Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Screening For Colorectal Cancer in the Age of Simulation Models: A Historical Lens

In late March, *The Washington Post*, citing unnamed officials, quoted Dr Anthony S. Fauci's response to the demands of modeling the novel coronavirus epidemic: "I've looked at all the models. I've spent a lot of time on the models. They don't tell you anything. You can't really rely upon models."[1] The provocative claim from the Director of the National Institute of Allergy and Infectious Diseases was quickly seized upon by critics of the Coronavirus Task Force's recommendations. Simulation models, however, are certainly not new and have played a prominent role in epidemiology for many decades.[2] Less visibly, they have also come to dominate other areas of medicine, especially preventive medicine. When the American Cancer Society (ACS) issued a modification of the United States Preventive Services Task Force (USPSTF) recommendations on colorectal cancer (CRC) screening guidelines in 2018, the society justified its actions by rerunning a simulation model. The new recommendation by the ACS encountered opposition and was not universally embraced.[3–6] Tellingly, no new empirical studies of screening were involved in the revision: the updated recommendations emanated from a recalibration of the incidence of CRC within a simulation model.

In this commentary, we report on the historical evolution of simulation modeling as applied to screening for CRC. We highlight how simulation modeling has come to be regarded by some as equivalent to evidence-based data, but we caution against overreach by relying too heavily on modeling relative to empirical testing.

Computer simulation models sit uncomfortably alongside evidence provided by controlled observational studies and formal randomized trials. Simulations are at once experimental—researchers can adjust parameters and observe different outcomes—and yet also fundamentally nonempirical, in that the parameters, inputs, and assumptions need not have strong empirical bases of support. According to Douglas Owens of Stanford University and the USPSTF, the introduction of simulation modeling into CRC screening recommendations helped address questions that clinical trials had left unanswered.[7] But as Eric Winsberg's survey of the rise of computer simulation has noted, "It is a mistake to think of simulations simply as tools for unlocking hidden empirical content."[8] Simulations do not produce empirical evidence; they enable the creation of "present futures" or alternative visions of the future to help decision making in the present.[9]

As recently as 1996, however, simulation models were not used at all in screening recommendations. The rapid rise in their importance over the following 2 decades, and their central role in shaping current screening recommendations, should occasion reflection about how and why simulation models have become an influential tool for medical decision making.

The Rise of Screening Models

Although the introduction of electronic computers and algorithms into medicine broadly dates back to the 1950s, the idea of simulating the effects of screening with computers arose only in the early 1970s.[10] In 1973, E.G. Knox, a professor of social medicine at the University of Birmingham in England proposed introducing computer simulation to provide a more rigorous economic justification for preventive measures such as screening. Knox's first simulation program was for cervical cancer screening and constituted 650 statements of Fortran code. It used the insight that each individual in a population could be modeled as moving between a series of states, from one of "normal," to others of "occult invasive disease," "clinical invasive disease," "dead of other causes," "treated early-invasive," "dead of the clinical cancer," and so on. The simulation model assigned probabilities of transferring between these states. Knox began with a cohort of 10,000 people with a known life-table taken from the Registrar-General's reports for England and Wales in 1970 and used a series of punched cards fed into an electronic computer to model transitions between 26 disease states.[11]

The basic elements of Knox's simulation have remained in many subsequent models: vital statistics data are combined with empirical evidence about transitions between disease states to predict the number of deaths or life-years lost. This simulation can then be rerun after introducing a preventive screening intervention. Empirical evidence about screening, such as the effect of screening on incidence and stage of tumors, can be incorporated and the change in deaths or life-years lost computed, and ultimately weighed against the financial and medical costs of the screening intervention. There are of course different varieties and subtleties among models—state transitions can be treated as continuous or discrete, for example—but we still use Knox's essential idea that a simulation should be run twice, once with and once without an intervention to estimate the intervention's effects on progression of disease.

