JNCI J Natl Cancer Inst (2021) 113(2): djaa104

doi: 10.1093/jnci/djaa104
First published online August 6, 2020
Editorial

Evaluating a New Cancer Screening Blood Test: Unintended Consequences and the Need for Clarity in Policy Making

David F. Ransohoff, MD*

Department of Medicine (Gastroenterology and Hepatology), Department of Epidemiology, and Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

***Correspondence to:** David F. Ransohoff, MD, Department of Medicine (Gastroenterology and Hepatology), Department of Epidemiology, and Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Campus Box 7080, Chapel Hill, NC 27599-7080, USA (e-mail: ransohof@med.unc.edu).

A blood test for cancer screening has been the "holy grail" ever since the carcino-embryonic antigen blood test in the 1960s was claimed to have nearly 100% sensitivity and specificity—but turned out not to (1)—for colorectal cancer (CRC). Nonadherence with CRC screening recommendations, now approximately 40%, might be reduced if a blood test were available to persons unwilling to have current tests that may be unpleasant (colonoscopy and its preparation) or distasteful (stool collection).

The process of screening test approval and adoption involves policy-making decisions by leading professional institutions. The US Food and Drug Administration (FDA) in 2016 approved the methylated SEPT9 DNA plasma assay, m*SEPT9*, also known as Epi proColon® (Epigenomics AG, Berlin, Germany), after they initially decided not to approve it in 2014. Currently, the Centers for Medicare and Medicaid Services (CMS) are considering a National Coverage Determination regarding payment (2). Ultimately, the test will be considered in practice recommendations by the US Preventive Services Task Force (USPSTF) and others.

In this issue of the Journal, researchers use a cost-effectiveness analysis (CEA) to compare 4 innovative CRC-screening strategies. They conclude that "for people who are unwilling to be screened with fecal immunochemical testing (FIT) or colonoscopy, annual screening with the m*SEPT9* is the test of choice given its cost-effectiveness profile compared with CTC (computed tomographic colonography), PillCam (capsule endoscopy) and mtSDNA (multitarget stool DNA, comprised of FIT and stool DNA, also known as Cologuard® [Exact Sciences Corporation, Madison, WI])" (3).

Modeling analyses, often used by CMS and USPSTF, are designed to simulate a randomized controlled clinical trial (RCT) that quantitatively assesses outcomes—that is, harms and benefits of screening such as CRC mortality reduction—to compare new testing strategies with others like FIT and colonoscopy that are supported by evidence from RCTs and are considered to be cost-effective. Modeling typically incorporates a new test's performance features, particularly sensitivity and specificity, to project outcomes of benefits, harms, effort, and sometimes cost.

This editorial considers policy decisions about m*SEPT9* that may have consequences that are unintended, unusual, and clinically significant. Three questions are addressed: What are m*SEPT9*'s sensitivity and specificity, and what are implications of a high false-positive rate (20%) for a test intended to be used by persons "unwilling" to have a colonoscopy? How does a test with a low degree of discrimination manage to achieve such a high degree of cost-effectiveness in CEA modeling? And what important unintended consequences may arise as a result of upcoming policy determinations about m*SEPT9* by CMS and USPSTF?

What Are m*SEPT9*'s Sensitivity and Specificity, and What Are the Implications of a High False-Positive Rate (20%) for a Test Intended to Be Used by Persons "Unwilling" to Have a Colonoscopy?

In its 2014 evaluation, the FDA raised concerns about m*SEPT9*'s sensitivity, specificity, and adherence that will become relevant to policy-making decisions of the CMS and USPSTF as well as to doctors and patients. The FDA in 2014 did not approve m*SEPT9* following an FDA Advisory Committee's divided vote. Studies presented by principal investigators Potter (4) and Johnson (5), in support of the sponsor's application, showed 68% sensitivity of m*SEPT9* for CRC and specificity of 80% in a screening setting. Test performance was based on a post hoc analysis of plasma samples selected from 7941 specimens (4) collected prospectively in persons scheduled for screening colonoscopy in the PRECEPT study, which had evaluated an earlier version of the m*SEPT9* assay (6).

