

**EVALUATION OF BLOOD-BASED  
COLORECTAL CANCER SCREENING  
AND DIAGNOSTIC ASSAYS:  
CONSIDERATIONS AND BEST PRACTICES**



## SUMMARY

One of the most promising strategies to combat malignant disease is early detection, in which cancers are identified and eliminated at an earlier, more treatable stage. Early detection is made possible through population-level screening and surveillance prior to clinical appearance of symptoms, which often signal more advanced local, regional, or distant disease. Colorectal cancer (CRC) screening is a prime example of the potential of early detection. The well-described, stepwise progression to CRC enables successful identification of high-risk precursor lesions, and guideline-supported screening strategies have been shown to decrease CRC incidence. Nevertheless, many patients still go unscreened and CRC remains a leading cause of cancer-related deaths.

Recently, significant advances in medical technology have renewed interest in the development of blood-based CRC screening and diagnostic assays. Blood-based cancer screening could be effective, convenient, and inexpensive, and many biomedical companies are developing assays with promising preliminary results. However, it can be challenging to properly contextualize the scientific information describing the assays and more accurately determine their potential clinical feasibility and impact on patient care.

The purpose of this article is to highlight key considerations for critically evaluating the scientific premise, research methodology, clinical application, and commercialization potential of blood-based cancer screening and diagnostic assays, with an emphasis on CRC and adenoma detection. The content represents an evidence synthesis and expert opinions from a discussion board of scientists, physicians, and other subject-matter experts affiliated with Exact Sciences and Mayo Clinic.

Ultimately, better understanding the clinical applicability and relevance of the scientific evidence on blood-based cancer assays that are under development will strengthen confidence both in interpreting reported data and identifying candidates with real capability to detect cancer early and potentially save lives.

## INTRODUCTION

Colorectal cancer (CRC) is the second deadliest cancer in the United States,<sup>1</sup> and regular screening to detect CRC is an effective strategy to reduce its incidence and improve overall survival.<sup>2</sup> There are different screening modalities that are recommended by guideline organizations, including relatively invasive structural strategies such

as colonoscopy (COL) and flexible sigmoidoscopy (SIG), and non-invasive stool-based strategies such as the guaiac-based fecal occult blood test (gFOBT), fecal immunochemical test (FIT) and the multi-target stool DNA (mt-sDNA) assay, Cologuard®.<sup>2,3</sup> These established screening modalities have been introduced, endorsed and adopted for widespread clinical application based on published data that refers to performance, adherence, access, and other relevant attributes of assay development.<sup>3-12</sup> Accordingly, newer candidate CRC screening modalities such as blood-based assays must be similarly supported by robust scientific evidence to demonstrate clinical utility. They should also complement and expand upon the current landscape of screening modalities, including providing a compelling value proposition during shared decision-making between patients and their providers, to significantly contribute to the future of CRC screening.<sup>13</sup>

When evaluating published evidence related to blood-based screening assays, the discussion board members recommended first considering where the evidence falls in the assay development process, and second the publication type and strength of the corresponding evidence.

In general, there are three categories of assay development studies: (1) exploratory, in which initial assay development is described and modeled; (2) pivotal, in which assay performance is determined and usually compared against current standard-of-care in the intended use population; and (3) post-market, in which assay effectiveness is assessed and modeled in real-world clinical practice (**Figure 1**). (These categories generally align with the National Cancer Institute's Early Detection Research Network's five phase framework for biomarker development, in which Phase IV studies would be considered pivotal.<sup>14</sup>) These categories describe a broad continuum of scientific evidence, so identifying where a particular study falls along this continuum enables a better contextualization and interpretation of the potential clinical utility of a given assay or methodology.

Second, there are three main types of scientific publications—abstracts, preprints, and manuscripts—and each type has a different expectation of evidence completeness (**Table 1**). For example, data presented in an abstract are generally more premature and less scientifically rigorous than what is required for publication in a peer-reviewed journal. Abstracts are characterized by novelty and timeliness and should signal the development of a corresponding manuscript, without the implication of providing conclusive data on assay performance or capability.

