VALUE IN HEALTH ▮ (2018) ▮▮▮–▮▮▮

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

Critical Analysis of Markov Models Used for the Economic Evaluation of Colorectal Cancer Screening: A Systematic Review

Nicolas Silva-Illanes, MD, MPP, MPH[1,*], *Manuel Espinoza, MD, MSc, PhD*[2]

[1]*School of Public Health, University of Chile, Santiago, Chile;* [2]*HTA Unit, Centre for Clinical Research UC, Pontifical Catholic University of Chile, Santiago, Chile*

ABSTRACT

Background: The economic evaluation of colorectal cancer screening is challenging because of the need to model the underlying unobservable natural history of the disease. **Objectives:** To describe the available Markov models and to critically analyze their main structural assumptions. **Methods:** A systematic search was performed in eight relevant databases (MEDLINE, Embase, Econlit, National Health Service Economic Evaluation Database, Health Economic Evaluations Database, Health Technology Assessment database, Cost-Effective Analysis Registry, and European Network of Health Economics Evaluation Databases), identifying 34 models that met the inclusion criteria. A comparative analysis of model structure and parameterization was conducted using two checklists and guidelines for cost-effectiveness screening models. **Results:** Two modeling techniques were identified. One strategy used a Markov model to reproduce the natural history of the disease and an overlaying model that reproduced the screening process, whereas the other used a single model to represent a screening program. Most of the studies included only adenoma-carcinoma sequences, a few included de novo cancer, and none included the serrated pathway. Parameterization of adenoma dwell time, sojourn time, and surveillance differed between studies, and there was a lack of validation and statistical calibration against local epidemiological data. Most of the studies analyzed failed to perform an adequate literature review and synthesis of diagnostic accuracy properties of the screening tests modeled. **Conclusions:** Several strategies to model colorectal cancer screening have been developed, but many challenges remain to adequately represent the natural history of the disease and the screening process. Structural uncertainty analysis could be a useful strategy for understanding the impact of the assumptions of different models on cost-effectiveness results.

Keywords: colorectal cancer, economic evaluation, Markov models, screening, structural uncertainty.

Introduction

Screening is an essential strategy for the secondary prevention of colorectal cancer (CRC). Several screening modalities are available, including colonoscopy (COL), rectosigmoidoscopy, virtual colonoscopy, guaiac-based fecal occult blood testing (gFOBT), fecal immunochemical testing (FIT), stool DNA testing, and capsule endoscopy.

Over the past two decades, several systematic reviews have concluded that CRC screening is a cost-effective intervention [1–4]. Nevertheless, the studies disagreed as to which screening strategy is most cost-effective or has the best incremental cost-effectiveness ratio (ICER) for a given cost-effectiveness threshold or willingness to pay.

The natural history of CRC is a process much more complicated than initially thought. Until the past two decades it was known that most colorectal adenocarcinomas originated from adenomas, through the adenoma-carcinoma sequence, or were "de novo," without a pre-existing lesion [5]. Recently, it has been identified that serrated lesions, initially considered as hyperplastic polyps without malignant potential, could be the precursors of up to one-third of CRCs and the cause behind some cancers initially considered de novo [6–8].

One particular challenge associated with the economic evaluation of CRC screening is that disease modeling requires accounting for many parameters on the natural progression of potentially malignant lesions that are not directly observable (so-called *deep parameters*) [9–11]. The main deep parameters in the natural history of CRC are adenoma *dwell time* (time from the adenoma incidence to its transformation into asymptomatic CRC) and CRC *sojourn time* (time from the onset of preclinical or asymptomatic CRC to its transition to symptomatic cancer and detection). Both parameters are random variables with an unknown distribution in the population. In turn, the sojourn time and the screening diagnostic test accuracy determine the *lead time* (the time during which screening advances the diagnosis compared with no screening) and, consequently, change the CRC stage distribution and determine the improvement in prognosis.

* *Address correspondence to:* Nicolás Silva-Illanes, School of Public Health, University of Chile, Independencia Avenue 939, Metropolitan Region, Santiago 8380453, Chile.

E-mail: nicosilva@uchile.cl

1098-3015$36.00 – see front matter

https://doi.org/10.1016/j.jval.2017.11.010

The sensitivity of screening tests is also a deep parameter that is difficult to measure and could be defined as a function of the progression of preclinical lesions [11]. Conversely, surface parameters are directly observable parameters, for example, screening participation, CRC survival, and death from other causes [9].

Similarly, models could be classified as *surface models* if they consider only observable events such as CRC incidence, prevalence, and mortality, whereas *deep models* incorporate the hypothesis about the disease process and the underlying disease dynamics that generate the observable events [12].

Previous efforts to characterize different modeling strategies are available [13,14]; these studies, however, do not necessarily provide an in-depth evaluation of the model structure, assumptions, and parameterization. Likewise, collaborative efforts among groups of modeling experts and consortiums of investigators have produced in-depth comparative evaluations of various CRC screening models [15–17]. Nevertheless, these efforts have been focused on microsimulation models.

Even though Markov models have been widely used to simulate CRC screening, to our knowledge, there have not been previous reviews focused on this modeling technique. This study provides a systematic review of the Cohort state transition models that have been used for the economic evaluation of CRC screening, with the aim of describing and analyzing the modeling strategies and their main structural assumptions. This review could be used to inform future cost-effectiveness studies as well as to identify possible sources of structural uncertainty between models.

Methods

A systematic literature review was performed. The inclusion criteria were as follows: 1) full economic evaluations (including cost-effectiveness, cost-utility, and cost-benefit studies); 2) comparing any CRC screening technique(s), and 3) using a Markov model applied to the general population or individuals with normal risk. The search was conducted using the following databases: MEDLINE, Embase, EconLit, National Health Service Economic Evaluation Database, Database of Abstracts of Reviews of Effects, and Health Technology Assessment database. Articles were limited to original reports published in English from 1990 to December 2015. We extended the search strategy to specific journals. Additional articles were identified through the references of the articles reviewed in full text and previous reviews. The full electronic search strategy is included in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2017.11.010.

A three-step selection process was performed. First, duplicates were removed, and clearly irrelevant studies were excluded on the basis of their titles. Second, abstracts were screened on the basis of the inclusion criteria. Finally, full-text copies of the remaining articles were obtained, and a third screening was performed to determine their eligibility. Two researchers screened abstracts and reviewed full-text copies independently. For those studies that used a previously published model, only the original model was considered. In cases in which an updated version of a model was developed by the same group of authors, the article that described the model the most completely was analyzed.

On the basis of previous definitions [12], two modeling strategies were identified. A deep model strategy and a surface model strategy. The model structures were analyzed according to three main modeling dimensions: the screening process, the modeling of deep parameters, and the clinical benefit of screening. After this process, the structure of those models sharing similar characteristics was reproduced using diagrams.

The parameterization and other features of the models were analyzed following the good practices checklist proposed by Karnon et al. [14] for cost-utility modeling of screening programs. Several dimensions used in the comparative workshop carried out by the Institute of Medicine were also included [18]. The costs of screening tests and CRC treatment and the results of each model in terms of incremental costs and ICERs were converted into 2016 US dollars and adjusted according to purchasing power parity using currency conversions from the International Monetary Fund database.

The information obtained is presented as a narrative synthesis and several comparative tables.

Results

The process of study selection is displayed in Figure 1. Overall, the search yielded 1730 hits. Title and abstract screening identified 163 articles for full-text assessment. After full-text analysis, a total of 34 models were included in the review. The main characteristics of the studies, the states and routes modeled, the screening test evaluated, screening age band, perspective, and time horizon are presented in Table 1.

General Characteristics and Modeling Strategies

Regarding the route of carcinogenesis, all models (n = 34) focused on the adenoma-carcinoma sequence, only eight included the de novo cancer pathway, and none included the serrated pathway (Table 1). Two studies also explicitly included lifetime latent cancers that will never be detected by symptoms and have no impact on survival. Adenoma regression was not included in any study, but some models included progressive and nonprogressive adenomas.

Regarding the screening process, the deep model strategy (n = 27 of 34 [79%]) used two superimposed models: a Markov model to reproduce the natural history of CRC and a second model that reproduced the screening protocol (Fig. 2) [19–46]. The surface model strategy (n = 7 of 34 [21%]) used a single Markov model to represent a CRC screening program (Fig. 3) [47–53]. A comparison of the main characteristics of both strategies is presented in Table 2. (A detailed analysis of the general characteristics and modeling strategies is provided in Supplemental Materials.)

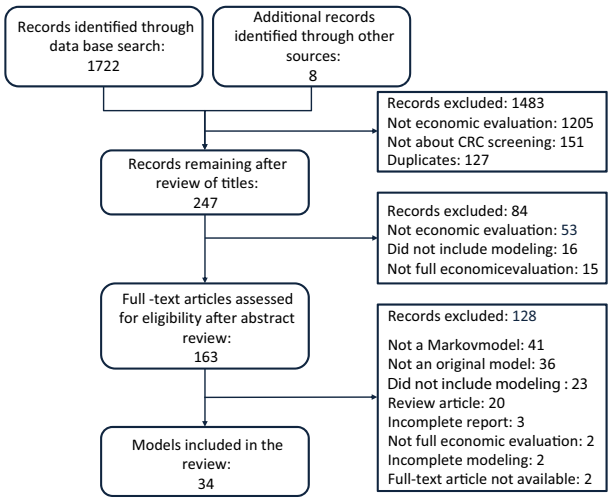

Fig. 1 – PRISMA diagram. CRC, colorectal cancer; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

Table 1 – Main characteristics of the reviewed studies.

