Multicancer Screening: One Size Does Not Fit All

Girish Putcha, MD, PhD[1]; Alberto Gutierrez, PhD[2]; and Steven Skates, PhD[3]

Cancer is the second leading cause of death in the United States,[1] and early detection represents one of the best hopes for reducing cancer-associated morbidity and mortality. However, current screening strategies are limited by suboptimal adherence, low sensitivity for early-stage disease, high false positive rates, and varied cost-effectiveness. Blood-based multicancer screening tests offer an appealing means of streamlining the screening process to improve adherence and cost-effectiveness. Recent studies evaluated two such tests,[2,3] yet substantial barriers remain for test development, validation, and implementation. For example, both tests demonstrate low sensitivity for early-stage cancer (with no characterization of performance for precancerous lesions) and unclear risk-benefit. Furthermore, these tests will require significant changes to clinical workflows. In an effort to realize the promise of a multicancer test, we discuss notable challenges and potential solutions for implementation in a clinically responsible manner.

DIFFERENT INTENDED USES REQUIRE DIFFERENT TEST PERFORMANCE CHARACTERISTICS

Cancer sounds like one disease but is actually many: there are more than 100 different cancers, each with multiple subtypes reflecting different underlying molecular pathophysiologies. Cancers are caused by dysregulation of diverse biological pathways, evolve at varying rates, have distinct diagnostic workups, and can have vastly different clinical outcomes.[4] Indeed, this heterogeneity affects screening, diagnosis, and treatment of each cancer.[4,5] Just as the benefits of an effective cancer screening test are clear (ie, reducing cancer-related morbidity and mortality), so too are the harms (ie, adverse effects resulting from false positives and false negatives). It is precisely because the invasiveness and therapeutic effectiveness of diagnosis and treatment—and therefore the consequences for the patient of inaccurate results—differ significantly for each cancer that there is no one-size-fits-all performance requirement for a multicancer screening test. For example, the impact to the patient of a false positive screen for colorectal cancer (CRC) resulting in an unnecessary colonoscopy is meaningfully different from the unnecessary major abdominal surgery that results from a false positive screen for pancreatic or ovarian cancer. Performance requirements are cancer-specific and depend on how and where in the care pathway the test is used (Fig 1). For a first-line test, some cancers may require greater sensitivity at a clinically acceptable specificity, whereas others may require very high specificity at a clinically acceptable sensitivity because of the benefits and risks of the subsequent diagnostic workup. Furthermore, performance characteristics depend on whether the test precedes, complements, or follows an accepted method of screening, or represents a new frontline screen for an otherwise unscreened cancer, either in an asymptomatic, average-risk, or symptomatic, high-risk individual.

CLINICAL UTILITY OF MULTICANCER SCREENING HAS NOT BEEN DEMONSTRATED IN THE SAME POPULATION

Clinical utility for any test requires that the results change patient management and improve health outcomes. A cancer screening test must accurately identify the cancer at a time when clinical intervention is beneficial to the patient without causing undue harm. These considerations underpin cancer screening recommendations from guideline bodies, such as the US Preventive Services Task Force (USPSTF). Following thorough risk-benefit assessments, the USPSTF recommends screening for only four cancers (cervical, colorectal, breast, and lung) and even goes so far as to recommend against screening for three cancers (ovarian, pancreatic, and thyroid), noting that there is no net benefit of screening or that the harms outweigh the benefits.[6] Even among the four cancers recommended for population screening, the populations recommended are not identical, which begs the question: What population should be screened with a multicancer screening test and is there clinical utility in that population?

PATH TO IMPLEMENTATION OF MULTICANCER SCREENING TESTS IN REAL-WORLD CLINICAL PRACTICE IS UNCLEAR

Integration into patient care pathways is another important consideration. For example, will the test replace, be a gatekeeper to, or supplement existing cancer screening tests? Any new cancer screening paradigm faces educational and behavioral barriers from patients, providers, and payers, but multicancer screening tests will face unique challenges. Since the diagnostic and therapeutic odyssey following a positive result depends on anatomically localizing the cancer, identification of the tissue of origin is required for any such test. Without localization, patients will

Author affiliations and support information (if applicable) appear at the end of this article.