Knox's work obviously required the existence of widely available electronic computers, but it was more directly a product of reforms in social medicine meant to address the increasing costs of chronic disease. Knox's original model was published as part of a symposium on the future of preventive medicine within the National Health Service, and the importance of modeling studies would only increase with the creation of the Canadian Task Force on the Periodic Health Examination in the late 1970s and the USPSTF in 1984.[12,13] During the 1980s, computer simulations of screening regimens proliferated as researchers sought to understand their potential benefits and costs.[14–16]

Until the twenty-first century, however, computer simulations were not seen as providing evidence definitive enough to shape screening recommendations. Randomized controlled trials continued to provide the highest level of empirical evidence for the USPSTF and the Canadian Task Force on Preventive Health Care.[17] The 1996 *Guide to Clinical Preventive Services* of the USPSTF did not rely on even established simulation models and 2 years later USPSTF representatives advocating for the use of models still acknowledged "important obstacles" because models were "invariably complex and involve numerous assumptions and subjective judgments."[18,19]

A major turning point was the development of the Cancer Intervention and Surveillance Network (CISNET) in 2000. CISNET is a consortium of modelers supported by the National Cancer Institute who, while independently modeling the effects of different cancer control interventions for breast, prostate, colorectal, esophagus, and lung cancers, collaborate whenever possible by addressing a similar question with common inputs and outputs so models can be compared. Thus, models and modelers might act in effect as a panel of experts, each weighing and assessing the empirical evidence, fleshing out disagreements in the service of producing a consensus recommendation.

In the case of CRC, 3 models are part of the CISNET group: MIcrosimulation SCreening Analysis Colorectal Cancer Model (MISCAN-Colon), Simulation Model of Colorectal Cancer (SimCRC), and Colorectal Cancer Simulated Population model for Incidence and Natural history (CRC-SPIN). Models for CRC screening were often based on those designed for breast and cervical cancers, but emerged later because empirical studies on the effects of CRC screening arrived later. The first CRC models were created in the 1990s,[20,21] incorporating the first randomized controlled studies of CRC screening.[22–25] The resulting simulation models within the CISNET group are similar in many regards: all are stochastic microsimulation models (meaning that instead of cohorts they model an individual's probabilistic transition between disease states) and all take into account the natural history of CRC as well as Surveillance, Epidemiology, and End Results (SEER) incidence and mortality data.

Simulation Models in Contemporary Preventive Medicine

In preparation for its 2008 recommendations on CRC screening, the USPSTF augmented their systematic evidence review for the first time with models provided by CISNET researchers, prompting praise for the use of state-of-the-art technology, but also concern that modelers had made some "surprising choices," including not quality adjusting the estimates of benefits or accounting for patient and institutional costs of adherence.[26] By 2016, the USPSTF's CRC recommendations relied heavily on CISNET researchers to provide information on the starting and stopping ages, as well as recommended intervals for screening methods. In particular, CISNET researchers modeled the experience of 40-year-old individuals with no previous cancer diagnosis, and aggregated those experiences over an entire hypothetical cohort to measure both the benefit (in life-years gained [LYG]) and the cost (in number of lifetime colonoscopies required) of screening. Effectiveness was defined by an "efficient frontier," highlighting the strategies that provided the "largest incremental increase in LYG per additional colonoscopy performed."[27] They concluded "the strategies of colonoscopy every 10 years, annual FIT (fecal immunochemical testing), SIG (flexible sigmoidoscopy) every 10 years with annual FIT, and CTC (computerized tomographic colonography) every 5 years performed from ages 50 to 75 years provided similar LYG and a comparable balance of benefit and screening burden."[28]

All 3 models predicted that starting screening at age 45 would result in an increase in LYG, but the USPSTF rejected recommending earlier initiation dates. They reasoned that the increase in LYG relative to the burden of increased colonoscopy examinations was small. Combining an earlier starting age with an extension of the subsequent screening interval from every 10 to every 15 years would have maintained a similar estimate of required lifetime colonoscopies but 1 of the 3 models (MISCAN-Colon) predicted this solution would cause a loss in LYG. In the end, the USPSTF considered these findings and concluded that "the evidence best supports a starting age of 50 years for the general population, noting the modest increase in LYG by starting screening earlier, the discordant findings across models for extending the screening interval when the age at which to begin screening is lowered, and the lack of empirical evidence in younger populations."[28]