Model (first author), country	Perspective/time horizon/age at screening and surveillance	Screening test(s) evaluated	Type of model	Multiple lesions modeling strategy, adenoma state(s) modeled	Cancer state(s) modeled	Modeling strategy for each segment of the colon	Route(s) modeled
Barouni et al. [21], Iran	• Third-party payer • Lifetime • 50–75 y	• gFOBT • FIT • RS • COL • Barium enema • CTC	Deep model	• Index lesion was modeled • Adenoma – Low-risk – Advanced (>9 mm and/or villous component and/or high-grade dysplasia)	• Cancer (undiagnosed and diagnosed) – Localized – Regional – Distant	Not differentiated	Adenoma-carcinoma sequence: 100%
Berchi et al. [40], France	• Health system • Lifetime • 50–74 y	• gFOBT • FIT	Deep model	• Index lesion was modeled • Adenoma – <1 cm – >1 cm	• Diagnosed cancer – Dukes A – Dukes B – Dukes C – Metastasis – Follow-up	Not differentiated	Adenoma-carcinoma sequence: 100%
Bishop et al. [46], Australia	• Third-party payer • Lifetime • 55–75 y	• FIT	Deep model	• Index lesion was modeled • Benign polyp • Large but nonprogressive adenoma • Large, progressive adenoma	• Cancer (undiagnosed and diagnosed) – Dukes A – Dukes B – Dukes C – Dukes D	Not differentiated	Adenoma-carcinoma sequence: 100%
Chen et al. [22], Taiwan	• Societal • Time horizon not defined • Cohort beginning at 50 y, end age not defined	• FIT	Deep model	• Index lesion was modeled • PCDP • Treated PCDP	Diagnosed cancer	Not differentiated	Preclinical screen-detectable–clinical cancer sequence: 100%
Dan et al. [47], Singapore	• Societal • Lifetime • 50–75 y	• FIT • RS • COL • Barium enema • Fecal DNA testing • CTC	Surface model	• Index lesion was modeled • Polyp	Diagnosed cancer	Not differentiated	Adenoma-carcinoma sequence: 100%
Di Bidino et al. [34], Italy	• Health system • Lifetime • Cohort beginning at 50 y, end age not defined	• gFOBT • RS • COL • Barium enema • CTC	Deep model	• Index lesion was modeled • Adenoma – <1 cm – >1 cm	• Diagnosed cancer – Dukes A – Dukes B – Dukes C – Dukes D – Follow-up	Not differentiated	Adenoma-carcinoma sequence: 100%
Eddy [23], United States	• Not clearly defined • Lifetime • 50–75 y	• gFOBT • RS • COL • Barium enema	Deep model	• Index lesion was modeled • Adenoma – <1 cm – >1 cm	• Cancer (undiagnosed and diagnosed) – Dukes A – Dukes B – Dukes C	Not differentiated	Adenoma-carcinoma sequence: 93% De novo cancer: 7%

continued on next page

Table 1 – *continued*

Model (first author), country	Perspective/time horizon/age at screening and surveillance	Screening test(s) evaluated	Type of model	Multiple lesions modeling strategy, adenoma state(s) modeled	Cancer state(s) modeled	Modeling strategy for each segment of the colon	Route(s) modeled
Frazier et al. [24], United States	• Societal • Lifetime • 50–85 y	• gFOBT • RS • COL • Barium enema	Deep model	• Index lesion was modeled • Nonadenomatous polyp • Adenoma – Low-risk – High-risk	• Cancer (undiagnosed and diagnosed) – Localized – Regional – Distant	The model included distal and proximal lesions separately	Adenoma-carcinoma sequence: 100%
Gyrd-Hansen et al. [48], Denmark	• Health system • Lifetime • 50–74 y	• gFOBT	Surface model	• Index lesion was modeled • Adenoma – <1 cm – >1 cm	Diagnosed cancer	Not differentiated	Adenoma-carcinoma sequence It is unclear whether de novo cancer was also modeled
Hassan et al. [25], United States	• Societal • Lifetime • 50–75 y	• gFOBT • FIT • RS • COL • FUSE • CTC • CE	Deep model	• Index lesion was modeled • Hyperplastic polyp • Adenoma (advanced and nonadvanced) – ≤5 mm – 6–9 mm – >1 cm	• Cancer (undiagnosed and diagnosed) – Localized – Regional – Distant	Not differentiated	Adenoma-carcinoma sequence: 90% De novo cancer: 10%
Heitman et al.* [41], Canada	• Health system; publicly funded • Lifetime • 50–75 y	• gFOBT • FIT • RS • COL • CTC	Deep model	• Index lesion was modeled • Adenoma (advanced and nonadvanced) – <1 cm – >1 cm	• Diagnosed cancer – Stage I – Stage II – Stage III – Stage IV	Not differentiated	Adenoma-carcinoma sequence: 100%
Heresbach et al. [36], France	• Third-party payer • 30 y • 50–74 y	• gFOBT • FIT • COL • CTC	Deep model	• Index lesion was modeled • Adenoma (<6 mm to >6 mm) – Low-risk – Advanced (villous component, high-grade dysplasia, carcinoma in situ, or >1 cm)	• Diagnosed cancer – Stage I – Stage II – Stage III – Stage IV	Not differentiated	Adenoma-carcinoma sequence: 100%
Huang et al. [49], China	• Third-party payer • 35 y • 50–70 y	• gFOBT	Surface model	• Index lesion was modeled • Polyp	Diagnosed cancer	Not differentiated	Adenoma-carcinoma sequence: 100%
Khandker et al. [27], United States	• Third-party payer • Lifetime • 50–85 y	• gFOBT • RS • COL • Barium enema	Deep model	• Index lesion was modeled • Hyperplastic polyp • Adenoma – Small – Large	• Cancer (undiagnosed and diagnosed) – Local – Regional – Distant	The model included distal and proximal lesions separately	Adenoma-carcinoma sequence: 100%

continued on next page

Lejeune et al. [50], France	• French social security system • Lifetime • 50–85 y	• gFOBT • FIT	Surface model	• Index lesion was modeled • Advanced adenoma (>1 cm, villous component, or severe dysplasia)	Diagnosed cancer	Not differentiated	Adenoma-carcinoma sequence It is not described whether de novo cancer was also modeled
Macafee et al. [39], England and Wales	• Health care provider • 50 y • 60–69 y	• gFOBT	Deep model	• Index lesion was modeled • Adenoma – Low-risk – High-risk	• Cancer (undiagnosed, diagnosed, and surveillance) – Early – Regional – Advanced	Not differentiated	Adenoma-carcinoma sequence: 100%
Neilson and Whynes [42], United Kingdom	• Third-party payer • 90 y • Cohort followed from 40 to 74 y	• gFOBT • RS	Deep model	• Index lesion was modeled • Adenoma – Progressive – Slow progressive growth that will not transit to cancer – Nonprogressive	• Cancer (undiagnosed and diagnosed) – Early – Late	Not differentiated	Adenoma-carcinoma sequence and de novo cancer. It is not described whether the proportions of cancers originated from each route
O'Leary et al. [28], Australia	• Health system, financed by government • 10 y • 50–64 y	• gFOBT • RS • COL	Deep model	• Index lesion was modeled • Adenoma – <10 mm – >10 mm	• Cancer (undiagnosed and diagnosed) – Dukes A – Dukes B – Dukes C – Dukes D – Survivor	Not differentiated	Adenoma-carcinoma sequence and de novo cancer
Park et al. [45], Korea	• Korean national health system • 80 y • Cohort beginning at 50 y	• gFOBT • RS • COL • Barium enema	Deep model	• Index lesion was modeled • Polyp	• Cancer (undiagnosed and diagnosed) – Early – Late – Survivor	Not differentiated	Adenoma-carcinoma sequence and de novo cancer
Pence et al. [37], United States	• Health system, financed by government • Not described • 50–75 y	• COL • Aspirin • Calcium	Deep model	• Index lesion was modeled • Adenoma – Low-risk – High-risk	Diagnosed cancer	Not differentiated	Adenoma-carcinoma sequence It is not described whether de novo cancer was also modeled
Sharp et al. [29], Ireland	• Third-party payer (Health Service Executive) • 100 y • 55–74 y	• gFOBT • FIT • RS	Deep model	• Index lesion was modeled • Adenoma – Low-risk – High-risk	• Cancer (undiagnosed and diagnosed) – Stage I – Stage II – Stage III – Stage IV	The model included distal and proximal lesions separately	Adenoma-carcinoma sequence: 86% De novo cancer: 14%
Shimbo et al. [43], Japan	• Third-party payer • 35 y • Cohort beginning at 40 y	• gFOBT • FIT • RS	Deep model	• Index lesion was modeled • Benign polyp • Polyp with potential to evolve into cancer – Low-risk	• Cancer (undiagnosed and diagnosed) – Dukes A – Dukes B – Dukes C	Not differentiated	Adenoma-carcinoma sequence: 50% De novo cancer: apparently 50%