Accepted on February 25, 2021 and published at ascopubs.org/journal/po on April 1, 2021: DOI https://doi.org/10.1200/PO.20.00488

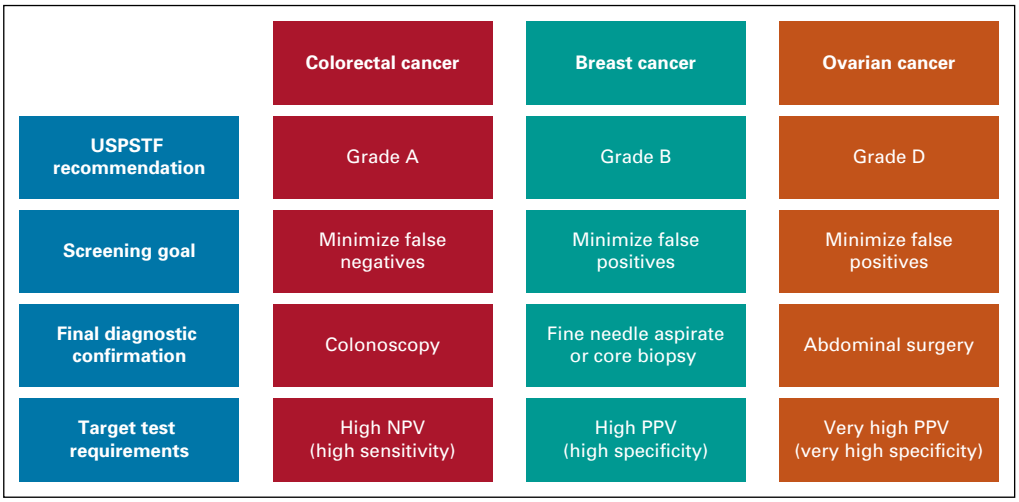

FIG 1. One size does not fit all in cancer screening. According to the USPSTF, CRC has an A recommendation (high certainty that net benefit is substantial), breast cancer has a B recommendation (high certainty that net benefit is moderate, or moderate certainty that the net benefit is moderate to substantial), and ovarian cancer has a D recommendation (moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits). The second row describes the screening goal and outlines the impact to the patient. The third row depicts the current diagnostic test for each cancer, and the fourth row describes target test characteristics. CRC, colorectal cancer; NPV, negative predictive value; PPV, positive predictive value; USPSTF, US Preventive Services Task Force.

understandably feel anxious. Undergoing additional standard-of-care screening tests, or even non–standard-of-care ones like positron emission tomography-computed tomography, to determine the tissue of origin may not only exacerbate this anxiety (especially if the tissue of origin remains unknown) and increase costs but also result in incidental findings that trigger new diagnostic odysseys.

CHALLENGES REMAIN FOR THE DEVELOPMENT AND VALIDATION OF MULTICANCER SCREENING TESTS

Given the low prevalence of individual cancers in an asymptomatic, average-risk population, large sample sizes are required to accurately assess test validity and utility. For example, in a validation study for a stool-based CRC screening test,[7] nearly 13,000 patients were enrolled to obtain only 65 cancers—a number that may be sufficient for validation but is grossly insufficient for test development. As a result, case-control designs using post-diagnosis samples are common. Although such designs may be necessary for test development, they are inappropriate for validation because they consistently overestimate clinical performance in the intended use population. For example, many case-control studies, especially those using post-diagnosis samples, suffer from spectrum bias across multiple dimensions. Cases and controls often include only the extreme ends of the disease spectrum (ie, cancer and healthy), not precancerous and benign lesions or other comorbidities; even when they do, the distribution is often different from that observed in the screening population (eg, more late-stage cancers). Moreover, many such studies use controls who are only negative by history or self-report.

SOLUTIONS TO THESE CHALLENGES EXIST BUT MUST BE ADDRESSED BEFORE IMPLEMENTATION

The potential impact of multicancer screening tests on health outcomes and the economics of cancer screening is compelling, although significant challenges for introducing them into mainstream clinical practice remain. These challenges can and should be addressed before implementation to safeguard patients. First, the test performance requirements for individual cancers must be established from robust risk-benefit analyses. In particular, modeling studies that define optimal cancer-specific test performance characteristics can inform development goals, study designs, and implementation of appropriate care pathways. Second, test development and validation for low-prevalence cancers could benefit from novel evaluation and test deployment strategies. For example, building the infrastructure to enable a learning healthcare system in which one can get diagnostic confirmation from real-world data sources (eg, medical and billing records) after completion of a multicancer screening test will permit population-scale studies that can complement traditional clinical trials. Such clinical data feedback loops may also enable future improvements in the same tests over time, especially for tests based on artificial intelligence and/or machine learning. Third, inclusion of appropriate surrogate study end points (eg, cancer stage shift, morbidity reduction, and quality-of-life improvements) in addition to mortality reduction should be considered. Earlier detection reduces cancer-specific[1] and, in some cases, overall mortality[8]; therefore, for cancers that demonstrate a strong correlation between reduction in late-stage incidence and