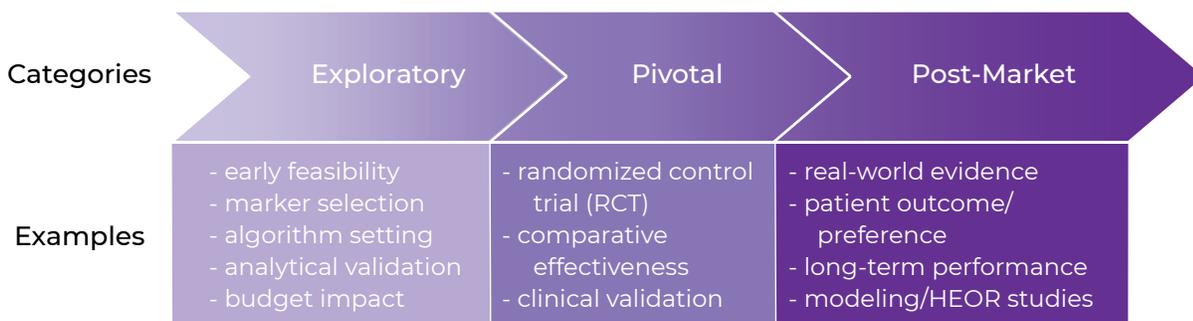
In order to improve the quality of scientific publications, multiple study reporting guidelines have been developed—CONSORT for randomized trials, STARD for diagnostic/prognostic studies, and STROBE for observational studies—that should be applied when evaluating corresponding publications. These guidelines usually provide a checklist of essential items that should be included in publications to enhance their completeness and transparency. The latest versions of these guidelines are available on the EQUATOR

(Enhancing the QUALity and Transparency Of health Research) Network’s website (<https://www.equator-network.org/>).

The discussion board members also developed a list of supplemental questions to consider when evaluating scientific data, many of which are particularly relevant to blood-based screening assays (**Table 2**). Specific considerations related to CRC blood-based assays are described throughout the article.

**FIGURE 1. SEQUENTIAL CATEGORIES OF STUDIES ASSOCIATED WITH THE DEVELOPMENT OF DIAGNOSTIC AND SCREENING ASSAYS.**

Guidelines for best-practices reporting for many of these studies can be found online through the EQUATOR (Enhancing the QUALity and Transparency Of health Research) Network (<https://www.equator-network.org/>). HEOR, health economic and outcomes research.



**TABLE 1. CHARACTERISTICS OF THE MOST COMMON TYPES OF SCIENTIFIC PUBLICATIONS.**

Publication Type	Definition	Location	Length	Peer-reviewed?	Quality of Evidence
Abstract	Short description of emergent research data	Scientific conference (poster or oral presentation)	Short	Partial <sup>1</sup>	Low
Preprint	Early version of manuscript (often before journal submission)	Public repository (e.g. medRxiv)	Long	No	Medium
Manuscript	Narrative description of scientific data	Scientific journal	Long	Yes	High

<sup>1</sup>Selection panels review abstracts prior to acceptance to a meeting or publishing online.

**TABLE 2. EXAMPLE CONSIDERATIONS TO EVALUATE BLOOD-BASED BIOMARKER STUDIES.**

Characteristic	Question to Consider	Rationale for Consideration
Study design	Do the study demographics reflect the population of interest?	Results are less applicable to intended-use population (e.g. lower-than-average patient age, healthy patient bias)
	Does the intended use of the assay reflect the study population?	Assays designed for early cancer detection requires significant inclusion of early-stage cancer samples
Study population	Were cases/controls collected under the same conditions?	Differences between case and control collection procedures (e.g. different technicians, sites, transport/storage conditions) could explain outcomes
	How were cases/controls determined?	Different guidelines classify cases differently; control study participants might be healthier than intended-use population; standard-of-care might be different by country or site
	How many sites were included?	Greater site diversity increases the study's relevance to population-level testing
	What established guidelines are followed for cancer diagnosis?	Different criteria exist for standardizing disease diagnosis, pathological review, tumor staging, etc.
	How were cancers defined pathologically?	Central pathology review provides greater consensus results than relying on clinical report alone
	How are biopsy samples described?	Collection timing and volume would impact biomarker availability
Data analysis	How were samples blinded?	Blinding patient status to test outcomes can reduce bias
	How were the samples analyzed?	Different algorithm fitting strategies (AI, partition, logistic) would produce different outcomes; FDA guidelines should be followed
Results and outcomes	How does this data differ from prior publications, if applicable?	Differences in analysis, study population, etc. should be clearly described
	Is there a biological explanation to test performance?	Biomarker availability is impacted by tumor etiology and development
Clinical feasibility and impact	How many samples will be required for future development?	Establishing clinical utility might require additional samples, especially for early-stage cancer
	What is the assay's clinical feasibility?	Implementation of patient/HCP clinical workflow might be burdensome or expensive

## STUDY DESIGN



**Terms like ‘good’ don’t even apply.... There’s value to any phase of study in the development process, as long as, one, the data are accurately reported and, two, appropriately represented.**

— Paul J Limburg, MD, MPH; Mayo Clinic

Study design is the methodological and/or analytical framework underpinning a study's results and outcomes. Study design is one of the most important

considerations for an accurate interpretation of results because it significantly impacts data generalizability and applicability. Many national guideline organizations, including the U.S. Preventive Services Task Force, have defined standards for evaluating the strength of published evidence based on study design.<sup>15</sup>

As previously described, assay development usually involves the sequential description and publication of studies that fall within the exploratory, pivotal, and post-market categories (**Figure 1**). Deviations from this conventional pathway are not intrinsically detrimental but they should be explained with appropriate context in the study's publication.