continued on next page

Table 1 – *continued*

Model (first author), country	Perspective/time horizon/age at screening and surveillance	Screening test(s) evaluated	Type of model	Multiple lesions modeling strategy, adenoma state(s) modeled	Cancer state(s) modeled	Modeling strategy for each segment of the colon	Route(s) modeled
				– High-risk	– Survivor		
Ladabaum et al.* [30], United States	• Not clearly defined • 100 y • 50–80 y	• gFOBT • COL • RS • CTC • Fecal DNA testing	Deep model	• Index lesion was modeled • Adenoma – Small – Large	• Cancer (undiagnosed and diagnosed) – Localized – Regional – Distant	Not differentiated	Adenoma-carcinoma sequence: 85% De novo cancer: apparently 15%
Sonnenberg et al. [51], United States	• Third-party payer • Lifetime • Cohort beginning at 50 y	• gFOBT • COL • RS • CTC	Surface model	• Index lesion was modeled • Polyp	Diagnosed cancer	Not differentiated	Adenoma-carcinoma sequence It is not described whether de novo cancer was also modeled
Tappenden et al. [31], England	• National Health Service (UK) • Lifetime • 50–70 y	• gFOBT • RS	Deep model	• Index lesion was modeled • Adenoma – Low-risk – High-risk	• Cancer (undiagnosed and diagnosed) – Dukes A – Dukes B – Dukes C – Dukes D	The model included distal and proximal lesions separately	Adenoma-carcinoma sequence: 100%
Telford et al. [32], Canada	• Third-party payer • Lifetime • 50–75 y	• gFOBT • FIT • RS • Fecal DNA testing • Barium enema • CTC • COL	Deep model	• Index lesion was modeled • Adenoma – Low-risk – High-risk	• Cancer (undiagnosed and diagnosed) – Localized – Regional – Distant	Not differentiated	Adenoma-carcinoma sequence: 100%
Tsoi et al. [53], Asia	• Not defined; apparently third-party payer • Lifetime • 50–80 y	• gFOBT • RS • COL	Surface model	• Index lesion was modeled • Polyp	Diagnosed cancer	Not differentiated	Adenoma-carcinoma sequence It is not described whether de novo cancer was also modeled
van Rossum et al. [52], The Netherlands	• Third-party payer • Not described • 50–75 y	• gFOBT • FIT	Surface model	• Index lesion was modeled • Advanced adenoma	Diagnosed cancer	Not differentiated	Adenoma-carcinoma sequence It is not described whether de novo cancer was also modeled
Vijan et al. [44], United States	• Third-party payer • 100 y • 50–75 y	• gFOBT • RS • CTC • COL	Deep model	• Index lesion was modeled • Polyp – Hyperplastic – Low-risk – High-risk	• Cancer (undiagnosed and diagnosed) – Localized – Regional – Disseminated – Survivor	Not differentiated	Adenoma-carcinoma sequence: 100%

continued on next page

Wagner et al. [19], United States	• Third-party payer • Lifetime • 65–85 y	• gFOBT • RS	Deep model	• Index lesion was modeled • Adenoma – Progressive – Nonprogressive	• Cancer – Early latent – Early progressive – Late (undiagnosed and diagnosed)	Not differentiated	Adenoma-carcinoma sequence: 57% De novo cancer: 43%
Wagner et al. [20], United States	• Third-party payer • 85 y • 65–85 y	• gFOBT • RS • COL • Barium enema	Deep model	• Index lesion was modeled • Adenomatous polyp	• Cancer (progressive latent[†]) – Initial – Late (undiagnosed and diagnosed)	Not differentiated	Adenoma-carcinoma sequence: 70%. De novo cancer: 30%
Wang et al. [38], China	• Not defined • Not described • 50–80 y	• COL	Deep model	• Index lesion was modeled • Adenoma – Nonadvanced – Advanced (>1 cm, villous component, or severe dysplasia)	• Cancer (screen-detected or detected by usual means) – Early – Early, postcurative resection – Advanced	Not differentiated	Adenoma-carcinoma sequence It is not described whether de novo cancer was also modeled
Whyte et al.[*] [33], England	• National Health Service (UK) • 100 y • 55–74 y	• gFOBT • RS	Deep model	• Index lesion was modeled • Adenoma – Low-risk – High-risk (three adenomas <1 cm or at least one adenoma >1 cm)	• Cancer (undiagnosed and diagnosed) – Dukes A – Dukes B – Dukes C – Dukes D	Not differentiated	Adenoma-carcinoma sequence: 100% De novo cancer is listed, but apparently was not modeled or had a very low probability postcalibration
Wu et al. [26], Taiwan	• Third-party payer • Lifetime • 50–75 y	• gFOBT • COL • RS • Fecal DNA testing	Deep model	• Index lesion was modeled • Adenoma – Small – Large	• Cancer (undiagnosed and diagnosed) – Early – Late	Not differentiated	Adenoma-carcinoma sequence: 100%

CE, capsule endoscopy; COL, colonoscopy; CRC, colorectal cancer; CTC, computed tomographic colonography (or virtual colonoscopy); FIT, fecal immunochemical testing; FUSE, full spectrum endoscopy; gFOBT, guaiac-based fecal occult blood testing; PCDP, preclinical screen-detectable phase; RS, rectosigmoidoscopy.

[*] These models were applied in several studies (see Supplemental Materials for more details).

[†] It is not clearly described whether latent cancers are only initial-state types of cancer.

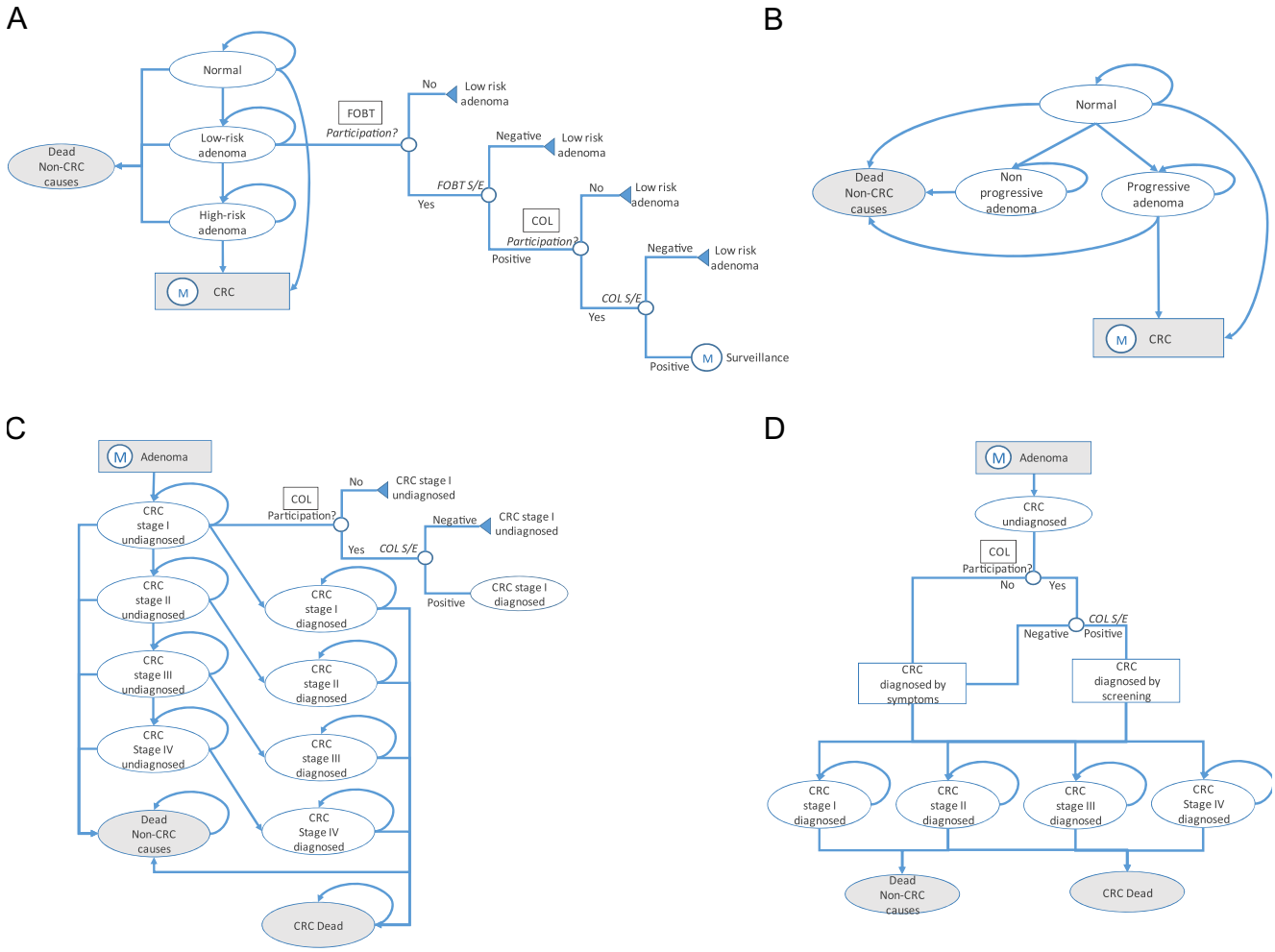

Fig. 2 – Deep model diagrams. Comparison of model structures implemented to simulate adenoma and CRC natural history. COL, colonoscopy; CRC, colorectal cancer; FOBT, fecal occult blood test; M, Markov model; S/E, sensitivity and specificity.

In most of the deep model studies (n = 20 of 27 [74%]), adenoma dwell time was an output of the model [21–41]. In these models, the natural history of adenomas is reproduced and governed by three transition probabilities: adenoma incidence, transition from low-risk to high-risk adenoma, and transition from high-risk adenoma to cancer (Fig. 2A). Alternatively, some models (n = 6 of 27 [22%]) used dwell time as an input of the model [19,20,42–45]. The most common structure of these models includes a nonprogressive adenoma state that will never develop into cancer, and a progressive adenoma state (Fig. 2B). Progressive adenomas transition into cancer according to adenoma dwell time. Few studies (n = 2 of 28 [7%]) did not explicitly describe their model structure [19,20]. All the reviewed studies modeled only index lesions (the more advanced of multiple lesions).