mortality, the use of cancer stage shift as a surrogate end point provides reasonable evidence of clinical utility and can reduce study duration.[9] Fourth, the careful creation and annotation of large sample and/or data resources (including real-world data) by groups like the Early Detection Research Network or the National Cancer Institute are critical to properly develop and validate tests for certain rare cancers. Finally, we must carefully reconsider the appropriate regulatory tools (eg, premarket controls and postmarket commitments) and reimbursement frameworks (eg, coverage with evidence development) that will allow and even foster development of multicancer screening tests while protecting patients. By recognizing and systematically addressing these significant challenges, the promise of multicancer screening tests can and will be realized.

AFFILIATIONS

[1]Freenome, South San Francisco, CA
[2]NDA Partners, Rochelle, VA
[3]Harvard Medical School and Massachusetts General Hospital, Boston, MA

CORRESPONDING AUTHOR

Girish Putcha, MD, PhD, Freenome, 279 E Grand Ave, South San Francisco, CA 94080; Twitter: @Freenome;
e-mail: authors@freenome.com.

AUTHOR CONTRIBUTIONS

Conception and design: All authors
Data analysis and interpretation: All authors
Manuscript writing: All authors
Final approval of manuscript: All authors
Accountable for all aspects of the work: All authors

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated unless otherwise noted. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/po/author-center.

Open Payments is a public database containing information reported by companies about payments made to US-licensed physicians (Open Payments).

Girish Putcha
Employment: Freenome
Leadership: Freenome
Stock and Other Ownership Interests: Freenome

Alberto Gutierrez
Employment: NDA Partners
Stock and Other Ownership Interests: NDA Partners
Consulting or Advisory Role: Freenome, GRAIL, Guardant Health, Mercy Bioscience, and Flagship Pioneering

Steven Skates
Employment: Massachusetts General Hospital
Stock and Other Ownership Interests: SISCAPA Assay Technologies
Consulting or Advisory Role: Abcodia, Guardant Health, Freenome, Mercy BioAnalytics
Research Funding: Mercy BioAnalytics, Massachusetts General Hospital co-licensed software to Abcodia

No other potential conflicts of interest were reported.

ACKNOWLEDGMENT

The authors gratefully acknowledge editorial support from Lauren N. Carroll and Signe Fransen.

REFERENCES

1. Siegel RL, Miller KD, Jemal A: Cancer statistics, 2020. CA Cancer J Clin 70:7-30, 2020
2. Liu MC, Oxnard GR, Klein EA, et al: Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA. Ann Oncol 31:745-759, 2020
3. Lennon AM, Buchanan AH, Kinde I, et al: Feasibility of blood testing combined with PET-CT to screen for cancer and guide intervention. Science 369:eabb9601, 2020
4. Loud JT, Murphy J: Cancer screening and early detection in the 21st century. Semin Oncol Nurs 33:121-128, 2017
5. Pinsky PF: Principles of cancer screening. Surg Clin North Am 95:953-966, 2015
6. Smith RA, Andrews KS, Brooks D, et al: Cancer screening in the United States, 2018: A review of current American Cancer Society guidelines and current issues in cancer screening. CA Cancer J Clin 68:297-316, 2018
7. Imperiale TF, Ransohoff DF, Itzkowitz SH, et al: Multitarget stool DNA testing for colorectal-cancer screening. N Engl J Med 370:1287-1297, 2014
8. Swartz, AW, Eberth, JM, Josey, MJ, et al: Reanalysis of all-cause mortality in the US Preventative Services Task Force 2016 evidence report on colorectal cancer screening. Ann Intern Med 167:602-603, 2017
9. Cuzick, J, Cafferty, FH, Edwards, R, et al: Surrogate endpoints for cancer screening trials: General principles and an illustration using the UK Flexible Sigmoidoscopy Screening Trial. J Med Screen 14:178-185, 2007