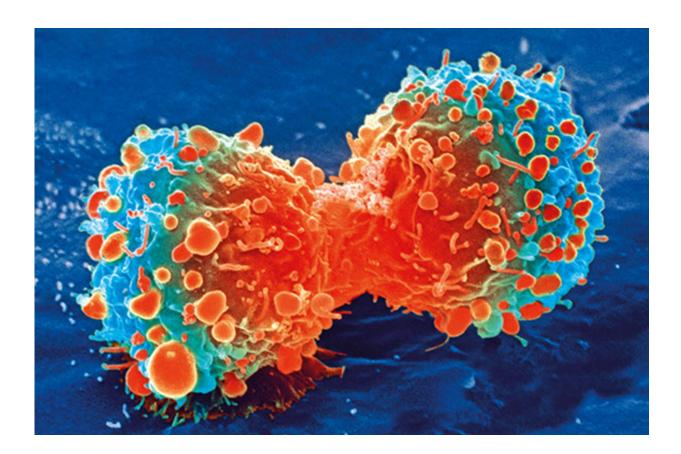# Study investigates assumptions about colonoscopy accuracy

April 5 2022

Cancer cell during cell division. Credit: National Institutes of Health

If you're over 45, chances are your doctor has told you to get screened for colorectal cancer. While it's a seemingly routine recommendation these days, a great deal of scientific analysis goes into the details.

Regularly updated guidelines first released in the mid-1990s outline the age range when people should be checked for disease, how often and which tests should be considered.

Computer modeling helps inform these screening policies and many other aspects of cancer detection and care. A recently published study using high performance computing resources at the U.S. Department of Energy's (DOE) Argonne National Laboratory investigates modeling assumptions for a common colorectal cancer screening method: colonoscopy. The finding could ultimately be used to improve both modeling accuracy and future decisions about screening regimens.

"Even though we found that we may be overstating the sensitivity of colonoscopy, it still seems to be quite effective at finding many precursors of cancer," says Carolyn Rutter of RAND

The stakes for effective screening protocols are high: Colorectal cancer is the second leading cause of cancer deaths, behind lung cancer. Though overall cases and deaths have declined in recent decades—thanks in part to increased testing—colorectal cancer rates among people under 50 are on the rise. Last year, the U.S. Preventive Services Task Force, an independent panel that issues the guidelines, revised the starting age for colorectal screening from age 50 to 45.

An unexpected result

During a colonoscopy, doctors use a camera to look inside the large intestine for both cancers and adenomas, which are lesions that have the potential to become cancer. The new research, which is a collaboration among scientists at the RAND Corporation, Kaiser Permanente and Argonne, concluded that colonoscopies may detect fewer small adenomas than currently believed. Fortunately, these small adenomas very rarely become cancerous.

The researchers initially planned to look at how well the Colorectal Cancer Simulated Population model for Incidence and Natural history (CRC-SPIN) predicted the occurrence of adenomas. Using the model, they set out to validate a trial from the early 1990s that examined whether wheat bran supplements prevented a recurrence of adenomas in people found to have one or more of the tumors in a baseline colonoscopy. (The study's conclusion: Supplements didn't have an effect.)

Replicating real-life experiments with computer models helps improve the models. But the team hit a snag: The CRC-SPIN model, a well-established tool in use for more than a decade, failed to match the wheat bran fiber trial's results. The model found too few adenomas in people who were at relatively high adenoma risk.

"The question then became, why is our model wrong? It's working for all these other things," said Carolyn Rutter, a senior statistician at RAND and lead author of the study. Rutter's colleague, Pedro Nascimento de Lima, an assistant policy researcher at RAND and a visiting graduate student at Argonne, suggested using the high performance computing power at the Argonne Laboratory Computing Resource Center (LCRC) to investigate.

Questioning the assumptions

When researchers use a model like CRC-SPIN, they start with a set of assumptions. To figure out why the model failed to validate the wheat bran fiber study, Rutter and colleagues had to figure out which set of assumptions would produce the right data—sort of like solving a math problem by starting with the answer. The LCRC's Bebop cluster, a powerful system of computing processors, made this endeavor possible. By efficiently running a large number of models at the same time, Bebop allowed for an investigation of a broad set of assumptions.

"Doing the work at Argonne allows us to go from one single set of assumptions to thousands," said Nascimento de Lima. "Without computational power, you cannot even try to do a systematic analysis of what assumptions worked with the model—and with the new data we are receiving, what assumptions don't work."

The model better matched the real-world study when the team lowered the assumed sensitivity values, which describe the chance that an adenoma is identified across different size categories during the colonoscopy. When sensitivity was assumed to be very high, the model simulated the discovery and removal of too many adenomas at the beginning of the study, especially when it came to the smallest ones. That left too few to grow and be detected later in the study, creating a mismatch between the model's results and the wheat bran fiber trial.

The research suggests that modeling studies have been assuming colonoscopies are more accurate than they really are—"too good to be true," as the study's name says. But the results do not negate colonoscopy's value as a cancer detection and prevention tool.

"Even though we found that we may be overstating the sensitivity of colonoscopy, it still seems to be quite effective at finding many precursors of cancer," Rutter said.

The findings assume the CRC-SPIN model is correct, but the team will continue to analyze the model for other possible explanations. For example, it could be that people who tend to develop adenomas also tend to have faster-growing adenomas. In that case, even a highly sensitive colonoscopy would miss them early on, and the model isn't accounting for this.

"Extending the model to examine limitations is much more difficult than examining colonoscopy sensitivity," the researchers write in the paper,

"but ultimately must be part of model evaluation."

Better models, more robust policy

Argonne computational scientist Jonathan Ozik, a co-author on the paper, worked with Rutter and Nascimento de Lima to run the experiments with the CRC-SPIN model on LCRC's Bebop, but his larger goal is to make this sort of research easier.

"We really like to work collaboratively with people, but our ultimate goal is to help people to do this type of work on their own," Ozik said, "so they can decide which experiments they want to run, what high performance computing resources they want to run them on, and then advance the science."

With this goal in mind, Ozik and a team of collaborators from Argonne have developed a framework called Extreme-scale Model Exploration with Swift (EMEWS). In addition to enabling the experiments run for the colonoscopy study, EMEWS has helped researchers support COVID-19 decision making in Chicago and Illinois and simulate precision cancer treatment scenarios. The framework also enabled Nascimento de Lima, Rutter and collaborators to evaluate 78 different reopening strategies for California in the wake of the COVID-19 pandemic.

"We are trying to create an ecosystem of computationally intensive research that can be done now but is not always thought of as the obvious next step by researchers," Ozik said.

Rutter and team will continue to probe the CRC-SPIN model for potential improvements and new insights on, for example, adenoma growth in people under 50. Their findings could eventually inform the U.S. Preventive Services Task Force when it makes future

recommendations on colorectal screening.

"Disease modeling has become important because there is just so much data," Rutter said. "This research is part of a bigger movement toward synthesizing all of that information and using it to present outcomes from different health policy approaches."

The study was published in January in the journal *Cancer Epidemiology, Biomarkers & Prevention*.

  More information: Carolyn M. Rutter et al, Too Good to Be True? Evaluation of Colonoscopy Sensitivity Assumptions Used in Policy Models, *Cancer Epidemiology, Biomarkers & Prevention* (2021). DOI: 10.1158/1055-9965.EPI-21-1001

Provided by Argonne National Laboratory