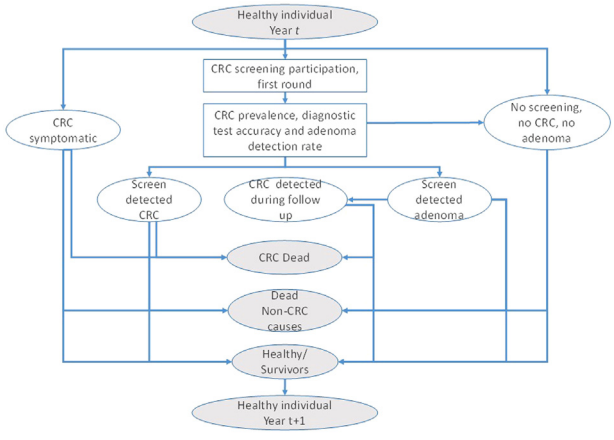

Fig. 3 – Surface model diagram. Model structure used to simulate a CRC screening program. CRC, colorectal cancer.

In most of the deep models (n = 25 of 27 [92%]), the benefit of CRC incidence reduction derived from the early detection of adenomas relied on the interaction of the natural history model and the model that represented the screening process (Fig. 2A) [21,22,24–46]. Adenoma detection depends on the sensitivity of the screening tests and the participation rate. Once an adenoma has been detected, the model assumes that it is resected and subjected to colonoscopic surveillance. Most of these models assumed that adenomas are potentially detectable from their first appearance and that sensitivity is a function of adenoma size and/or histology, but because adenoma growth is not modeled continuously, the probability of detection is a discrete function of each adenoma state. One model assumed that each screening test has a "nonrandom false-negative rate," defined as how long before cancer invasion a test can first detect an adenoma, and a "random false-negative rate" equivalent to the sensitivity of the test [23].

Table 2 – Summary of main approaches to CRC Markov modeling.

Characteristic	Deep models	Surface models
Number of reviewed models	27 (79%)	7 (21%)
Description	Incorporate hypothesis about the disease process and the underlying disease dynamics that generate the observed events	Consider only observable events such as CRC incidence, prevalence, and mortality
Modeling strategy	Two superimposed models: natural history of the disease and screening protocol	A single model that directly represents the screening process
Main strengths	• Are more flexible to evaluate combination of different screening tests • Allow to evaluate different adenoma surveillance strategies • Could be updated according to changes observed in the CRC stage distribution at diagnosis or changes in population cancer survival • Incorporate lead time (or the time by which screening advances the diagnosis compared with absence of screening) in the model structure • Could be used to evaluate screening schemes that incorporate strategies according to polyp size thresholds	• Require fewer parameters than deep models • Do not need to simulate CRC incidence, because it is one of the inputs of the model • Cancer sojourn time and adenoma dwell time are inputs of the model
Limitations	• Require more parameters than surface models • CRC incidence, adenoma prevalence, adenoma dwell time, and cancer sojourn time are outputs of the model that need to be validated	• The analysis of combination of different screening tests is challenging • Do not incorporate cancer lead time in the model structure • Estimate adenoma detection on the basis of positivity rate and positive predictive value rather than on the sensitivity of the screening tests • Updating changes in the cancer stage distribution at diagnosis and changes in population cancer survival is challenging • The evaluation of screening schemes that incorporate strategies according to different polyp size thresholds is challenging
Shared challenges	• It is difficult to incorporate time-dependent transition probabilities in the adenoma and cancer progression • It is difficult to model multiple synchronic lesions	

CRC, colorectal cancer.

There are three strategies to model the clinical benefit of early CRC detection in deep models. Each is based on the assumption that CRC stages are associated with different levels of health-related quality of life and survival rates and therefore clinical benefit depends on the stage distribution of incident (symptomatic) versus screen-detected cases.

Studies with a structure as that displayed in Figure 2C (n = 16 of 27 [59%]) included an undiagnosed state and a diagnosed state for each CRC stage [21–33,42–44,46]. When an undiagnosed case is detected through screening, it is transferred to the diagnosed state, and the progression to more advanced stages is blocked. Sojourn time is an output of the model and is governed by the transition probabilities between undiagnosed CRC stages and the transition probabilities from undiagnosed to diagnosed CRC stages (Fig. 3A), whereas lead time is related to the time advanced in the screen-detected cases. Few studies modeled early- and late-stage CRC, but included a transition from undiagnosed to diagnosed state for only late-stage CRC, although it was assumed that early-stage CRC has a window for detection before entering into late-stage CRC [19,20]. In the other study, the sojourn time modeling strategy was not explicitly described [23].

Models that adopted a structure as shown in Figure 2D (n = 3 of 27 [11%]) did not model CRC progression [40,41,45]. The proportion of patients with CRC in each stage depends on the diagnosis strategy, with more favorable CRC distributions contingent on the screening test evaluated. Sojourn time and lead time could not be estimated from the model because they do not take into account the time that screening advances the diagnosis compared with no screening.

A less common strategy (n = 1 of 27 [4%]) included transitions between CRC stages while assuming that each CRC stage has a without-screening state or a screen-detected state with different survival rates [38]. Few studies (n = 4 of 27 [15%]) modeled CRC progression between stages, but insufficient information is provided regarding sojourn time and lead time because only diagnosed CRC states are modeled [34–37].

The surface model strategy does not reproduce the natural history of adenoma [47–53]. Alternatively, the model predicts the number of adenomas that would be diagnosed through screening in each round of the screening program, and CRC incidence is adjusted downward accordingly.

In these models, the benefit of early CRC detection is characterized in a simpler structure including only two CRC states: CRC

detected by symptoms and CRC detected by screening. Higher survival rates are considered in the former (Fig. 3). Sojourn time and lead time are underlying factors in the CRC prevalence modeled in each screening round. Only few studies (n = 2 of 7 [29%]) reported sojourn time estimations [47,50]. One study (14%) did not describe how the clinical benefit of early CRC detection was modeled [53].

Parameterization of CRC and Adenoma Prevalence

Deep models

Most studies simulated a cohort starting at the age of 50 or 55 years. The initial health state distribution of the cohort requires estimating the adenoma and CRC prevalence at this age.

One strategy to obtain prevalence parameters was to perform estimations directly from the model (n = 3 of 27 [11%]) [29,31,33]. The remaining studies (n = 24 of 27 [89%]) obtained adenoma and CRC prevalence from the literature. Some models (n = 6 of 27 [22%]) used local epidemiological data sources to inform adenoma prevalence [28,34,38,41,44,46] or CRC prevalence [26,28,34,38,41,46]. Few studies (n = 6 of 27 [22%]) performed a systematic search for parameters of adenoma prevalence [19,20,29,31,33,35]. A more detailed description is presented in Table 3.

Surface models

Screening in surface models involved the detection of CRC either by the interaction of the screening test's diagnostic accuracy and CRC prevalence (n = 2 of 7 [29%]) [48,50] or by using the positivity rate and the positive predictive value of the screening tests (n = 2 of 7 [29%]) [47,52]. Few studies (n = 2 of 7 [29%]) estimated CRC prevalence by modeling on the basis of CRC incidence and sojourn time estimates. Some studies (n = 3 of 7 [43%]) described the sensitivity of screening tests for CRC but did not provide sufficient information on how CRC was detected [49,51,53]. Most studies (n = 4 of 7 [57%]) used local epidemiological data to inform their models [48–50,52]. Adenoma prevalence was not required as an input in these studies.

Parameterization of Adenoma Natural History and CRC Progression

Deep models

Most of the studies (n = 12 of 27 [44%]) modeled adenoma incidence using different transition probabilities by age group [19,22,24,26,27,29,31–33,35,36,39] and one study used a time-varying probability function dependent on age [26]. Few models (n = 2 of 27 [7%]) used age-dependent probabilities to model transitions between adenomas and from adenoma to cancer [25,33] or de novo cancer incidence [30]. Progression between CRC stages and diagnosed/undiagnosed states was estimated using either transition probabilities or the average duration of each stage [19,20,23,27,42–46], and no studies included age-dependent transition probabilities in these parameters. Two studies included time-in-state–dependent probabilities. One used a semi-Markov process in which a probability function determined the holding time required to elapse before a movement to the next state [42]. The second study included higher transition probabilities from adenoma to cancer conditional on the time spent in the adenoma state [45]. As has been described elsewhere [14], another study could be considered a combination of an analytical-numerical method and simulation because it used a state transition model with probabilities obtained by numerical integration of a set of differential equations [23]. In this model, adenoma dwell time was a parameter required to estimate the "preclinical interval" or the time required for an adenoma to develop from first being potentially detectable by screening to the appearance of signs and symptoms. Few models (n = 4 of 27 [15%]) performed a systematic search of evidence to inform any of these parameters [19,29,31,33].

Surface models

Instead of modeling the natural history of adenomas and adenoma prevalence, these studies directly modeled adenoma detection rates in each screening round. CRC incidence reduction is modeled assuming that a proportion of future CRC cases are avoided because of adenoma detection, whereas the benefit of CRC early detection is due to improving survival in screen-detected cases. None of these studies performed a systematic literature search to inform these parameters.

Other Model Features

Postpolypectomy surveillance

Among the deep model studies that included more than one type of polyp, most modeled (n = 11 of 24 [46%]) differential surveillance schemes depending on polyp characteristics [21,24,26,27,29,31,32,35,39,41,44], whereas several studies (n = 10 of 24 [42%]) failed to clearly describe the surveillance scheme [22,23,28,40,42,43,47,50,52,53].

Several studies (n = 14 of 34 [41%]) included the possibility of recurrent adenomas after polypectomy [24,26,27,29,31–33,35,36,38,41,44,45,51]. Some surface model studies (n = 2 of 7 [29%]) did not address polyp recurrence, but instead included a probability of developing CRC during the surveillance period after polypectomy [48,50] or considered the number of polypectomies expected during the surveillance period (n = 1 of 7 [14%]) [51].