The ACS's review 2 years later addressed the question of initiation age by homing in on the discordant findings across models. In particular, they noted the apparent outlier, the MISCAN-Colon simulation model, would have agreed with the other 2 about the benefit of starting screening at age 45 if it had incorporated new information on the increased incidence of CRC in younger people.[29] All 3 CISNET models had used SEER incidence rates taken from 1975 to 1979 (providing an incidence rate before the institution of appreciable screening activity, hence more reflective of the underlying natural history). The ACS research team used MISCAN-Colon to simulate "a cohort of adults aged 40 years in 2015, and assumed that this cohort had a 1.591-fold increased CRC incidence across all ages compared with the original model."[30] The incidence multiplier was derived from comparing SEER data from 1975 with more recent incidence rates. Rerunning the MISCAN-Colon model with updated assumptions resulted in the ACS Guideline Development Group providing a qualified recommendation that screening begin at 45, while maintaining initiation at 50 as a strong recommendation.[31,32]

Discussion

In this instance, the use of a simulation model for CRC screening

recommendations engendered controversy rather than resolved questions. Although there was indeed new epidemiologic evidence on CRC incidence published between 2016 and 2018,[29] there were no new empirical studies about the effects of earlier screening. In fact, as the ACS group acknowledged, there has never been any substantial research on the effects of starting screening before age 50 for CRC.[32] Although the model predicts benefits in earlier detection of CRC, a number of unintended consequences could potentially follow,[6] primary of which is the diversion of resources away from higher risk populations for whom the relative benefit of screening is substantially greater.[33]

Moreover, the agreement or disagreement of models is not a straightforward measure of the strength or weakness of evidence. It is difficult to know whether differences in predictions are due to differences in model design or model inputs.[34] Two models that have been accurately validated using existing studies may have quite different predictions about future outcomes. A 2016 study by CISNET-CRC modelers, for example, noted that the essential but unobservable process of transition from benign adenoma to malignant tumor, or dwell time, is simulated by all the models but such a value is an output rather than input or assumption. As a result, even if models are carefully calibrated to match existing data, information about how sensitive predictions are to different dwell-time distributions, for example, remains unknown.[35]

Given the expense of trials and the difficulty of combining multiple parameters simultaneously, such as age of initiation, screening interval, and preferred modality—modeling may indeed be an essential component for decision making. The ease of modeling compared with large-scale clinical trials, as well as the ability of models to combine evidence from many different sources also favors their use. Models have significant limitations, however, from their imposition of assumptions about the natural history of disease, the screening process, and the behavior of individuals, to the possibility of producing results with misleading levels of precision.

Most important, no matter how elaborate the design of the virtual laboratory created by a model[36] or the quality of the experiments within, empirical studies remain essential. As has become abundantly clear in the case of statins[37] and hormone replacement therapy,[38] in the absence of large and unexpected effects, only randomized controlled trials of adequate size performed in applicable settings can provide reliable information about an intervention's effectiveness. As Dr Fauci said recently, "All models are just models. When you get new data, you change them."[39] New epidemiologic data on incidence rates can certainly justify updating a model's assumptions, which then might lead the model to make different predictions about the effects of screening, but without formal randomized trials of the new intervention, we have no new empirical information about effectiveness.

Computer simulations have subtly expanded what counts as evidence-based medicine. Balancing reliance on models with careful assessment of the systematic and empirical evidence is the core responsibility of organizations developing and delivering prescriptions for public policy. Careful deliberation is required to avoid over reliance on modeling. Models are no substitute for real-world data,[40] and their conclusions remain bound by the availability of rigorous empirical studies.

