During this examination the
adenoma is detected, and as a result both CRC and CRC death are prevented. Hence, integrating all 3 modules for Patient A, screening prolongs life by the amount indicated by the blue arrow. Patient B also develops an adenoma, and although this adenoma does progress into preclinical cancer, Patient B would never have been diagnosed with CRC in a scenario without screening (see Life history 2). However, during the screening examination simulated in the screening module, indicated by the red arrow, CRC is screen-detected in stage I. Hence, in this patient screening results in over-diagnosis of CRC: it detects a cancer that would never have been diagnosed in a scenario without screening. Hence, integrating all 3 modules in this patient, screening does not prolong life but results in additional LYs with CRC care (over-treatment), as indicated by the red arrow.

Model application: simulating the Kaiser Permanente Northern California member population

To derive point estimates of per-lesion sensitivity of colonoscopy for each quintile of adenoma detection in the Kaiser Permanente Northern California (KPNC) data, we simulated the screened populations in each quintile in terms of the age distribution at the time of screening. Population size was inflated in the model to 1 million lives per adenoma detection rate (ADR) quintile to reduce random variability in model outcomes. Two main simplifications were that: (1) although we simulated the age distribution of patients per ADR quintile, inter-provider differences in terms of patient risk factors such as age and sex were assumed to be negligible. Thus, apart from the different age distributions per ADR quintile, all simulated patients were selected randomly from an average-risk US population; (2) it was assumed that patients did not get screened previously, whereas the data included some individuals with a negative prior colorectal cancer test (≥10 years ago). Any misclassification was assumed to be non-differential given random assignment to each ADR quintile.

To validate the model including the point estimates for colonoscopy sensitivity in terms of the predicted interval cancer incidence after screening, we also simulated the follow-up time as included in the KPNC data. Because the incidence rate is variable over time and depends on whether a person had adenomas detected at baseline, we exactly replicated the person-years of follow-up after 1, 2, …, 10 years, stratifying patients with a positive and negative baseline colonoscopy (for adenomas). Because the interval cancers in the data included cancers detected by opportunistic screening or surveillance colonoscopies, we also simulated the proportion of patients with a repeat colonoscopy in years 1, 2, …, 10.

The 5 different sets of parameters for per-lesion sensitivity by polyp size were derived to reproduce the average ADR for each quintile. The parameters were constrained by assuming that: (1) sensitivity for cancer was 98% across all quintiles; (2) sensitivity for medium to large adenomas varied less than for small adenomas, and increased according to a fixed rule from the lowest to the highest quintile (fixed detection likelihood (sensitivity/[1-sensitivity]) ratios for adjacent quintiles) while matching estimates for average practice in the middle quintile (85% for medium adenomas, 95% for large adenomas);18 (3) maximum sensitivity for adenomas was 98%. Sensitivity for adenomas was then varied to match ADR values with 0.1 point precision. The estimates were independent of adenoma location. From the lowest to the highest ADR quintile, resultant sensitivity was 14.7% in quintile 1, 41.0% in quintile 3 and 98% in quintile 5 for small adenomas, 39.6 to 98% for medium adenomas, and 88.0 to 98% for large adenomas (see Table 2 for estimates per ADR quintile).
KPNC data on cancer diagnoses after colonoscopy were compared to the cancer incidence predicted by the model. The model closely reproduced observed incidence in the lower four ADR quintiles, but underestimated incidence in the upper quintile (eFigure 7-8).
eFigure 1. The stages of disease in the semi-Markov model

- No Lesion
- Adenoma ≤ 5 mm
- Adenoma 6-9 mm
- Adenoma ≥ 10 mm
- Preclinical stage I
- Preclinical stage II
- Preclinical stage III
- Preclinical stage IV
- Clinical stage I
- Clinical stage II
- Clinical stage III
- Clinical stage IV
- Cancer death

* Cancer stages were based on the 5th edition Cancer Staging Manual from the American Joint Committee on Cancer.¹⁹
eFigure 2. Simulated versus observed adenoma prevalence in selected autopsy studies (with 95% confidence intervals)a