Costs

Most of the models (n = 17 of 34 [50%]) included the costs of CRC treatment and surveillance during the survivor state [19,22,24,26,27,29–33,35,40,41,43,45,47,50]. Other sources of costs were included only in some studies, such as costs associated with terminal CRC (end-of-life costs) (n = 7 of 34 [21%]) [23,24,26,30,32,43,50], costs associated with CRC recurrence (n = 3 of 34 [9%]) [31,32,38], and administrative costs associated with running the screening program (n = 9 of 34 [26%]) [28,29,33,40,47–50,52]. Few studies (n = 4 of 34 [12%]) also included indirect costs [22,35,41,47]. Costs of CRC treatment vary between models from $1,538 to $43,837 for localized cancer and from $5,453 to $120,131 for late-stage cancer. The costs of screening test also are highly heterogeneous between studies (see Supplemental Materials for a more detailed description).

Utilities

The most common reported outcome referred to life-years gained. From the studies that reported quality-adjusted life-years per CRC stage, two used Health Utility Index scores from a study of CRC survivors [30,47], and three used quality-adjusted life-year values from a study that applied a standard gamble exercise to postpolypectomy patients [31,32,41]. Two studies used both information sources [29,53]. One study used a single utility estimate for CRC that was not specific to CRC stage [33].

CRC survival

Most of the studies (n = 20 of 34 [59%]) used population-based CRC survival data [21–24,26,27,29–36,40,41,44,46,51,52]. Most studies modeled complications associated with endoscopy screening such as bleeding, colon perforation, and death. The main differences between studies were regarding the inclusion of other causes of death besides cancer for CRC states (n = 13 of 34 [38%]) [21–23,26,27,29–31,33,34,36,41,42], consideration of CRC relapse (n = 2 of 34 [6%]) [21,32], and the assumption that patients were

Table 3 – Methods of estimation (or sources of information) used for the transition probabilities of adenoma and CRC natural history.

Parameters	Method of estimation (E) or primary source of information (PS)	Primary studies	Models
Adenoma prevalence and/or cancer prevalence	(PS) Studies of autopsy data	[55–64]	[20,24,25,30,40,44]
	(PS) Cross-sectional studies or randomized controlled trials of low-sensitivity screening tests (FOBT and RS)	[65–67]	[28,34,46]
	(PS) Cross-sectional studies of high-sensitivity screening tests (COL and CTC)	[65–80]	[25,32,35–38,45]
	(PS) Results from CRC screening programs	[73,81,82]	[52]
	(PS) Meta-analysis of cross-sectional COL screening studies	[83]	[41]
	(E) Calibration based on CRC incidence, adenoma prevalence, and the cumulative probability of developing cancer according to studies of in situ polyp surveillance	NA	[26]
	(E) Prevalence of disease accumulated over the prescreening period (i.e., 30–50 y) estimated by filtering the cohort through the health states of the natural history of disease	NA	[29,31,33]
	(E) Modeled on the basis of CRC incidence and sojourn time estimates	NA	[48,50]
Adenoma incidence	(PS) Observational study of patients undergoing repeated colonoscopies; incidence rate was estimated by newly developed adenomas from patients with no lesions on the first examinations	[84]	[36]
	(E) Calibration against adenoma prevalence by age group	NA	[24,30,32,44,45,51]
	(E) Estimation based on statistics of CRC incidence by age group, cumulative risk of developing cancer among adenoma carriers, and average dwell time. The probability of a person developing progressive adenoma was derived by tracking the age-specific incidence of CRC to the likely time of adenoma development	NA	[26,46]
	(E) Simultaneous calibration of all the parameters of the natural history of disease using a Bayesian method	NA	[29,31,33]
Transition between adenomas	(PS) Observational studies on surveillance of patients with polyps left in situ	[56–60,85–88]	[24,35,40]
	(PS) A three-state Markov model based on data from a case-cohort study of subjects who underwent screening colonoscopy and were followed linking patient data to cancer registry data	[89]	[38]
	(E) Calibration against adenoma prevalence by age group	NA	[30,32,41,44]
	(E) Simultaneous calibration of all the parameters of the natural history of disease using a Bayesian method	NA	[29,31,33]
Transition from adenoma to cancer	(PS) Observational studies on surveillance of patients with polyps left in situ	[85]	[24,35,45]
	(PS) Cross-sectional study that estimates transition probabilities combining data from a national screening colonoscopy registry and cancer registry data	[90]	[36]
	(PS) A three-state Markov model based on data from a case-cohort study of subjects who underwent screening colonoscopy and were followed linking patient data to cancer registry data	[89]	[38]
	(E) Calibration against CRC incidence data	NA	[30,32]
	(E) Calibration based on CRC incidence, adenoma prevalence, and the cumulative probability of developing cancer according to studies of in situ polyp surveillance	NA	[26,44]
	(E) Simultaneous calibration of all the parameters of the natural history of disease using a Bayesian method	NA	[29,31,33]
Transition probability between cancer stages	(E) Calibration against CRC stage distribution using data from cancer registries	NA	[23–25,32–34,44,49]
	(PS) Multistage Markov model using data from a CRC screening program	[91]	[22,26]
	(PS) Expert opinion	NA	[18,19,23,27,30,42,43]
	(E) Simultaneous calibration of all the parameters of the natural history of disease using a Bayesian method	NA	[29,31,33]
	(PS) Multistate Markov model using data from CRC screening programs	[91]	[22,26]

continued on next page

Table 3 – *continued*

Parameters	Method of estimation (E) or primary source of information (PS)	Primary studies	Models
Transition probability from asymptomatic to symptomatic CRC	(E) Calibration against CRC stage distribution using data from cancer registries	NA	[24,27,32,44]
	(E) Estimation based on the proportion of incident cases over the cumulative proportion of cases by stage, on the basis of data from cancer registries	NA	[28,46]
	(E) Simultaneous calibration of all the parameters of the natural history of disease using a Bayesian method	NA	[29,31,33]
CRC stage distribution	(PS) Results from a CRC screening program	*	[40]
	(PS) RCT of the efficacy of FOBTs	[92–96]	[40,41]
	(PS) Cross-sectional studies of colonoscopy	[71,73]	[41]
	(PS) Tumor registry	†	[40]
CRC incidence reduction due to adenoma detection	(PS) Surveillance studies of patients with polyps left in situ and surveillance studies of patients with resected adenomas	[85,96]	[48–50,52]
Adenoma detection rate	(PS) RCTs	[93]	[48]
	(PS) Estimated using the positive predictive value of confirmatory colonoscopy and the positivity rate of gFOBT	[97,98]	[50,52]

COL, colonoscopy; CRC, colorectal cancer; CTC, computed tomographic colonography (or virtual colonoscopy); gFOBT, guaiac-based occult blood testing; NA, not applicable; RCT, randomized controlled trial; RS, rectosigmoidoscopy.

* Data obtained from a screening program in Calvados (reference not provided).

† Data obtained from Tumor Registry of Calvados (reference not provided).

cured after a given survival period (n = 15 of 34 [44%]) [19,21,28,30,32–36,40,41,44,46,51,52].

Validation and calibration

Two alternatives for external validation were identified. The first strategy consisted of performing an external validation of model-predicted incidence, mortality, and/or cumulative CRC risk against data from cancer registries [28,31,33,45,50,52]. Another common technique was external validation against the results of clinical trials of screening [23,25,27,38,43,45,50]. Among deep models, only one reported adenoma dwell time and sojourn time predictions [39].

Few models (n = 4 of 34 [12%]) performed cross-validation, adjusting some parameters to replicate other studies and comparing the results between them [24,26,42,48]. Some studies (n = 9 of 34 [26%]) did not perform cross-validation [22,23,27,34,41,44,45,47,49]. The remaining studies simply compared the results of their models with those from other studies.

Calibration was used in several models to estimate the natural history of disease parameters. The most common strategy was manual calibration (i.e., an arbitrary variation of parameters by the analyst to obtain better predictions from the model) [19,21,24,25,27,28,30,44,45]. Only three studies [29,31,33] used a formal calibration method and described the process in terms of the objectives, search algorithm, acceptance criteria, and stopping rules.

Modeling right versus left colon

One model assumed a different proportion of CRC incidence for the left and right segments [33]. Few models (n = 4 of 34 [18%]) included an independent state for each colon segment, assuming a different proportion of lesions for each [24,27,29,31]. No studies used specific transition probabilities for disease progression of each segment of the colon. In addition, some of the aforementioned studies assumed a probability of finding a sentinel polyp in the distal colon that was dependent on the probability of having a cancerous lesion in the proximal colon [24,29,31].

Diagnostic test accuracy

Several studies (n = 10 of 34 [29%]) performed a systematic evidence search for diagnostic test accuracy [19,20,23,29,31,33,40,41,44,46]. One of these studies also relied on expert opinion because of lack of evidence for several parameters [29]. One study used calibration to estimate the sensitivity and specificity parameters, and it was also the only to include differential specificity by age for gFOBT and FIT [33]. Subsequent analysis of the same model also incorporated a decrease in the gFOBT sensitivity between initial and subsequent screening rounds [54].

The diagnostic accuracy parameters are highly heterogeneous between studies. For example, sensitivity of gFOBT to detect CRC varies among studies from 13% to 72%, from 48% to 94% for FIT, from 90% to 100% for COL, from 62% to 97% for sigmoidoscopy, and from 85% to 97% for virtual colonoscopy (see Supplemental Materials for a detailed description of the screening test parameters).