There are different study designs—both traditional designs (retrospective, prospective, case-control, cohort, etc.) and more novel, biomarker-specific designs (enrichment, randomize-all, umbrella, basket, etc.)—that are associated with or appropriate for different phases of assay development, with their own strengths and limitations.<sup>16,17</sup> For example, exploratory studies are not generally required to meet the more rigorous study design expectations of pivotal studies, so retrospective case-control studies are often sufficient and appropriate for the exploratory phase of assay development. In addition, study design is also inextricably linked to a study's intended purpose, meaning that studies designed for regulatory agency approval (e.g. from the Food and Drug Administration) can be significantly different from those designed exclusively for publication in peer-reviewed journals. For these reasons, the strengths and limitations of the study, specifically with respect to study design, should be transparently described in the publication.

As long as the results are accurately described, no study type is intrinsically better or more useful than any other, although higher-order studies such as prospective, cross-sectional studies or randomized clinical trials (RCTs) are considered to generate a higher level of evidence than retrospective studies. In general, even a single, well-designed study within the development continuum is insufficient to completely address an assay's clinical applicability. Even large RCTs that function as pivotal studies will require additional, corroborating post-market studies to support clinical validity.

An important study design consideration is to provide a baseline reference against which to compare outcomes of a novel biomarker assay. In many cases, the reference test could be a well-established screening tool that is already clinically adopted as part of standard care. Although colonoscopy is considered an imperfect criterion standard for CRC screening, it is a well-accepted reference to determine true “positive” or “negative” results for a non-invasive assay.<sup>18</sup>

Notably, colonoscopy is not the only source of comparison—there are additional performance standards and screening modalities that can be used (e.g. FIT), depending on the study—but other modalities have traditionally been discounted as having less discriminating ability compared to colonoscopy. Without a high-quality reference standard, it is difficult to anchor the study in current clinical practice or effectively compare its outcomes with similar publications that use different assays. To establish validity, the performance of one assay against another should be made within the same study, setting and population. A non-head-to-head comparison between a new assay and the historical performance of another established assay should be treated as conjecture.

## STUDY POPULATION

One of the most important considerations for assay development is that the evidence generation strategy should center around a clearly defined target (or intended use) population. Usually, this is an iterative process: studies more closely approximate the intended use population throughout assay development, culminating in the pivotal study in which the study participants closely reflect the patients with whom the assay will be clinically used.

In retrospective case-control studies, controls should not be composed of individuals with an average age less than the recommended screening start age, and cases should include individuals who are otherwise eligible for guideline-supported screening. For CRC screening, the intended use population are individuals at average risk for the development of CRC (i.e., no personal history of advanced colorectal neoplasia, iron deficiency anemia, inflammatory bowel disease, pertinent polyposis syndromes or family history of early CRC). Ideally, cases should include all stages of pre-cancer and cancer, not just advanced-stage disease, and control samples should include non-malignant disease states that could occur in the intended use population (so-called “diseased controls,” which for CRC could include non-advanced adenomas, hyperplastic polyps, diverticular disease, melanosis coli, etc.).

There are multiple sources of bias that can be introduced through the selection of the case or control populations that are often underrecognized or insufficiently described. Selection bias can occur during subject recruitment, in which cases may have different exposures or risks than controls and have not been appropriately measured or accounted for. For example, cases might be recruited from oncology clinics and may not be representative of the patients seen in primary care. Another potential bias occurs during sample acquisition, in which the collection, transportation, storage, and processing of specimens can impact the stability or degradation of a particular class of biomarkers. Overall, the interpretability of a study is hampered considerably when the potential biases in population selection or sample acquisition are not properly described.

Because of their significance and prominence in assay development, pivotal studies require additional considerations,<sup>19</sup> although these considerations are not exclusive to this study category. Pivotal studies should be managed by an external principal investigator without direct conflicts of interest, and assay performance targets should be prespecified and with sufficient statistical power to demonstrate the primary outcome and highest-priority secondary outcomes. In addition, the study protocol should receive input from subject-matter experts with clinical, guideline, reimbursement, and regulatory expertise to ensure the study appropriately

considers these perspectives and to facilitate rapid adoption. The study eligibility criteria should also reflect the specified target population; for example, the pivotal trial for a blood-based, average-risk CRC screening assay should enroll asymptomatic subjects without active symptoms, prior diagnoses, predisposing conditions, or a family history indicative of increased CRC risk. Both data acquisition (i.e. clinical operators) and data analysis should be blinded, and once data collection is complete, the data should be analyzed and/or validated by an external statistician.

