Mathematical models and cost-effective screening strategies for colorectal cancer

Robert A. Smith PhD

Previously published at www.cmaj.ca

∞∞∞ See related research article by Telford and colleagues, page 1307, and see also research article by Simunovic and colleagues, page 1301

In 1993 the results of the Minnesota Colon Cancer Control Study provided clear evidence of the efficacy of screening for colorectal cancer with a guaiac-based fecal occult blood test.[1] By 1997 two leading groups in the United States cited both the strong evidence from the Minnesota trial and inferential evidence from nonexperimental studies to endorse not only annual screening with the fecal occult blood test, but also additional screening options, including colonoscopy every 10 years, double contrast barium enema every 5 years, flexible sigmoidoscopy every 5 years, and the combination of a fecal occult blood test annually and a flexible sigmoidoscopy every 5 years.[2,3] None of these tests was given priority over the others based on the argument that (a) they all were effective, (b) not all tests were equally preferred or accessible and (c) there was uncertainty over the relative balance of costs, benefits and harms in an ongoing colorectal cancer screening program.

Key points

- Given the number of options for colorectal cancer screening, the use of modelling to estimate outcomes and costs of programs is particularly useful for planning screening strategies.
- The underlying assumptions, methods and results of mathematical modelling require careful scrutiny by policy-makers, as well as sensitivity analysis of key model inputs.
- Policy-makers need to appreciate that predicted outcomes of programs depend on high-quality test performance and consistently high rates of compliance with screening and follow-up recommendations.

Although the recommendations were straightforward enough for individual decisions about screening, program planners have relied on the results of mathematical modelling, such as those reported by Telford and colleagues,[4] to estimate relevant costs and outcomes of screening programs. The use of modelling has been particularly relevant to screening strategies for colorectal cancer given the growing number of screening options, the potential for early detection and prevention, the consistent observation that all strategies meet conventional criteria for cost-effectiveness, and the low likelihood of additional prospective randomized controlled trials (RCTs) to compare alternative screening strategies with existing technology. Yet, the underlying assumptions, methods and results of mathematical modelling require careful scrutiny by policy-makers and a realistic appraisal as to whether the performance benchmarks are achievable and at what costs. Ultimately, modelling will be most productive as a guide to health policy when there is as much emphasis on the program's performance benchmarks as there is on the expected outcomes.

The earliest models were relatively crude. More recent simulations have evolved in sophistication to the point where they are now playing a role in defining screening recommendations.[5] This important new development in evidence-based medicine is not without controversy. Although evidence from prospective RCTs is preferred, the appetite of funding agencies for additional RCTs is diminished once the efficacy of screening has been shown. However, given the investment in costs and time, RCTs may not always be necessary or even advisable once efficacy has been established. In general, if high-quality data on test performance are available, the use of mathematical modelling is an efficient approach to examine the costs and outcomes of improved performance, alternative screening strategies, or tailored approaches to screening based on risk.

Ironically, even if a screening test receives a favourable evaluation in an RCT, policy-makers may still be hesitant to endorse it because of unanswered questions. In the United States, this was the case with both fecal DNA testing and colonography using computed tomography.[6,7] Each method showed acceptable test performance, but concerns about costs and potential harms were deciding factors in withholding approval for reimbursement. The RCT data and additional data from observational studies pertaining to lingering uncertainties about these technologies will likely be modelled to ultimately determine whether these tests are cost-effective options for screening all adults, or certain subpopulations.

Still, we should not be so easily seduced by the appearance of precision and rigour of mathematical modelling. Like any research exercise, the underlying assumptions, methods and results of modelling require careful scrutiny. The more complex the model, the longer the list of assumptions about such

Robert Smith is the Director of Cancer Screening at the American Cancer Society, Atlanta, USA.