Comparison of Cost-Effectiveness Results and Model Predictions between Modeling Strategies

There is high variability in the results and predictions between models (see Appendix Figures 1–5 in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2017.11.010 for a more detailed description). On average, deep models predict higher incidence reduction for annual gFOBT and annual FIT (36% vs. 22% and 46% vs. 27%, respectively), and SIG every 5 years (43% vs. 33%); higher incidence reduction for once-only SIG and once-only COL (10% vs. 19% and 27% vs. 28%, respectively); and similar results for colonoscopy every 10 years (54%). In terms of CRC

mortality reduction, on average, deep models predict lower estimates for annual gFOBT (46% vs. 18%), annual FIT (58% vs. 25%), biennial FIT (31% vs. 25%), SIG every 5 years (46% vs. 32%), once-only COL (31% vs. 28%), and COL every 10 years (56% vs. 36%); higher estimates for once-only SIG (13% vs. 16%); and similar estimates for biennial gFOBT (16 vs. 17%). In terms of incremental costs and incremental life-years gained compared with no screening, there is no clear pattern between surface and deep models. Nevertheless, the ICER compared with no screening is higher in surface models for all the screening strategies analyzed, with stronger differences in the case of annual FOBT ($66,089 vs. $ 8,286), annual FIT ($28,867 vs. $5,185), once-only SIG ($29,334 vs. $119), once-only COL ($39,526 vs. −$42), and COL every 10 years ($17,890 vs. $4,299). These results need to be interpreted with caution because there was a smaller number of surface models compared with deep models as well as high variability within groups, and there were several parameters besides the model structures that could have a major impact on the results.

Discussion

Following previous definitions, two categories of model structures were identified: deep models and surface models. The deep model strategy was the most common technique used. This strategy provides a more thorough characterization of the disease and screening processes, although it requires the estimation of numerous deep parameters. Surface models, however, require fewer specifications. One restriction of this approach is that the natural history of the adenoma-carcinoma sequence is not modeled directly, thereby limiting the analysis of different screening tools and polyp surveillance protocols. Similarly, the use of different survival rates for screen-detected and incident cases entails the assumption of an underlying CRC stage distribution that is fixed and specific to each screening test, which undermines the flexibility of these models.

Most observational studies used to inform the adenoma natural history had a longitudinal design in which cumulative risk of adenoma progression was estimated through in situ polyp surveillance. As has been noted previously, this study design has some limitations. First, differential attrition during surveillance could introduce bias [15,31]. Second, the rate at which large adenomas become malignant cannot be accurately estimated because it would require information concerning the point at which the polyp becomes incident [15,31].

Some evidence suggests that the positive predictive value of screening test decreases between the initial and repeat screening rounds [99–101]. Failing to incorporate this evidence could overestimate the effectiveness of screening. The only study that incorporates this aspect estimates a decrease in the sensitivity of gFOBT in the repeat screen [54], whereas recent evidence from another modeling study suggests an opposite conclusion for FIT —that the diminishing performance is attributed primarily to the progressive removal of people with neoplasia from the screen-eligible population [102]. Further research in this area is needed.

Failing to incorporate time- or age-dependent parameters in the natural history of the disease could represent a source of bias in several of the analyzed studies. As was described in the results, only two studies included time-in-state–dependent probabilities for adenoma dwell time. This restriction does not allow to model the risk of progression to CRC conditional on the time that an individual has remained in the adenoma state. As has been suggested, if the assumption of constant progression is erroneous, models would favor shorter screening intervals because of an overestimation of the early rate of progression [15].

When studies rely on cross-sectional data, age effects could not be disentangled from birth-cohort effects, because people of different ages are born in different periods. This could be a limitation when the calibration of transition probabilities is based on adenoma prevalence figures. To address this potential confounding, some studies have used birth-cohort analysis based on large screening databases [103,104] that could provide useful information for future cost-effectiveness models.

Current knowledge reveals that CRC is a disease with different biological origins [105]. Thus, failing to include all routes of carcinogenesis, such as the serrated pathway, could lead to an overestimation of screening effectiveness [106]. Evidence also suggests that proximal and distal CRCs could be considered distinct clinicopathological entities [107,108]. Future models should consider including different natural history parameters for each segment of the colon.

Another critical issue is CRC survival modeling. Most of the studies obtained parameters of specific survival or relative survival from population cancer registries. In both cases, this reflects the probability of dying specifically from CRC and, therefore, requires taking into account the risk of death from other causes. Models that failed to include this could have produced a biased estimation of effects.

Calibration is a useful strategy for estimating deep parameters [109,110] and is commonly used in cancer models [111–113]. One of the limitations observed is that most of the studies failed to comply with reporting recommendations [113] and relied on manual calibration techniques.

In terms of validation, it would be helpful if alternative model predictions are reported, such as the number of cancers prevented, cancer stage–specific incidence predictions, as well as adenoma dwell time and cancer sojourn time estimations. This could increase transparency regarding a model's structural assumptions. Another useful measure that has been proposed for this purpose is maximum clinical incidence reduction [15,16], which is the model-simulated reduction in CRC incidence through the end of the cohort follow-up period, at which point it is assumed that detection and management of all preclinical cancer cases are complete.

As presented in Table 3, state transition models face several limitations, some of which could be overcome using microsimulation models. In this type of model, it is easier to incorporate time-dependent transition probabilities in the adenoma and cancer progression. It is also possible to model more accurately different segments of the colon and several synchronic lesions. In terms of validation, because in microsimulation models the history of every person modeled is recorded, underlying model parameters, such as lead time, mean sojourn time, and adenoma dwell time, could be directly estimated. Use of microsimulation models could also be better to analyze screening strategies targeted to different subgroup populations.

There were limitations in our review that should be stated. In terms of the selected studies, only articles written in English were analyzed. In terms of the model analysis, even though a comprehensive characterization of the studies was achieved, because model development is a very complex process, describing all the modeling elements in detail is not feasible. Structural uncertainty could be partially understood from this review, but for a better analysis different models should be reproduced using specific standard assumptions (i.e., expected levels of adherence to the screening, medicals costs, and test performance characteristics) to provide a critical insight into the impact of different model structures.

This last issue is critical because of the diversity of models, assumptions, and data sources used. Collaborative efforts among groups of modeling experts and standardized disease-specific models [114] are promising strategies for improving the consistency and relevance of health decision-making inferences.

Conclusions

Two different strategies and many models have been developed to evaluate the cost effectiveness of CRC screening. Studies have several differences between them in terms of the sources of information and assumptions regarding deep parameters. There are still many challenges associated with adequately representing the natural history of the disease and the screening process and evaluating the structural uncertainty of the models.

Acknowledgments

We thank Dr. Paul Tappenden for his comments that greatly improved the article. We also thank Dr. Óscar Arteaga, Tania Alfaro, and Carla Castillo-Laborde for their comments on the final version of the article. We thank Sage Language Services for providing editing services on the final version of the manuscript.

Source of financial support: Financial support for this study was provided entirely by a research grant from the School of Public Health, University of Chile, and a grant from Chile's National Fund for Health Research and Development (*Fondo Nacional de Investigación y Desarrollo en Salud*, grant no. FONIS SA13I20255). The funding agreement ensured the authors' independence in designing the study, interpreting the data, and writing and publishing the report.

Supplemental Materials

Supplemental material accompanying this article can be found in the online version as a hyperlink at http://dx.doi.org/10.1016/j.jval.2017.11.010 or, if a hard copy of article, at www.valueinhealthjournal.com/issues (select volume, issue, and article).