## DATA ANALYSIS

Regardless of design, all studies should have a well-described statistical analysis plan (SAP), which for publications would be located either in the methods or supplemental materials. Depending on study category, elements of an effective SAP should include a pre-specified estimate of performance, a detailed description of the power calculation, and a discussion of how modeling was performed, either introducing the potential possibility of overfitting or describing how overfitting was prevented. The concern is that overfitting optimizes the performance of an assay or algorithm in a particular dataset while jeopardizing its performance in future studies. The most effective approach for a new diagnostic assay or algorithm is to include external validation datasets, either an unevaluated dataset of patients not used in its development or a random splitting of the existing dataset into training and testing sets, in which performance is reported to support claims of diagnostic accuracy.

When building an algorithm to distinguish between “positive” or “negative” samples, subdividing data into training/validation datasets and test datasets is considered the gold standard to reduce bias and maximize performance. Similar to the overall study population, the test set (and ideally the training set) should resemble the intended use population, and different types of test sets are more appropriate depending on the algorithm. Distinguishing positivity and negativity can be difficult with progressive diseases like CRC but should be defined in the publication.

Finally, for clinical trials, the SAP should be prespecified to maximize study integrity and should match the methods and results in the corresponding manuscript. These prespecified analyses should incorporate the study population considerations described above and should attempt to minimize the potential for bias and variation that can be found in diagnostic accuracy studies.<sup>20</sup> Any additional, non-prespecified analyses should be considered exploratory only and explained thoroughly in the publication, although unexplained discrepancies in published studies are disappointingly common.<sup>21,22</sup>

## RESULTS AND OUTCOMES

Similar to study population considerations, defining “healthy” or “normal” and “disease” is important to estimate assay performance characteristics. For example, with respect to specificity, “criterion-standard test negative” samples should be described in a manner that convincingly demonstrates an absence of the target condition (including pre-cancerous lesions) so that the full spectrum of assay performance can be understood and compared to other screening tests.



**There are not many full-length peer-reviewed manuscripts published on CRC blood tests or adenoma detection tests.... There's just a lot of stuff that gets released in abstract-form or press-release form that never gets fully published.**

— John B Kisiel, MD, Mayo Clinic

Additionally, there is an important biological consideration to performance. The discussion board members explained that because early or advanced adenomas have limited access to the blood stream, their biomarkers could be difficult to identify in blood plasma. In this instance, controlling for potential artifacts is critical to avoid the identification of non-specific biomarkers associated with the sample collection process or other factors that are not representative of disease pathogenesis. (There are multiple historical examples of published cancer biomarkers that were subsequently considered to be methodological artifacts.<sup>23</sup>)

The most important investigative step in assay development is the pivotal study, which is principally designed to demonstrate safety and effectiveness within the assay's intended use population. Its primary endpoints should describe its clinical performance, including sensitivity and specificity. Secondary endpoints can include a determination of superiority or non-inferiority to standard-of-care and would need to be appropriately powered. Historical test performance of a comparator/reference standard from a previous publication should not be used as a benchmark comparison in separate studies of assay performance.

This concept of test performance comparison is especially important during reference standard selection and for direct comparison in the pivotal study, particularly regarding how the outcomes from the reference standard will be interpreted relative to the new assay. For example, for CRC-based assays, how will advanced adenomas be considered by the new assay (i.e. are they expected to be categorized as “positive” or “negative” similar to CRC)? Additionally, at minimum, the sensitivity

and specificity of a novel blood-based screening assay for CRC detection should be non-inferior to the current guideline-supported screening strategy with the lowest test performance, which is currently FIT.<sup>24</sup> Ideally, evidence of non-inferiority should be established in a head-to-head study in which the performance of the novel CRC blood-based assay is compared to the performance of FIT in the same study population. FIT performance characteristics should be clearly described in the publication because differences in the study population, FIT brand, and test positivity cutoff value can significantly influence sensitivity and specificity.<sup>25</sup>

There are many characteristics of CRC detection that are important to clinicians but are underreported in many scientific studies. For example, lesion size, especially for pre-cancers, is rarely reported; the inclusion of larger lesions can have significant impact on biomarker availability (larger lesions translate to greater biomarker abundance). Information about CRC and precursor sidedness is also important to include. Proximal tumors are thought to be biologically different from distal cancers<sup>26</sup>—and are historically less frequently detected or prevented by the criterion standard, colonoscopy<sup>27</sup>—and therefore sidedness can impact assay performance.