CMAJ 2010. DOI:10.1503/cmaj.100986

DOI:10.1503/cmaj.100986

elements as test performance, expected outcomes, resources and costs. The description of a model's methodology generally is so lengthy that only an abbreviated summary is included in an article, with the more detailed description usually published in an online appendix. Further, it is common that the methodologic rigour of the model is not as evident in the manner in which model assumptions and inputs are chosen, which too often are not derived from systematic reviews and thus choices may be debatable.[8]

One approach to overcome uncertainty over model inputs is a sensitivity analysis, which has been described as a core methodologic obligation in cost-effectiveness analyses.[9] The sensitivity analysis will be most informative if it includes both best- and worst-case scenarios. Should the efficacy of a screening test, or its performance benchmarks, be derived from a single study in an expert setting, a meta-analysis of studies or evidence from community practice? Here the modeller faces the choice of building the model on the basis of best-case assumptions versus what is achieved in average clinical settings. Although a sensitivity analysis can ensure that the model is informative to a range of perspectives, the number of potentially relevant sensitivity analyses can be daunting. Even conscientious investigators will probably not satisfy everyone. The statistician George Box commented that "… all models are wrong, but some are useful."[10] In essence, Box reminds us that all models, as attempts to simulate reality, are inherently more or less defective, given the number of parameters in a model, the varying precision of measurements, unknown covariates, interactions that are not well understood, and the fact that measurement error and uncertainty may be magnified when decades of interventions and outcomes are being simulated.

Telford and colleagues have made an important contribution to the literature on screening for colorectal cancer by addressing several key elements in their model, including noncompliance with screening intervals, the costs of modern chemotherapy, the estimation of quality-adjusted life-years and the performance of sensitivity analyses on key model inputs.[4] Of particular relevance are the sensitivity analyses in which they varied the accuracy of test performance in quasi best- and worst-case scenarios, and their examination of the contribution of screening to disease control when the cost of state-of-the-art treatment increases exponentially.

It can be argued that their model has overestimated the accuracy of both fecal occult blood testing and colonoscopy, at least in terms of current performance outside of expert settings. Thus, their predicted reductions in the incidence and mortality of colorectal cancer may be higher than would realistically be achieved if a program based on fecal occult blood testing or colonoscopy were launched today.

Yet, these performance characteristics have been achieved in clinical settings. Thus, a clear message to Canadian policy-makers is needed that emphasizes what must be in place to achieve the predicted outcomes. Screening involves multiple steps, each of which is vulnerable to failures that can collectively diminish the potential to reduce the incidence and mortality of colorectal cancer. In particular, test performance and rates of compliance with screening and follow-up recommendations are critical indicators for ongoing evaluation and corrective action. If there is consensus that the model inputs are achievable, these indicators must be monitored on an ongoing basis to achieve the predicted outcomes of a program. As Telford and colleagues have shown, investing in one technology over another may ultimately be determined to have been the wrong choice if the assumptions of the model cannot be met.

This article was solicited and has not been peer reviewed.

Competing interests: None declared.

REFERENCES

1. 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.
2. Byers T, Levin B, Rothenberger D, et al. American Cancer Society guidelines for screening and surveillance for early detection of colorectal polyps and cancer: update 1997. American Cancer Society Detection and Treatment Advisory Group on Colorectal Cancer. *CA Cancer J Clin* 1997;47:154-60.
3. Winawer SJ, Fletcher RH, Miller L, et al. Colorectal cancer screening: clinical guidelines and rationale. *Gastroenterology* 1997;112:594-642.
4. Telford JJ, Levy AR, Sambrook JC, et al. The cost-effectiveness of colorectal cancer screening in Canada. *CMAJ* 2010;182:1307-13.
5. Zauber AG, Lansdorp-Vogelaar I, Knudsen AB, et al. Evaluating test strategies for colorectal cancer screening: a decision analysis for the US Preventive Services Task Force. *Ann Intern Med* 2008;149:659-69.
6. 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.
7. Johnson CD, Chen MH, Toledano AY, et al. Accuracy of CT colonography for detection of large adenomas and cancers. *N Engl J Med* 2008;359:1207-17.
8. Campbell H, Briggs A, Buxton M, et al. The credibility of health economic models for health policy decision-making: the case of population screening for abdominal aortic aneurysm. *J Health Serv Res Policy* 2007;12:11-7.
9. Weinstein MC, Siegel JE, Gold MR, et al. Recommendations of the Panel on Cost-effectiveness in Health and Medicine. *JAMA* 1996;276:1253-8.
10. Box GEP, Draper NR. *Empirical model-building and response surfaces*. New York (NY): Wiley; 1987.

Correspondence to: *Dr. Robert A. Smith, Director of Cancer Screening, American Cancer Society, Ste. 600, 250 Williams St., Atlanta GA 30303, USA; robert.smith@cancer.org*