Observed results are only shown for the two largest studies on which the model has been calibrated. MISCAN-colon has additionally been calibrated to 8 other autopsy studies.
eFigure 3a-e. Simulated versus observed colorectal cancer incidence in 1975-1979 Surveillance Epidemiology and End Results program data.
Figure 4a-b. Simulated versus observed distal colorectal cancer incidence in the intervention group of the UK Flexible Sigmoidoscopy Trial.
PATIENT A: BENEFIT FROM SCREENING
PATIENT B: OVER-DIAGNOSIS FROM SCREENING

Demography module

*Life history 1*
(without CRC)

| Birth | non-CRC death |

Natural history module

*Life history 2*
(developing CRC)

| Birth | Adenoma ≤ 5 mm | Adenoma 6-9 mm | Adenoma ≥ 10 mm | Preclinical CRC stage I | Preclinical CRC stage II | non-CRC death |

Screening module

*Life history 3*
(with screening)

| Birth | Adenoma ≤ 5 mm | Adenoma 6-9 mm | Adenoma ≥ 10 mm | Preclinical CRC stage I | Screen detected | CRC stage I | non-CRC death |

Screening detects CRC life-years with overtreatment
eFigure 6. Bootstrap analysis for average cancer incidence and adenoma detection rates at Kaiser Permanente Northern California.\textsuperscript{a}

\textsuperscript{a} We performed a parametric bootstrap analysis for average observed adenoma detection and incidence rates per ADR quintile (100,000 scenarios, 10,000 shown). Incidence was varied along the lognormal distribution (with Poisson standard errors) and adenoma detection was varied along the normal distribution (with binomial standard errors). Weak and strong association scenarios represent the resulting 2.5\textsuperscript{th} (average 2.4-6\textsuperscript{th}) and 97.5\textsuperscript{th} (97.4-6\textsuperscript{th}) percentile of bootstrap scenarios in terms of the linear regression coefficient for incidence to ADR.
eFigure 7a-b. Simulated versus observed average cancer incidence and adenoma detection rates at Kaiser Permanente Northern California.

© 2015 American Medical Association. All rights reserved.
Abbreviations: Ad. = Adenoma; prev. = prevalence; assoc. = association.

* In the base-case model the adenoma prevalence of 37% was insufficient to reproduce ADR levels observed for the upper quintile in the KPNC data (the curve stops below 37% due to imperfect sensitivity). The red scenario with higher simulated adenoma prevalence reproduced the observed ADR level and led to similar overall results as the base-case model (not shown).
eFigure 8a-e. Simulated versus observed cancer incidence rates (cumulative) in the Kaiser Permanente Northern California data.
### eTable 1. Sensitivity analysis results: The adenoma detection rate-outcome relationship for various modeling scenarios.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Average outcome difference per 5 percentage-point higher adenoma detection rate a</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cancer cases</td>
</tr>
<tr>
<td></td>
<td>Mean 95%CI</td>
</tr>
<tr>
<td>1. Base-case</td>
<td>Rel., %</td>
</tr>
<tr>
<td></td>
<td>Abs.</td>
</tr>
<tr>
<td>2. ADR-variation attr. to adenoma ≤5mm b</td>
<td>Rel., %</td>
</tr>
<tr>
<td></td>
<td>Abs.</td>
</tr>
<tr>
<td>3. ADR-variation attr. to adenoma of all sizes b</td>
<td>Rel., %</td>
</tr>
<tr>
<td></td>
<td>Abs.</td>
</tr>
<tr>
<td>4. ADR-variation attr. to exam completion b</td>
<td>Rel., %</td>
</tr>
<tr>
<td></td>
<td>Abs.</td>
</tr>
<tr>
<td>5. ADR-variation attr. to adenoma prevalence b</td>
<td>Rel., %</td>
</tr>
<tr>
<td></td>
<td>Abs.</td>
</tr>
<tr>
<td>6. Intensified surveillance c</td>
<td>Rel., %</td>
</tr>
<tr>
<td></td>
<td>Abs.</td>
</tr>
<tr>
<td>7. No surveillance</td>
<td>Rel., %</td>
</tr>
<tr>
<td></td>
<td>Abs.</td>
</tr>
<tr>
<td>8. Colonoscopy costs +50%</td>
<td>Rel., %</td>
</tr>
<tr>
<td></td>
<td>Abs.</td>
</tr>
<tr>
<td>9. No discounting</td>
<td>Rel., %</td>
</tr>
<tr>
<td></td>
<td>Abs.</td>
</tr>
</tbody>
</table>