## CLINICAL FEASIBILITY AND IMPACT

Discussion board members were particularly cognizant of the challenges to commercialize and implement a blood-based screening or diagnostic assay in the clinic (**Table 2**).<sup>28,29</sup> For pan-cancer blood-based assays, it is important to consider not only what cancers are being detected, but also if the assay performance is by cancer subtype and if the cancers are being detected at a treatable or non-treatable stage. In other words, identification of cancers alone is insufficient: patient outcomes must also be improved, such as a reduction in cancer incidence or mortality, which can be demonstrated through longitudinal, post-market studies. Additional evidence, such as confirmatory performance studies in a different study population or studies addressing design improvement or better understanding safety and effectiveness, can also strengthen an assay's clinical utility.

Other aspects that hinder commercialization of newly developed diagnostic assays are related to study design and intended use. When a compendium of studies describing an assay are presented to a regulatory agency for approval, certain limitations in study design that could increase the likelihood or speed of approval might simultaneously limit how the assay can be promoted or used. Moreover, the proposed testing frequency and assay characteristics will significantly impact commercial feasibility, not only through practical implementation but also cost effectiveness and longitudinal adherence. Characteristics such as the sample collection

requirements, the number of consumables used (e.g. tubes, reagents, etc.), and the assay's testing platform and the associated cost of goods sold (COGS) will greatly impact its commercial potential and ability to successfully integrate into clinical practice.

Notably, results from seemingly promising case-control studies do not often translate into successful blood-based assays because they cannot be replicated, succumbing to many of the biases described above. Many abstracts never materialize into peer-reviewed publications,<sup>30</sup> and even fewer blood-based assays described in these abstracts become clinically viable products.

Discussion board members recommended always reviewing each scientific publication separately and with a degree of scrutiny and circumspection. For example, reviewing a manuscript from a well-known, highly regarded biomedical journal may increase its anticipated credibility, but it does not automatically connote high-quality research. Each manuscript must be evaluated independent of related publications because improperly designed studies with insufficient descriptions exist even in top-tier journals.

In addition, although industry-sponsored research is essential for the development of clinically viable assays, considering potential sources of bias, including those found in disclosure statements and methodological descriptions, is essential to instill appropriate confidence in the data.

## CONCLUSION

Evaluating characteristics such as study design, subject population, results and outcomes, clinical feasibility, and public health impact will improve study interpretability and generate more realistic expectations around scientific data (**Table 3**). Reviewing these considerations will help enable a more thorough understanding of the capability of specific blood-based biomarker assays to detect CRC at the earliest point in carcinogenesis, thereby imparting the greatest positive impact on patient care.