The FDA was concerned that the 20% false-positive rate could cause many persons without CRC to be recommended for colonoscopy—a problem for a test intended for persons

Received: June 24, 2020; **Accepted:** July 21, 2020

unwilling to have colonoscopy or FIT. Another concern was that an advised colonoscopy might not be accepted after a positive m*SEPT9*. The FDA asked the sponsor to assess adherence among persons "unwilling" to have FIT or colonoscopy, that is, among intended screenees. The Liles study (7) assessed 182 compensated participants who completed the m*SEPT9* test. Among 30 participants with a positive m*SEPT9*, 23 had a colonoscopy or agreed to one by the end of the study (7). Clear detail is not provided, however, about how participants' "unwillingness" to have FIT or colonoscopy was assessed. A "scripted telephone interview" was conducted "to confirm selected inclusion criteria," but the wording of the script was not reported (7). Further, Liles (7) notes, "We did not explain false-positive rates of the fecal and blood tests to participants before enrollment; we wanted to deliver succinct counseling to simulate primary care." It is hard, then, to understand the magnitude of the problem of adherence without knowing "how unwilling" participants were to have colonoscopy at enrollment.

When the FDA approves a diagnostic test, it may issue a Summary of Safety and Effectiveness Data (SSED), where some of these issues might be addressed (8). The SSED is written by the sponsor and then approved by the FDA as part of the product's labeling. The 2016 SSED devotes most of its 41 pages to technical issues, including limits of detection, specimen stability, and reproducibility (8). Although clinical performance data are summarized from the Potter (4) and Johnson (5) publications, the SSED does not cite those publications or provide detail about possible FDA supervision or review of the Potter study (eg, about how blinding may have been assured in a post hoc analysis). Nor does the SSED discuss how well the FDA's adherence question may have been addressed by Liles (7). This lack of clarity is striking because the FDA's instruction template provided to sponsors writing a SSED states to "Describe any information submitted by the applicant in response to outstanding issues, and whether the response was found acceptable" (9).

FDA decision making has substantial impact because the FDA is regarded as a kind of "gatekeeper" in the overall test evaluation process, providing what is widely perceived as a "seal of approval." For example, in online public comments requested by the CMS to consider its National Coverage Determination for m*SEPT9*, one-half of the 67 comments mentioned "FDA approval" as a basis for supporting the test; for example, "The FDA approval ensures the 'safety and efficacy' of the test" (10). Similarly, investigators of the CEA in this issue of the Journal focus on "recently developed FDA-approved tests" (3).

The FDA's process for test approval differs substantially from its widely understood process for drug approval that requires strong evidence—that is, from an RCT—to measure benefits and harms and to quantify effectiveness (11). For test approval, the FDA does not require direct RCT evidence, nor does the FDA project benefits and harms using quantitative modeling as does the USPSTF. Although the word "effectiveness" commonly refers to long-term benefits and harms (12), the word when used by the FDA in evaluating a test may refer to test discrimination: "…analyses of effectiveness based on sensitivity and specificity" [(8), page 27]. Although FDA "approval" carries great weight, the implications of that approval are limited.

How Does a Test With a Low Degree of Discrimination Manage to Achieve Such a High Degree of Cost-Effectiveness in CEA Modeling?

The m*SEPT9* test has limited ability to discriminate cancer from no cancer because its positivity rate is 68% in persons with CRC and 20% in persons with no colonic disease. Because the false-positive rate is 20%—not 2% or 0.2%—the ratio of false positives to true positives is extraordinarily high: 37.8 false positives per true-positive result compared with 5.4 for FIT [ref (8) page 33], even though the 2 tests have similar sensitivity for CRC [68% for m*SEPT9* and 74% for FIT (3)]. Although colonoscopy following a positive m*SEPT9* may find adenomas that may be precursors to CRC, m*SEPT9* cannot be credited with finding them, because m*SEPT9* has the same 20% positivity rate both in persons with no colonic lesions and in persons with even advanced adenomas (4).