Abbreviations: ADR = adenoma detection rate; CI = confidence interval; Rel. = relative; Abs. = Absolute; attr. = attributable.

a Relative outcomes differences were estimated by linear regression and are compared to the lower ADR quintile. Absolute differences are presented as risk/number per 1,000 adults. Absolute cost differences are in US $ million.

b We evaluated four alternative causal models for the observed ADR differences across the quintiles: in scenario 2 all variation in ADR was attributed to sensitivity of colonoscopy for small adenomas under 5 mm, which varied from 5.4 in the lowest quintile to 98% in the highest quintile; in scenario 3 all ADR variation was attributed equally to sensitivity for small, medium and large adenomas, which varied from 26.0 to 98%; in scenario 4 it was assumed that the rate of completeness of colonoscopy along with differences in colonoscopy sensitivity accounted for the observed ADR-variations, varying from 75% to 98%; in scenario 4 adenoma prevalence was assumed to be up to a relative 25% higher with higher ADR.

© 2015 American Medical Association. All rights reserved.
Under intensified surveillance, we assumed that all patients with adenomas detected at colonoscopy underwent surveillance at 3 years after the procedure, and patients with a negative surveillance colonoscopy underwent surveillance at 5 years. For reference, in the base-case analysis, patients with adenomas detected at colonoscopy were referred for surveillance after 3 or 5 years, depending on the number and size of the adenomas detected. Likewise, patients with a negative surveillance colonoscopy were referred for a follow-up colonoscopy in 5 or 10 years, depending on whether the preceding interval was 3 or 5 years.
eTable 2. Modeled results: Effectiveness of screening colonoscopy according to quintile of adenoma detection rate (0% discounted).  

<table>
<thead>
<tr>
<th>Lifetime health outcomes</th>
<th>No screening</th>
<th>Screening; Quintiles of adenoma detection rate</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean 95%CI</td>
<td>Mean 95%CI</td>
<td>Mean 95%CI</td>
<td>Mean 95%CI</td>
<td>Mean 95%CI</td>
<td>Mean 95%CI</td>
</tr>
<tr>
<td>per 1,000 patients</td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Colorectal cancer outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colorectal cancer cases</td>
<td>66.8</td>
<td>(50.7-85.1)</td>
<td>48.1</td>
<td>(36.1-62.2)</td>
<td>38.8</td>
<td>(28.9-49.9)</td>
<td>32.9</td>
</tr>
<tr>
<td>Advanced cancer cases</td>
<td>32.5</td>
<td>(23.7-43.6)</td>
<td>13.4</td>
<td>(9.6-18.4)</td>
<td>10.2</td>
<td>(7.2-13.8)</td>
<td>8.3</td>
</tr>
<tr>
<td>Colorectal cancer deaths</td>
<td>27.8</td>
<td>(20.8-36.5)</td>
<td>11.8</td>
<td>(8.6-15.8)</td>
<td>9.0</td>
<td>(6.5-12)</td>
<td>7.5</td>
</tr>
<tr>
<td>Years of life lost</td>
<td>324.6</td>
<td>(242-429.1)</td>
<td>141.6</td>
<td>(102.5-190.9)</td>
<td>112.1</td>
<td>(80.9-149.9)</td>
<td>94.6</td>
</tr>
<tr>
<td>Effectiveness of screening</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prevented cancer cases</td>
<td>-</td>
<td></td>
<td>18.7</td>
<td>(13.7-24.5)</td>
<td>28.0</td>
<td>(21.3-36.3)</td>
<td>33.8</td>
</tr>
<tr>
<td>Prevented cancers deaths</td>
<td>-</td>
<td></td>
<td>16.0</td>
<td>(12.1-20.9)</td>
<td>18.8</td>
<td>(14.1-24.3)</td>
<td>20.3</td>
</tr>
<tr>
<td>Years of life saved</td>
<td>-</td>
<td></td>
<td>183.1</td>
<td>(137.7-241.1)</td>
<td>212.5</td>
<td>(158.2-277.1)</td>
<td>230.0</td>
</tr>
</tbody>
</table>