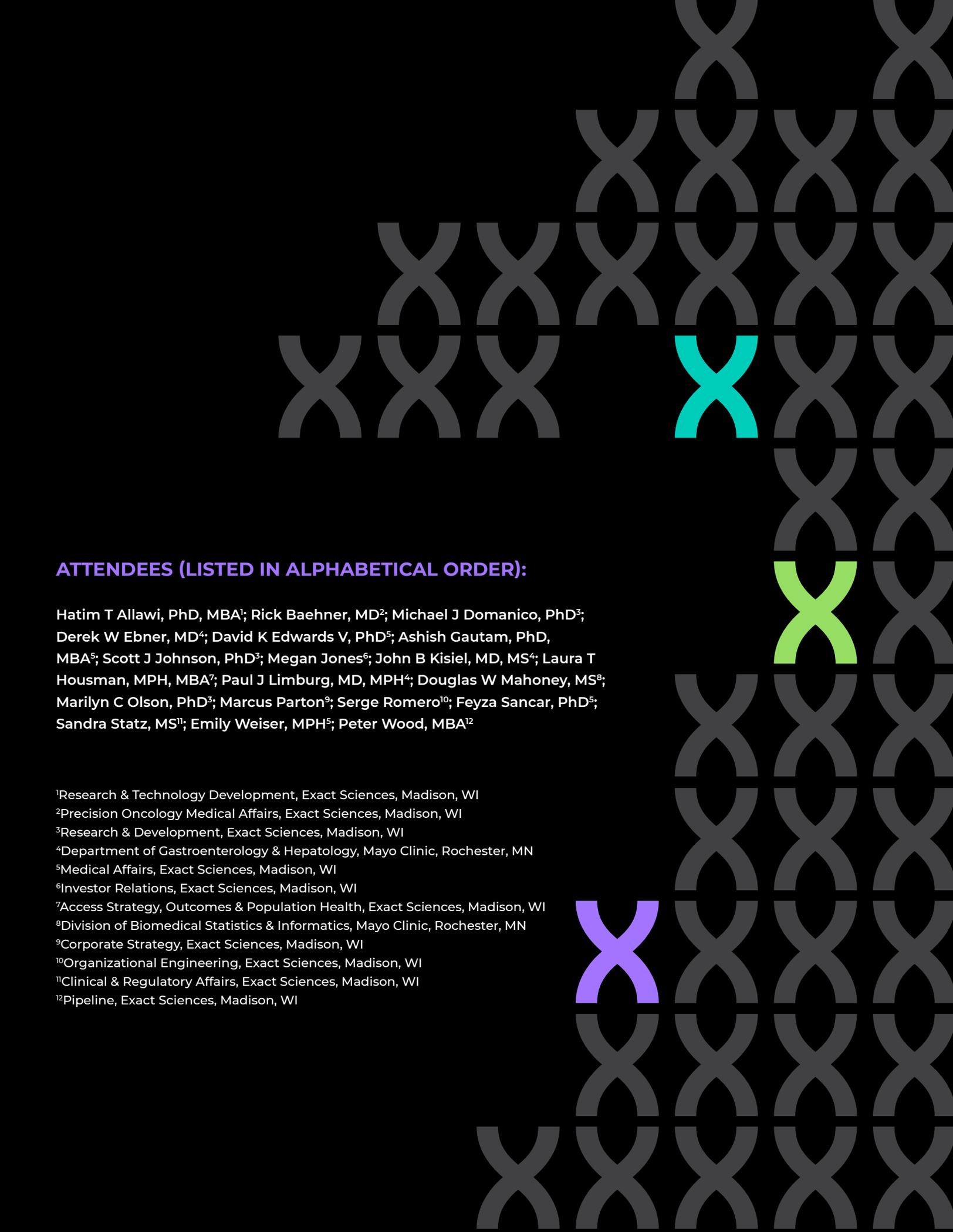
**TABLE 3. RECOMMENDED APPROACHES TO ADDRESS POTENTIAL PITFALLS IN STUDIES RELATED TO CRC BLOOD-BASED DIAGNOSTIC/SCREENING ASSAYS.**

RCT, randomized control trial; PI, principal investigator; COI, conflict of interest.

Category	Consideration	Potential Pitfall	Recommended Approach
<b>Manuscript characteristics</b>	<b>Quality</b>	Study methodology is poorly described	Follow reporting guidelines (based on study design)
	<b>Bias</b>	Study succumbs to common biases	Avoid known biases (e.g. selection bias, performance bias, etc.) through rigorous study planning and execution <sup>31</sup>
	<b>Clinical context</b>	Unproven claims about study's clinical impact in manuscript	Properly contextualize study (exploratory, pivotal, post-market); list study limitations in Discussion section
<b>Study design</b>	<b>Strengths and limitations</b>	Consequence of study design is not described in manuscript	Describe strengths and limitations of study design in manuscript; <sup>15,16</sup> examples include: <ul style="list-style-type: none"> <li>• RCTs: Randomization reduces bias; protocol should reflect clinical application and intended use population</li> <li>• Case-control studies: Efficient but prone to bias and unable to establish temporal trends</li> </ul>
	<b>Participant selection</b>	Mismatch between case/control patients (e.g. age, sex, race/ethnicity)	Justify differences in manuscript; match populations before analysis
		Patients differ from intended use population (e.g. high-risk)	Conduct sub-analysis limited to patients in intended use population; identify differences in manuscript (population must match for pivotal study)
		Cases/controls not clinically relevant (e.g. perfectly healthy controls, only advanced-stage cases)	Enroll control patients with non-malignant disease states and cases with all stages of pre-cancer and cancer
		Patients enrolled from countries with different standard-of-care than US	Enroll from multiple sites within US
		Patients enrolled only from oncology clinics	Include primary and specialty practice sites to optimize representation of screen-detected cancers in case samples
	<b>Sample collection</b>	Cases/controls were handled differently (e.g. collected, processed, shipped, stored)	Standardize shipping, handling, and sample processing; balance case and control selection to avoid collection bias; minimize reliance on third-party biobanks.
		Sample was collected after biopsy (e.g. adenomas that are removed)	Enroll some samples prospectively; wait seven days post-biopsy for sample collection
	<b>Pivotal study</b>	Additional considerations for pivotal studies were not followed	<ul style="list-style-type: none"> <li>• Have study managed by external PI without COI</li> <li>• Prespecify/appropriately power performance targets</li> <li>• Review study protocol from subject matter experts</li> <li>• Eligibility criteria reflects intended use population (e.g. asymptomatic, average-risk subjects)</li> <li>• Blind clinical operators and data analysis</li> <li>• Have external statistician analyze/validate data</li> </ul>
	<b>Data analysis</b>	<b>Model overfitting</b>	Lack of validation or independent "test" dataset for algorithm setting
<b>Assay performance</b>		Using different reference/criterion standard for measuring performance	Use a comparator test (e.g. FIT) and well-defined reference/criterion standard (e.g. central pathology, colonoscopy quality controls)
		Performance values (e.g. sensitivity and specificity) were significantly low	Test performance should be non-inferior to current guideline-supported screening strategy with the lowest test performance (e.g. FIT)
<b>Confidence in outcomes</b>		Imbalanced marker-to-sample ratio (events per predictor parameter)	Generally follow "one in ten rule" (one marker per 10 samples); justify any discrepancies
		Study outcomes are modified, and non-prespecified analyses are described as significant conclusions	Prespecify primary and secondary outcomes; describe non-prespecified analysis as exploratory only

## REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin*. 2020;70(1):7-30.
2. 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(4):250-281.
3. U. S. Preventive Services Task Force, Bibbins-Domingo K, Grossman DC, et al. Screening for Colorectal Cancer: US Preventive Services Task Force Recommendation Statement. *JAMA*. 2016;315(23):2564-2575.
4. Winawer SJ, Zauber AG, Ho MN, et al. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. *N Engl J Med*. 1993;329(27):1977-1981.
5. Atkin WS, Edwards R, Kralj-Hans I, et al. Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. *Lancet*. 2010;375(9726):1624-1633.
6. Nishihara R, Wu K, Lochhead P, et al. Long-term colorectal-cancer incidence and mortality after lower endoscopy. *N Engl J Med*. 2013;369(12):1095-1105.
7. 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.
8. Heresbach D, Manfredi S, D'Halluin P N, Bretagne JF, Branger B. Review in depth and meta-analysis of controlled trials on colorectal cancer screening by faecal occult blood test. *Eur J Gastroenterol Hepatol*. 2006;18(4):427-433.
9. 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.
10. Inadomi JM, Vijan S, Janz NK, et al. Adherence to colorectal cancer screening: a randomized clinical trial of competing strategies. *Arch Intern Med*. 2012;172(7):575-582.
11. Nielson CM, Vollmer WM, Petrik AF, Keast EM, Green BB, Coronado GD. Factors Affecting Adherence in a Pragmatic Trial of Annual Fecal Immunochemical Testing for Colorectal Cancer. *J Gen Intern Med*. 2019;34(6):978-985.
12. Weiser E, Parks PD, Swartz RK, et al. Cross-sectional adherence with the multi-target stool DNA test for colorectal cancer screening: Real-world data from a large cohort of older adults. *J Med Screen*. 2020;969141320903756.
13. Melson JE, Imperiale TF, Itzkowitz SH, et al. AGA White Paper: Roadmap for the Future of Colorectal Cancer Screening in the United States. *Clin Gastroenterol Hepatol*. 2020.
14. Pepe MS, Etzioni R, Feng Z, et al. Phases of biomarker development for early detection of cancer. *J Natl Cancer Inst*. 2001;93(14):1054-1061.
15. U. S. Preventive Services Task Force. Section 4.2 Internal and External Validity Assessment of Individual Studies. In. *USPSTF Procedure Manual*. 2020.
16. Frieden TR. Evidence for Health Decision Making - Beyond Randomized, Controlled Trials. *N Engl J Med*. 2017;377(5):465-475.
17. Janiaud P, Serghiou S, Ioannidis JPA. New clinical trial designs in the era of precision medicine: An overview of definitions, strengths, weaknesses, and current use in oncology. *Cancer Treat Rev*. 2019;73:20-30.
18. Zhao S, Wang S, Pan P, et al. Magnitude, Risk Factors, and Factors Associated With Adenoma Miss Rate of Tandem Colonoscopy: A Systematic Review and Meta-analysis. *Gastroenterology*. 2019;156(6):1661-1674 e1611.
19. U.S. Food & Drug Administration Center for Devices and Radiological Health. Design Considerations for Pivotal Clinical Investigations for Medical Devices: Guidance for Industry, Clinical Investigators, Institutional Review Boards and FDA Staff. 2013. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/design-considerations-pivotal-clinical-investigations-medical-devices>.
20. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Group Q-S. A systematic review classifies sources of bias and variation in diagnostic test accuracy studies. *J Clin Epidemiol*. 2013;66(10):1093-1104.
21. Goldacre B, Drysdale H, Dale A, et al. COMPare: a prospective cohort study correcting and monitoring 58 misrepresented trials in real time. *Trials*. 2019;20(1):118.
22. Cro S, Forbes G, Johnson NA, Kahan BC. Evidence of unexplained discrepancies between planned and conducted statistical analyses: a review of randomised trials. *BMC Med*. 2020;18(1):137.
23. Diamandis EP. Cancer biomarkers: can we turn recent failures into success? *J Natl Cancer Inst*. 2010;102(19):1462-1467.
24. Imperiale TF, Gruber RN, Stump TE, Emmett TW, Monahan PO. Performance Characteristics of Fecal Immunochemical Tests for Colorectal Cancer and Advanced Adenomatous Polyps: A Systematic Review and Meta-analysis. *Ann Intern Med*. 2019.
25. Brenner H, Qian J, Werner S. Variation of diagnostic performance of fecal immunochemical testing for hemoglobin by sex and age: results from a large screening cohort. *Clin Epidemiol*. 2018;10:381-389.
26. Jensen CE, Villanueva JY, Loaiza-Bonilla A. Differences in overall survival and mutation prevalence between right- and left-sided colorectal adenocarcinoma. *J Gastrointest Oncol*. 2018;9(5):778-784.
27. Lim DR, Kuk JK, Kim T, Shin EJ. Comparison of oncological outcomes of right-sided colon cancer versus left-sided colon cancer after curative resection: Which side is better outcome? *Medicine (Baltimore)*. 2017;96(42):e8241.
28. Diamandis EP. The failure of protein cancer biomarkers to reach the clinic: why, and what can be done to address the problem? *BMC Med*. 2012;10:87.
29. Ren AH, Fiala CA, Diamandis EP, Kulasingam V. Pitfalls in cancer biomarker discovery and validation with emphasis on circulating tumor DNA. *Cancer Epidemiol Biomarkers Prev*. 2020.
30. Scherer RW, Meerpohl JJ, Pfeifer N, Schmucker C, Schwarzer G, von Elm E. Full publication of results initially presented in abstracts. *Cochrane Database Syst Rev*. 2018;11:MR000005.
31. Pannucci CJ, Wilkins EG. Identifying and avoiding bias in research. *Plast Reconstr Surg*. 2010;126(2):619-625.