The way that a test with such a low level of discrimination can be so cost effective in a CEA is simple to understand. A test with a 20% positivity rate when applied yearly [the "most cost-effective" strategy among the 4 new tests, according to the CEA (3)] leads to a cumulative positivity rate of 50% at 3 years and 70% at 5 years. Thus, virtually all participants are enrolled in colonoscopy within a few years of receiving yearly m*SEPT9*. A testing program that soon puts everyone into colonoscopy will necessarily have cost-effectiveness roughly similar to that of colonoscopy.

The mechanism by which chance may result in a positive outcome has been called "serendipity" and is not new in the example of m*SEPT9*. For m*SEPT9*, however, the magnitude of the mechanism is dramatic because its false-positive rate is so high. Other examples of serendipity include detection of small colonic adenomas by guaiac-based fecal occult blood screening (13) and detection of prostate cancer by digital rectal exam screening (14). Further, serendipity is responsible for a portion of the mortality reduction achieved in the US-based RCT (15) of guaiac-based stool screening for CRC (16).

Clarity in discussing such implications could help interpret results of the CEA as well as upcoming deliberations of the CMS and USPSTF. For example, is it logical to suggest, on the basis of the CEA, that "people unwilling to be screened with FIT or colonoscopy" should receive "annual screening with the m*SEPT9*, on the basis of its cost-effectiveness profile" (3), if "unwilling" people will be recommended, within a few years, to have colonoscopy?

In sum, although the m*SEPT9* test is technically cost-effective, it achieves cost-effectiveness by an unusual mechanism—a high false-positive rate that causes people, by chance, to become enrolled in colonoscopy screening.

What Important Unintended Consequences May Arise as a Result of Upcoming Policy Determinations About m*SEPT9* by CMS and USPSTF?

Because decisions by the CMS and USPSTF have such major policy implications, clinically significant unintended consequences might result from their determinations. What consequences should policy makers anticipate, deliberate, and clearly explain?

For doctors and patients, the implications of endorsing a screening strategy that operates by "chance" would be relatively straightforward to address, if perhaps a bit awkward. Doctors could explain to patients that using the test on a yearly basis would result in a 50% chance of having a colonoscopy within 3 years and 70% within 5 years, so that using the test is at best buying a small amount of time before colonoscopy is conducted.

For policy makers such as the CMS and USPSTF, the implications are more dramatic. Is this the mechanism that we, as

policy makers, want to use to get people to have a colonoscopy—that is, by practically "tricking" them because of a false-positive test result? There is something unsettling about this logic.

A further consequence—unusual, unintended, and potentially important—may be the impact of m*SEPT9* adoption on test developers like academic laboratories and companies. If a screening test can be successful and adopted because of its high false-positive rate, then why not simply lower the cutoff levels of existing tests such as FIT or mtSDNA or of the next new noninvasive test for CRC so that greater numbers of people are recommended to have colonoscopy? That interpretation could have a chilling effect on efforts to develop screening tests that actually achieve a high degree of discrimination. At an extreme, a test developer could see its goal as to create a test with a high positivity rate that people "believe in" and so, for example as in this instance, will receive colonoscopy if the test is positive. Are these the precedents that policy makers like CMS and USPSTF want to encourage?

Because approval decisions of leading policy-making institutions like the FDA, CMS, and USPSTF are so influential, their determinations must be clearly described and explained, including consideration of possible unintended consequences.

Funding

U24 CA086368/CA/NCI/NIH HHS/United States; Early Detection Research Network and UL1 TR001111/NCAT/NIH HHS/United States; and North Carolina Translational & Clinical Sciences Institute provided funding.

Notes

Role of the funders: The funders had no role in the writing of this editorial or the decision to submit it for publication.

Disclosures: I was senior author and chair of the Clinical Study Steering Panel of the 2014 PRESEPT study of Epigenomics' methylated SEPT9 DNA plasma assay (6) that provided specimens and clinical data for the post hoc m*SEPT9* analysis by Potter (4). I was a senior author of the 2014 DEEP-C study of Exact Sciences' mtSDNA test (17). For neither study did I receive, from a sponsor or any other source, financial support for consulting, salary, equity, or other. I am a member of FDA's Immunology Devices Panel, Center for Devices and Radiological Health (CDRH). In 2017 and 2018, I consulted for GRAIL, and for Exact Sciences on a topic not related to CRC screening. I consult for Delfi and for Mercy Bioanalytics about screening test development.