REFERENCES

[1] Lansdorp-Vogelaar I, Knudsen AB, Brenner H. Cost-effectiveness of colorectal cancer screening. Epidemiol Rev 2011;33:88–100.
[2] Pignone M, Saha S, Hoerger T, Mandelblatt J. Cost-effectiveness analyses of colorectal cancer screening: a systematic review for the U.S. Preventive Services Task Force. Ann Intern Med 2002;137:96–104.
[3] Kittrongsiri K, Chaikledkaew U. Economic evaluation of colorectal cancer screening: a systematic review. J Pharm Sci 2014;41:31–8.
[4] Patel SS, Kilgore ML. Cost effectiveness of colorectal cancer screening strategies. Cancer Control 2015;22:248–58.
[5] Bordacahar B, Barret M, Terris B, et al. Sessile serrated adenoma: from identification to resection. Dig Liver Dis 2015;47:95–102.
[6] Rex D, Ahnen D, Baron J, et al. Serrates lesions of the colorectum: review and recommendations from an expert panel. Am J Gastroenterol 2012;107:1315–30.
[7] Sweetser S, Smyrk T, Sinicrope F. Serrated colon polyps as precursors to colorectal cancer. Clin Gastroenterol Hepatol 2013;11:760–7.
[8] Yamagishi H, Kuroda H, Imai Y, Hiraishi H. Molecular pathogenesis of sporadic colorectal cancers. Chin J Cancer 2016;35:4.
[9] van Oortmarssen GJ, Boer R, Habbema JD. Modelling issues in cancer screening. Stat Methods Med Res 1995;4:33–54.
[10] van Ballegooijen M, Boer R, Zauber AG. Simulation of colorectal cancer screening: what we do and do not know and does it matter. Best Pract Res Clin Gastroenterol 2010;24:427–37.
[11] Prevost TC, Launoy G, Duffy SW, Chen HH. Estimating sensitivity and sojourn time in screening for colorectal cancer: a comparison of statistical approaches. Am J Epidemiol 1998;148:609–19.
[12] Bross IDJ, Blumenson LE, Slack NH, Priore RL. A two disease model for breast cancer. In: Forrest APM, Bunkler PB,eds. Prognostic Factors in Breast Cancer. Baltimore, MD: Williams & Wilkins, 1968. p. 288–300.
[13] Jeong KE, Cairns JA. Review of economic evidence in the prevention and early detection of colorectal cancer. Health Econ Rev 2013;3:20.
[14] Karnon J, Goyder E, Tappenden P, Mcphie S. A review and critique of modelling in prioritising and designing screening programmes. Health Technol Assess 2007;11:iii–v: ix.
[15] Kuntz KM, Lansdorp-vogelaar I, Rutter CM, et al. A systematic comparison of microsimulation models of colorectal cancer: the role of assumptions about adenoma progression. Med Decis Making 2011;31:530–9.
[16] van Ballegooijen M, Rutter CM, Knudsen AB, et al. Clarifying differences in natural history between models of screening: the case of colorectal cancer. Med Decis Making 2011;31:540–9.
[17] Pignone M, Russell L, Wagner J,editors. Economic Models of Colorectal Cancer Screening in Average-Risk Adults. Washington, DC: National Academies Press, 2005.
[18] Wagner J, Russell L, Pignone M,editors. Economic Models of Colorectal Cancer Screening in Average-Risk Adults: Workshop Summary. Washington, DC: National Academies Press, 2005.
[19] Wagner JL, Herdman RC, Wadhwa S. Cost effectiveness of colorectal cancer screening in the elderly. Ann Intern Med 1991;115:807–17.
[20] Wagner J, Tunis S, Brown M. Cost-effectiveness of colorectal cancer screening in average-risk adults. In: Young G, Rozen P, Levin B,eds. Prevention and Early Detection of Colorectal Cancer. London: WB Saunders, 1996. p. 321–56.
[21] Barouni M, Larizadeh MH, Sabermahani A, Ghaderi H. Markov's modeling for screening strategies for colorectal cancer. Asian Pac J Cancer Prev 2012;13:5125–9.
[22] Chen L-S, Liao CS, Chang SH, et al. Cost-effectiveness analysis for determining optimal cut-off of immunochemical faecal occult blood test for population-based colorectal cancer screening (KCIS 16). J Med Screen 2007;14:191–9.
[23] Eddy DM. Screening for colorectal cancer. Ann Intern Med 1990;113:373–84.
[24] Frazier AL, Colditz GA, Fuchs CS, Kuntz KM. Cost-effectiveness of screening for colorectal cancer in the general population. JAMA 2000;284:1954–61.
[25] Hassan C, Pickhardt PJ, Rex DK. A resect and discard strategy would improve cost-effectiveness of colorectal cancer screening. Clin Gastroenterol Hepatol 2010;8:865–9: 869.e1–3.
[26] Wu GH, Wang Y-M, Yen AM, et al. Cost-effectiveness analysis of colorectal cancer screening with stool DNA testing in intermediate-incidence countries. BMC Cancer 2006;6:136.
[27] Khandker RK, Dulski JD, Kilpatrick JB, et al. A decision model and cost-effectiveness analysis of colorectal cancer screening and surveillance guidelines for average-risk adults. Int J Technol Assess Health Care 2000;16:799–810.
[28] O'Leary BA, Olynyk JK, Neville AM, Platell CF. Cost-effectiveness of colorectal cancer screening: comparison of community-based flexible sigmoidoscopy with fecal occult blood testing and colonoscopy. J Gastroenterol Hepatol 2004;19:38–47.
[29] Sharp L, Tilson L, Whyte S, et al. Cost-effectiveness of population-based screening for colorectal cancer: a comparison of guaiac-based faecal occult blood testing, faecal immunochemical testing and flexible sigmoidoscopy. Br J Cancer 2012;106:805–16.
[30] Ladabaum U, Chopra CL, Huang G, et al. Aspirin as an adjunct to screening for prevention of sporadic colorectal cancer: a cost-effectiveness analysis. Ann Intern Med 2001;135:769–81.
[31] Tappenden P, Chilcott J, Eggington S, et al. Option appraisal of population-based colorectal cancer screening programmes in England. Gut 2007;56:677–84.
[32] Telford JJ, Levy AR, Sambrook JC, et al. The cost-effectiveness of screening for colorectal cancer. CMAJ 2010;182:1307–13.
[33] Whyte S, Chilcott J, Halloran S. Reappraisal of the options for colorectal cancer screening in England. Colorectal Dis 2012;14:e547–61.
[34] Di Bidino R, Cicchetti A, Corio M, et al. Impact of technology overlapping: a case study on colorectal cancer screening. Technol Health Care 2010;18:303–15.
[35] Hassan CE, Zullo A, Laghi A, et al. Colon cancer prevention in Italy: cost-effectiveness analysis with CT colonography and endoscopy. Dig Liver Dis 2007;39:242–50.
[36] Heresbach D, Chauvin P, Hess-Migliorretti A, et al. Cost-effectiveness of colorectal cancer screening with computed tomography colonography according to a polyp size threshold for polypectomy. Eur J Gastroenterol Hepatol 2010;22:716–23.
[37] Pence BC, Belasco EJ, Lyford CP. Combination aspirin and/or calcium chemoprevention with colonoscopy in colorectal cancer prevention: cost-effectiveness analyses. Cancer Epidemiol Biomarkers Prev 2013;22:399–405.
[38] Wang ZH, Gao Q-Y, Fang J-Y. Repeat colonoscopy every 10 years or single colonoscopy for colorectal neoplasm screening in average-risk Chinese: a cost-effectiveness analysis. Asian Pac J Cancer Prev 2012;13:1761–6.
[39] Macafee DA, Waller M, Whynes DK, et al. Population screening for colorectal cancer: the implications of an ageing population. Br J Cancer 2008;99:1991–2000.
[40] Berchi C, Bouvier V, Réaud J-M, Launoy G. Cost-effectiveness analysis of two strategies for mass screening for colorectal cancer in France. Health Econ 2004;13:227–38.