CHRISTOPHER J. PHILLIPS
Department of History
Carnegie Mellon University
Pittsburgh, Pennsylvania

ROBERT E. SCHOEN
Division of Gastroenterology
Hepatology and Nutrition
Departments of Medicine
and Epidemiology
University of Pittsburgh
Pittsburgh, Pennsylvania

References

1. Wan W, Dawsey J, Parker A, et al. Experts and Trump's advisers doubt White House's 240,000 coronavirus deaths estimate. The Washington Post, April 2, 2020. Available from: www.washingtonpost.com/health/2020/04/02/experts-trumps-advisers-doubt-white-houses-240000-coronavirus-deaths-estimate/.
2. Awerbuch T. Evolution of mathematical models of epidemics. Ann N Y Acad Sci 2004;740:232–241.
3. Bretthauer M, Kalager M, Weinberg DS. From colorectal cancer screening guidelines to headlines: beware! Ann Intern Med 2018;169:405–406.
4. Corley DA, Peek RM Jr. When should guidelines change? A clarion call for evidence regarding the benefits and risks of screening for colorectal cancer at earlier ages. Gastroenterology 2018;155:947–949.
5. Imperiale TF, Kahi CJ, Rex DK. Lowering the starting age for colorectal cancer screening to 45 years: who will come...and should they? Clin Gastroenterol Hepatol 2018;16:1541–1544.
6. Liang PS, Allison J, Ladabaum U, et al. Potential intended and unintended consequences of recommending initiation of colorectal cancer screening at age 45 years. Gastroenterology 2018;155:950–954.
7. Bauchner H. Conversations with Dr. Bauchner: screening for colorectal cancer. JAMA Network 2016. Available from: https://sites.jamanetwork.com/colon-cancer-screening/jama-author-interview-screening-for-colorectal-cancer.html.
8. Winsberg E. Science in the age of computer simulation. Chicago: University of Chicago Press, 2010.
9. Beckert J. Imagined futures: fictional expectations and capitalist dynamics. Cambridge, MA: Harvard University Press, 2016.
10. November J. Biomedical computing: digitizing life in the united states. Baltimore, MF: Johns Hopkins University Press, 2012.
11. Knox EG. A simulation system for screening procedures. In: McLachlan G, ed. The future—and present indicatives. London: Oxford University Press, 1973:18–55.