## ATTENDEES (LISTED IN ALPHABETICAL ORDER):

Hatim T Allawi, PhD, MBA<sup>1</sup>; Rick Baehner, MD<sup>2</sup>; Michael J Domanico, PhD<sup>3</sup>;  
Derek W Ebner, MD<sup>4</sup>; David K Edwards V, PhD<sup>5</sup>; Ashish Gautam, PhD,  
MBA<sup>5</sup>; Scott J Johnson, PhD<sup>3</sup>; Megan Jones<sup>6</sup>; John B Kisiel, MD, MS<sup>4</sup>; Laura T  
Housman, MPH, MBA<sup>7</sup>; Paul J Limburg, MD, MPH<sup>4</sup>; Douglas W Mahoney, MS<sup>8</sup>;  
Marilyn C Olson, PhD<sup>3</sup>; Marcus Parton<sup>9</sup>; Serge Romero<sup>10</sup>; Feyza Sancar, PhD<sup>5</sup>;  
Sandra Statz, MS<sup>11</sup>; Emily Weiser, MPH<sup>5</sup>; Peter Wood, MBA<sup>12</sup>

<sup>1</sup>Research & Technology Development, Exact Sciences, Madison, WI

<sup>2</sup>Precision Oncology Medical Affairs, Exact Sciences, Madison, WI

<sup>3</sup>Research & Development, Exact Sciences, Madison, WI

<sup>4</sup>Department of Gastroenterology & Hepatology, Mayo Clinic, Rochester, MN

<sup>5</sup>Medical Affairs, Exact Sciences, Madison, WI

<sup>6</sup>Investor Relations, Exact Sciences, Madison, WI

<sup>7</sup>Access Strategy, Outcomes & Population Health, Exact Sciences, Madison, WI

<sup>8</sup>Division of Biomedical Statistics & Informatics, Mayo Clinic, Rochester, MN

<sup>9</sup>Corporate Strategy, Exact Sciences, Madison, WI

<sup>10</sup>Organizational Engineering, Exact Sciences, Madison, WI

<sup>11</sup>Clinical & Regulatory Affairs, Exact Sciences, Madison, WI

<sup>12</sup>Pipeline, Exact Sciences, Madison, WI