Abbreviations: CI = confidence interval.

a Unlike the base-case results in Table 3 of the manuscript, these results were not discounted and represent actual expected lifetime benefits of colonoscopy screening.

b Adenoma detection rate (ADR) quintiles were derived from 57,588 colonoscopies performed by 136 gastroenterologists in Kaiser Permanente Northern California, a large integrated healthcare delivery system in the United States. The averages (and ranges) of ADR for quintiles 1 through 5 were 15.32% (7.35-19.05%), 21.27% (19.06-23.85%), 25.61% (23.86-28.40%), 30.89% (28.41-33.50%) and 38.66% (33.51-52.51%), respectively.
**eTable 3. Modeled results: Resources and complications for colonoscopy screening according to quintile of adenoma detection rate (0% discounted).** a b

<table>
<thead>
<tr>
<th>Resources per 1,000 patients</th>
<th>No screening</th>
<th>Screening; Quintiles of adenoma detection rate</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean 95%CI</td>
<td>Mean 95%CI</td>
<td>Mean 95%CI</td>
<td>Mean 95%CI</td>
<td>Mean 95%CI</td>
<td>Mean 95%CI</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Screening resources used**

- Total colonoscopies
- Screening colonoscopies
- Surveillance colonoscopies
  - Mean 95%CI: 1.229 (0.932-1.548), 1.612 (1.231-2.027), 1.835 (1.404-2.288), 2.160 (1.644-2.704), 2.426 (1.841-3.015)
- Colonoscopies with polypectomy (screening and surveillance)
  - Mean 95%CI: 1.309 (1.028-1.574), 1.601 (1.276-1.894), 1.759 (1.407-2.064), 1.976 (1.601-2.305), 2.132 (1.728-2.478)

**Serious GI-related complications**

- Mean 95%CI: 10.9 (7.3-15.4), 13.4 (9.2-18.5), 14.5 (9.9-19.7), 15.5 (10.7-21.2), 15.8 (10.8-21.5)

**Financial resources used (US $ million)**

- Total medical costs
  - Mean 95%CI: 6.5 (4.9-8.3), 8.6 (7.1-10.3), 8.1 (6.8-9.6), 7.8 (6.5-9.2), 7.5 (6.3-8.9), 7.3 (6.1-8.5)
- Screening costs
  - Mean 95%CI: 3.8 (3.3-4.2), 4.2 (3.6-4.7), 4.4 (3.8-5.1), 4.8 (4.1-5.5), 5.1 (4.3-5.8)
- Colonoscopy costs
  - Mean 95%CI: 3.7 (3.2-4.2), 4.1 (3.6-4.7), 4.3 (3.7-4.9), 4.7 (4.5-5.4), 5.0 (4.2-5.7)
- Complication costs
  - Mean 95%CI: 0.1 (0.0-0.1), 0.1 (0.1-0.1), 0.1 (0.1-0.1), 0.1 (0.1-0.1), 0.1 (0.1-0.1)
- Treatment costs
  - Mean 95%CI: 4.9 (3.6-6.3), 4.0 (2.9-5.2), 3.4 (2.5-4.5), 2.8 (2.3-3.7), 2.2 (1.6-2.9)
- Net screening costs
  - Mean 95%CI: 2.1 (1.6-2.6), 1.7 (1.2-2.3), 1.3 (0.5-2.1), 1.1 (0.2-1.9), 0.8 (-0.2-1.7)

**Abbreviations:** GI = gastrointestinal; CI = confidence interval; n.a. = not assessed.

a Unlike the base-case results in Table 4 of the manuscript, these results were not discounted and represent actual expected lifetime resources used in colonoscopy screening.

b Adenoma detection rate (ADR) quintiles were derived from 57,588 colonoscopies performed by 136 gastroenterologists in Kaiser Permanente Northern California, a large integrated healthcare delivery system in the United States. The averages (and ranges) of ADR for quintiles 1 through 5 were 15.32% (7.35-19.05%), 21.27% (19.06-23.85%), 25.61% (23.86-28.40%), 30.89% (28.41-33.50%) and 38.66% (33.51-52.51%), respectively.

© 2015 American Medical Association. All rights reserved.
eReferences


© 2015 American Medical Association. All rights reserved.