12. Johns DM, Bayer R. The paradox of authority — transformation of the USPSTF under the Affordable Care Act. N Engl J Med 2016; 375:1710–1712.
13. McGinnis JM, Woolf SH. Background and Objectives of the U.S. Preventive Services Task Force. J Gen Intern Med 1990;(5 Suppl):S11–S13.
14. Habbema JDF, van Oortmarssen GJ, Lubbe JThN, et al. The MISCAN simulation program for the evaluation of screening for disease. Comput Methods Programs Biomed 1984; 20:79–93.
15. Parkin DM. A computer simulation model for the practical planning of cervical cancer screening programmes. Br J Cancer 1985; 51:551–568.
16. Shun-Zhang Y, Miller AB, Sherman GJ. Optimising the age, number of tests, and test interval for cervical screening in Canada. J Epidemiol Community Health 1982;36:1–10.
17. Stevenson CE. Statistical models for cancer screening. Stat Methods Med Res 1995;4:18–32.
18. Atkins D, DiGuiseppi CG. Broadening the evidence base for evidence-based guidelines: a research agenda based on the work of the U.S. Preventive Services Task Force. Am J Prev Med 1998;14:335–344.
19. US Preventive Services Task Force. Guide to preventive services. 2nd ed. Washington, DC: U.S. Department of Health and Human Services, Office of Public Health and Science, Office of Disease Prevention and Health Promotion, 1996.
20. Eddy DM. Screening for colorectal cancer. Ann Intern Med 1999; 113:373–384.
21. Loeve F, Boer R, van Oortmarssen GJ, et al. The MISCAN-COLON simulation model for the evaluation of colorectal cancer screening. Computers and Biomedical Research 1999; 32:13–33.
22. Hardcastle JD, Chamberlain JO, Robinson MH, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet 1996;348:1472–1477.
23. Kronborg O, Fenger C, Olsen J, et al. Randomised study of screening for colorectal cancer with faecal-occult-blood test. Lancet 1996;348:1467–1471.
24. Mandel JS, Bond JH, Church TR, et al. Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Control Study. [Erratum appears in N Engl J Med 1993 Aug 26;329:672]. N Engl J Med 1993; 328:1365–1371.
25. Mandel JS, Church TR, Bond JH, et al. The effect of fecal occult-blood screening on the incidence of colorectal cancer. [letter; comment]. N Engl J Med 2000; 343:1603–1607.
26. Pignone M, Sox HC. Screening guidelines for colorectal cancer: a twice-told tale. Ann Intern Med 2008;149:680–682.
27. Knudsen AB, Zauber AG, Rutter CM, et al. Estimation of benefits, burden, and harms of colorectal cancer screening strategies: modeling study for the US Preventive Services Task Force. JAMA 2016;315:2595–2609.
28. US Preventive Services Task Force. Screening for colorectal cancer: US Preventive Services Task Force Recommendation Statement. JAMA 2016; 315:2564–2575.
29. Siegel RL, Fedewa SA, Anderson WF, et al. Colorectal cancer incidence patterns in the United States, 1974–2013. J Natl Cancer Inst 2017:109.
30. Peterse EFP, Meester RGS, Siegel RL, et al. The impact of the rising colorectal cancer incidence in young adults on the optimal age to start screening: microsimulation analysis I to inform the American Cancer Society Colorectal Cancer Screening Guideline. Cancer 2018; 124:2964–2973.
31. Meester RGS, Peterse EFP, Knudsen AB, et al. Optimizing colorectal cancer screening by race and sex: microsimulation analysis II to inform the American Cancer Society Colorectal Cancer Screening Guideline. Cancer 2018;124:2974–2985.
32. Wolf AMD, Fontham ETH, Church TR, et al. Colorectal cancer screening for average-risk adults: 2018 guideline update from the American Cancer Society. CA Cancer J Clin 2018;68:250–281.
33. Ladabaum U, Mannalithara A, Meester RGS, et al. Cost-effectiveness and national effects of initiating colorectal cancer screening for average-risk persons at age 45 years instead of 50 years. Gastroenterology 2019; 157:137–148.
34. Boer R, Plevritis S, Clarke L. Diversity of model approaches for breast cancer screening: a review of model assumptions by The Cancer Intervention and Surveillance Network (CISNET) Breast Cancer Groups. Stat Methods Med Res 2004;13:525–538.
35. Rutter CM, Knudsen AB, Marsh TL, et al. Validation of models used to inform colorectal cancer screening guidelines: accuracy and implications. Med Decis Making 2016; 36:604–614.
36. Alagoz O. Introduction to the Cancer Intervention and Surveillance Modeling Network (CISNET) breast cancer models. Medical Decision Making 2018;38(1S):3S–8S.
37. Collins R, Reith C, Emberson J, et al. Interpretation of the evidence for the efficacy and safety of statin therapy. Lancet 2016;388:2532–2561.
38. Glasziou P, Chalmers I, Rawlins M, et al. When are randomised trials unnecessary? Picking signal from noise. BMJ 2007;334:349–351.
39. McNeil DG Jr. The coronavirus in America: the year ahead. The New York Times (April 18, 2020). Available from: www.nytimes.com/2020/04/18/health/coronavirus-america-future.html.
40. Corrigan-Curay J, Sacks L, Woodcock J. Real-world evidence and real-world data for evaluating drug safety and effectiveness. JAMA 2018;320:867–868.

Received April 29, 2020. Accepted July 2, 2020.

Conflicts of interest
The author discloses no conflicts.

Funding
Early Detection Research Network UO1-CA152753 to RES.

Most current article

0016-5085/$36.00
https://doi.org/10.1053/j.gastro.2020.07.010