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.)

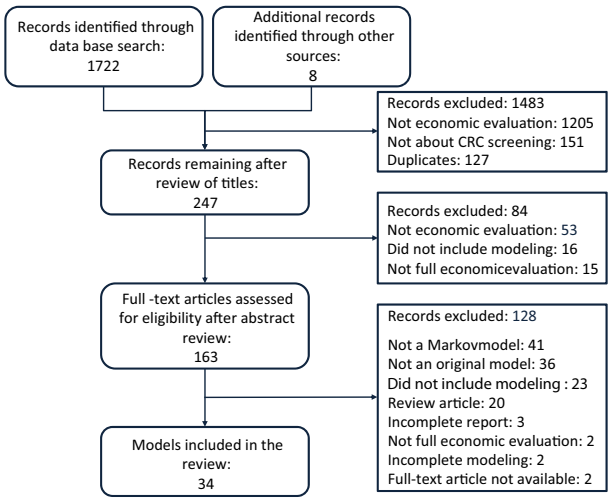

Fig. 1 – PRISMA diagram. CRC, colorectal cancer; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.