References

1. Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *N Engl J Med*. 1978;299(17):926–930.
2. Centers for Medicare and Medicaid Services. National Coverage Analysis (NCA) tracking sheet for screening for colorectal cancer—blood-based biomarker tests (CAG-00454N). https://www.cms.gov/medicare-coverage-database/details/nca-tracking-sheet.aspx?DocID=299&DocType=nca&DocName=Screening+for+Colorectal+Cancer+-+Blood-Based+Biomarker+Tests&NCAId=299&bc=AAAAAAAAABAA&. Accessed June 20, 2020.
3. Peterse EFP, Meester RGS, de Jonge L, et al. Comparing the cost-effectiveness of innovative colorectal cancer screening tests. *J Natl Cancer Inst* 2020.
4. Potter NT, Hurban P, White MN, et al. Validation of a real-time PCR-based qualitative assay for the detection of methylated SEPT9 DNA in human plasma. *Clin Chem*. 2014;60(9):1183–1191.
5. Johnson DA, Barclay RL, Mergener K, et al. Plasma Septin9 versus fecal immunochemical testing for colorectal cancer screening: a prospective multicenter study. *PLoS One*. 2014;9(6):e98238.
6. Church TR, Wandell M, Lofton-Day C, et al. Prospective evaluation of methylated SEPT9 in plasma for detection of asymptomatic colorectal cancer. *Gut*. 2014;63(2):317–325.
7. Liles E, Coronado G, Perrin N, et al. Uptake of a colorectal cancer screening blood test is higher than of a fecal test offered in clinic: a randomized trial. *Cancer Treat Res Commun*. 2017;10:27–31.
8. Food and Drug Administration. Summary of Safety and Effectiveness Data (SSED). Blood based colorectal cancer screening test. https://www.accessdata.fda.gov/cdrh_docs/pdf13/P130001b.pdf. Accessed June 20, 2020.
9. Food and Drug Administration. FDA Summary of Safety and Effectiveness Data (SSED) Office of Device Evaluation template. https://www.fda.gov/media/113810/download. Accessed June 20, 2020.
10. Centers for Medicare and Medicaid Services. View public comments for screening for colorectal cancer—blood-based biomarker tests. https://www.cms.gov/medicare-coverage-database/details/nca-view-public-comments.aspx?NCAId=299&NcaName=Screening+for+Colorectal+Cancer+-+Blood-Based+Biomarker+Tests&ExpandComments=n&DocID=299&DocType=nca&DocName=Screening+for+Colorectal+Cancer+-+Blood-Based+Biomarker+Tests&bc=AAAAAAAAABAA&. Accessed June 20, 2020.
11. Food and Drug Administration. Guidance for industry-providing clinical evidence of effectiveness for human drug and biological products. https://www.fda.gov/media/71655/download (Accessed June 20, 2020).
12. Cochrane AL. *Effectiveness and Efficiency: Random Reflections on Health Services*. London: Nuffield Provincial Hospitals Trust; 1972.
13. Ransohoff DF, Lang CA. Small adenomas detected during fecal occult blood test screening for colorectal cancer. The impact of serendipity. *JAMA*. 1990;264(1):76–78.
14. McNaughton Collins M, Ransohoff DF, Barry MJ. Early detection of prostate cancer. Serendipity strikes again. *JAMA*. 1997;278(18):1516–1519.
15. 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(19):1365–1371.
16. Lang CA, Ransohoff DF. Fecal occult blood screening for colorectal cancer. Is mortality reduced by chance selection for screening colonoscopy? *JAMA*. 1994;271(13):1011–1013.
17. Imperiale TF, Ransohoff DF, Itzkowitz SH, et al. Multitarget stool DNA testing for colorectal-cancer screening. *N Engl J Med*. 2014;370(14):1287–1297.