[41] Heitman SJ, Hilsden RJ, Au F, et al. Colorectal cancer screening for average-risk North Americans: an economic evaluation. PLoS Med 2010;7:e1000370.
[42] Neilson AR, Whynes DK. Cost-effectiveness of screening for colorectal cancer: a simulation model. IMA J Math Appl Med Biol 1995;12:355–67.
[43] Shimbo T, Glick HA, Eisenberg JM. Cost-effectiveness analysis of strategies for colorectal cancer screening in Japan. Int J Technol Assess Health Care 1994;10:359–75.
[44] Vijan S, Hwang I, Inadomi J, et al. The cost-effectiveness of CT colonography in screening for colorectal neoplasia. Am J Gastroenterol 2007;102:380–90.
[45] Park SM, Yun YH, Kwon S. Feasible economic strategies to improve screening compliance for colorectal cancer in Korea. World J Gastroenterol 2005;11:1587–93.
[46] Bishop J, Glass P, Tracey E, et al. Health Economics Review of Bowel Cancer Screening in Australia. Alexandria, New South Wales: Cancer Institute, 2008.
[47] Dan YY, Chuah BY, Koh DC, Yeoh KG. Screening based on risk for colorectal cancer is the most cost-effective approach. Clin Gastroenterol Hepatol 2012;10:266–71: e1.
[48] Gyrd-Hansen D, Søgaard J, Kronborg O. Colorectal cancer screening: efficiency and effectiveness. Health Econ 1998;7:9–20.
[49] Huang W, Liu G, Zhang X, et al. Cost-effectiveness of colorectal cancer screening protocols in urban Chinese populations. PLoS One 2014;9:e109150.
[50] Lejeune C, Arveux P, Dancourt V, et al. A simulation model for evaluating the medical and economic outcomes of screening strategies for colorectal cancer. Eur J Cancer Prev 2003;12:77–84.
[51] Sonnenberg A, Delcò F, Inadomi JM. Cost-effectiveness of colonoscopy in screening for colorectal cancer. Ann Intern Med 2000;133:573–84.
[52] van Rossum LG, van Rijn AF, Verbeek AL, et al. Colorectal cancer screening comparing no screening, immunochemical and guaiac fecal occult blood tests: a cost-effectiveness analysis. Int J Cancer 2011;128:1908–17.
[53] Tsoi KK, Ng SS, Leung MC, Sung JJ. Cost-effectiveness analysis on screening for colorectal neoplasm and management of colorectal cancer in Asia. Aliment Pharmacol Ther 2008;28:353–63.
[54] Kearns B, Whyte S, Chilcott J, Patnick J. Guaiac faecal occult blood test performance at initial and repeat screens in the English Bowel Cancer Screening Programme. Br J Cancer 2014;111:1734–41.
[55] 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;7:249–61.
[56] Rickert RR, Auerbach O, Garfinkel L, et al. Adenomatous lesions of the large bowel: an autopsy survey. Cancer 1979;43:1847–57.
[57] Vatn MH, Stalsberg H. The prevalence of polyps of the large intestine in Oslo: an autopsy study. Cancer 1982;49:819–25.
[58] Eide TJ, Stalsberg H. Polyps of the large intestine in Northern Norway. Cancer 1978;42:2839–48.
[59] Blatt LJ. Polyps of the colon and rectum. Dis Colon Rectum 1961;4:277–82.
[60] Williams AR, Balasooriya BA, Day DW. Polyps and cancer of the large bowel: a necropsy study in Liverpool. Gut 1982;23:835–42.
[61] Koretz RL. Malignant polyps: Are they sheep in wolves' clothing? Ann Intern Med 1993;118:63–8.
[62] 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;36:179–86.
[63] Stemmermann GN, Heilbrun LK, Nomura A, et al. Adenomatous polyps and atherosclerosis: an autopsy study of Japanese men in Hawaii. Int J Cancer 1986;38:789–94.
[64] Eide TJ. The age-, sex-, and site-specific occurrence of adenomas and carcinomas of the large intestine within a defined population. Scand J Gastroenterol 1986;21:1083–8.
[65] Collett JA, Olynyk JK, Platell CF. Flexible sigmoidoscopy screening for colorectal cancer in average-risk people: update of a community-based project. Med J Aust 2000;173:463–6.
[66] Segnan N, Senore C, Andreoni B, et al. Randomized trial of different screening strategies for colorectal cancer: patient response and detection rates. J Natl Cancer Inst 2005;97:347–57.
[67] Li S, Nie Z, Li N, et al. Colorectal cancer screening for the natural population of Beijing with sequential fecal occult blood test: a multicenter study. Chin Med J (Engl) 2003;116:200–2.
[68] Johnson DA, Gurney MS, Volpe RJ, et al. A prospective study of the prevalence of colonic neoplasms in asymptomatic patients with an age-related risk. Am J Gastroenterol 1990;85:969–74.
[69] DiSario JA, Foutch PG, Mai HD, et al. Prevalence and malignant potential of colorectal polyps in asymptomatic, average-risk men. Am J Gastroenterol 1991;86:941–5.
[70] Imperiale TF, Wagner DR, Lin CY, et al. Risk of advanced proximal neoplasms in asymptomatic adults according to the distal colorectal findings. N Engl J Med 2000;343:169–74.
[71] Lieberman DA, Weiss DG, Bond JH, et al. Use of colonoscopy to screen asymptomatic adults for colorectal cancer. N Engl J Med 2000;343:162–8.
[72] Yee J, Akerkar GA, Hung RK, et al. Colorectal neoplasia: performance characteristics of CT colonography for detection in 300 patients. Radiology 2001;219:685–92.
[73] Regula J, Rupinski M, Kraszewska E, et al. Colonoscopy in colorectal-cancer screening for detection of advanced neoplasia. N Engl J Med 2006;355:1863–72.
[74] Strul H, Kariv R, Leshno M, et al. The prevalence rate and anatomic location of colorectal adenoma and cancer detected by colonoscopy in average-risk individuals aged 40–80 years. Am J Gastroenterol 2006;101:255–62.
[75] Lin OS, Kozarek RA, Schembre DB, et al. Screening colonoscopy in very elderly patients: prevalence of neoplasia and estimated impact on life expectancy. JAMA 2006;295:2357–65.
[76] Imperiale TF, Wagner DR, Lin CY, et al. Results of screening colonoscopy among persons 40 to 49 years of age. N Engl J Med 2002;346:1781–5.
[77] Schoenfeld P, Cash B, Flood A, et al. Colonoscopic screening of average-risk women for colorectal neoplasia. N Engl J Med 2005;352:2061–8.
[78] Lieberman DA, Smith FW. Screening for colon malignancy with colonoscopy. Am J Gastroenterol 1991;86:946–51.
[79] Rex DK, Lehman GA, Hawes RH, et al. Screening colonoscopy in asymptomatic average-risk persons with negative fecal occult blood tests. Gastroenterology 1991;100:64–7.
[80] Liu H-H, Wu M-C, Peng Y, Wu M-S. Prevalence of advanced colonic polyps in asymptomatic Chinese. World J Gastroenterol 2005;11:4731–4.
[81] Segnan N, Senore C, Andreoni B, et al. Comparing attendance and detection rate of colonoscopy with sigmoidoscopy and FIT for colorectal cancer screening. Gastroenterology 2007;132:2304–12.
[82] Hüppe D, Hartmann H, Felten G, et al. Effectiveness of screening colonoscopy in a community-based study. Z Gastroenterol 2008;46:193–200.
[83] Heitman SJ, Ronksley PE, Hilsden RJ, et al. Prevalence of adenomas and colorectal cancer in average risk individuals: a systematic review and meta-analysis. Clin Gastroenterol Hepatol 2009;7:1272–8.
[84] Yamaji Y, Mitsushima T, Ikuma H, et al. Incidence and recurrence rates of colorectal adenomas estimated by annually repeated colonoscopies on asymptomatic Japanese. Gut 2004;53:568–72.
[85] Stryker SJ, Wolff BG, Culp CE, et al. Natural history of untreated colonic polyps. Gastroenterology 1987;93:1009–13.
[86] Hoff G, Foerster A, Vatn MH, et al. Epidemiology of polyps in the rectum and colon: recovery and evaluation of unresected polyps 2 years after detection. Scand J Gastroenterol 1986;21:853–62.
[87] Hofstad B, Vatn MH, Andersen SN, et al. Growth of colorectal polyps: redetection and evaluation of unresected polyps for a period of three years. Gut 1996;39:449–56.
[88] Knoernschild HE. Growth rate and malignant potential of colonic polyps: early results. Surg Forum 1963;14:137–8.
[89] Chen C-D, Yen M-F, Wang W-M, et al. A case-cohort study for the disease natural history of adenoma-carcinoma and de novo carcinoma and surveillance of colon and rectum after polypectomy: implication for efficacy of colonoscopy. Br J Cancer 2003;88:1866–73.
[90] Brenner H, Hoffmeister M, Stegmaier C, et al. Risk of progression of advanced adenomas to colorectal cancer by age and sex: estimates based on 840,149 screening colonoscopies. Gut 2007;56:1585–9.
[91] Chen TH, Yen MF, Lai MS, et al. Evaluation of a selective screening for colorectal carcinoma: the Taiwan multicenter cancer screening (TAMCAS) project. Cancer 1999;86:1116–28.
[92] Hardcastle JD, Chamberlain JO, Robinson MH, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet 1996;348:1472–7.
[93] Kronborg O, Fenger C, Olsen J, et al. Randomised study of screening for colorectal cancer with faecal-occult-blood test. Lancet 1996;348:1467–71.
[94] Mandel JS, Bond JH, Church TR, et al. Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Control Study. N Engl J Med 1993;328:1365–71.
[95] Imperiale TF, Ransohoff DF, Itzkowitz SH, et al. Fecal DNA versus fecal occult blood for colorectal-cancer screening in an average-risk population. N Engl J Med 2004;351:2704–14.
[96] Winawer SJ, Zauber AG, Ho MN, et al. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. N Engl J Med 1993;329:1977–81.
[97] 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.
[98] Faivre J, Dancourt V, Denis B, et al. Comparison between a guaiac and three immunochemical faecal occult blood tests in screening for colorectal cancer. Eur J Cancer 2012;48:2969–76.
[99] Parente F, Boemo C, Ardizzoia A, et al. Outcomes and cost evaluation of the first two rounds of a colorectal cancer screening program based on immunochemical fecal occult blood test in northern Italy. Endoscopy 2013;45:27–34.

[100] Benson VS, Atkin WS, Green J, et al. Toward standardizing and reporting colorectal cancer screening indicators on an international level: the International Colorectal Cancer Screening Network. Int J Cancer 2012;130:2961–73.
[101] Binefa G, Garcia M, Milà N, et al. Colorectal cancer screening programme in Spain: results of key performance indicators after five rounds (2000–2012). Sci Rep 2016;6:19532.
[102] Ladabaum U, Chen FW, Mannalithara A. Mo1711 sensitivity of repeated fecal immunochemical testing (FIT) over time: Does each "bite at the apple" stand the same chance? Gastroenterology 2016;150:S759.
[103] Brenner H, Altenhofen L, Stock C, Hoffmeister M. Incidence of colorectal adenomas: birth cohort analysis among 4.3 million participants of screening colonoscopy. Cancer Epidemiol Biomarkers Prev 2014;23:1920–8.
[104] Brenner H, Altenhofen L, Stock C, Hoffmeister M. Natural history of colorectal adenomas: birth cohort analysis among 3.6 million participants of screening colonoscopy. Cancer Epidemiol Biomarkers Prev 2013;22:1043–51.
[105] Tanaka T. Colorectal carcinogenesis: review of human and experimental animal studies. J Carcinog 2009;8:5.
[106] Greuter MJ, Xu XM, Lew JB, et al. Modeling the Adenoma and Serrated pathway to Colorectal CAncer (ASCCA). Risk Anal 2014;34:889–910.
[107] Minoo P, Zlobec I, Peterson M, et al. Characterization of rectal, proximal and distal colon cancers based on clinicopathological, molecular and protein profiles. Int J Oncol 2010;37:707–18.
[108] Iacopetta B. Are there two sides to colorectal cancer? Int J Cancer 2002;101:403–8.
[109] Briggs AH, Weinstein MC, Fenwick EA, et al. Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-6. Med Decis Making 2012;32:722–32.
[110] Karnon J, Vanni T. Calibrating models in economic evaluation: a comparison of alternative measures of goodness of fit, parameter search strategies and convergence criteria. Pharmacoeconomics 2011;29:51–62.
[111] Rutter CM, Miglioretti DL, Savarino JE. Bayesian calibration of microsimulation models. J Am Stat Assoc 2010;104:1338–50.
[112] Taylor DC, Pawar V, Kruzikas D, et al. Methods of model calibration: observations from a mathematical model of cervical cancer. Pharmacoeconomics 2010;28:995–1000.
[113] Stout NK, Knudsen AB, Kong CY, et al. Calibration methods used in cancer simulation models and suggested reporting guidelines. Pharmacoeconomics 2009;27:533–45.
[114] Frederix GW, van Hasselt JG, Schellens JH, et al. The impact of structural uncertainty on cost-effectiveness models for adjuvant endocrine breast cancer treatments: the need for disease-specific model standardization and improved guidance. Pharmacoeconomics 2014;32:47–61.