FRIENDS of CANCER RESEARCH

A FRIENDS OF CANCER RESEARCH WHITE PAPER

Multi-Cancer Early Detection Screening Tests: Considerations for Use of Real-World Data

Introduction

Cancers that are detected in late stages generally have a worse prognosis compared to cancers detected in earlier stages, when tumors are more amenable to effective and even curative interventions.[1] Currently there are a limited number of cancer types with available minimally-invasive standard of care (SOC) screening approaches to detect cancer earlier, and they are designed to detect only a single cancer type.[2] As a result, many cancers may go undetected or may be detected at later stages when treatment may not be as effective and outcomes are worse. The observed mortality benefit for screened cancers[3–6] raises the possibility that safe and effective screening tests for currently unscreened cancers may reduce cancer mortality for those cancer types.

Recent innovations enable the emergence of technologies that detect the presence of multiple types of cancer from a sample of blood, i.e., a liquid biopsy. Multi-cancer early detection (MCED) screening tests are a type of liquid biopsy intended to detect cancer-associated signals at early stages, including cancers with and cancers without SOC screening modalities. Given the novel nature and the unique challenges in clinical validation associated with multi-cancer screening approaches[7], there is an opportunity to explore innovative strategies for generating and assessing evidence to robustly characterize the safety and effectiveness of MCED screening tests.

The safety and effectiveness of cancer screening tests are usually demonstrated through evidence generation by clinical screening studies which use traditional data capture methods (e.g., electronic data capture, case report forms, patient reported outcomes) and occur in a pre-specified, selected population. To date, most screening studies designed to evaluate safety and effectiveness of FDA-approved single cancer screening devices have been prospective and observational studies.[8,9] Data for some long-term clinical outcome endpoints, such as overall survival and cancer-specific mortality, have been generated from prospective randomized controlled trials (RCTs) and rigorous epidemiologic studies in the post-market setting (**Figure 1**).[3–6]

Thank You to Our Contributors

This paper reflects discussions that occurred among stakeholder groups on various challenges and opportunities related to multi-cancer early detection screening tests. The topics covered in the paper, including recommendations, are intended to capture key discussion points and should not be interpreted to reflect alignment on all topics included in the white paper by all the contributors.

Carolyn Aldige, Prevent Cancer Foundation

Seema Singh Bhan, Exact Sciences Corporation

Deepshikha Bhandari, GRAIL

Christina A. Clarke, GRAIL

Teresa Coleman, IQVIA

Michael del Aguila, GRAIL

Ruth Etzioni, University of Washington and the Fred Hutchinson Cancer Research Center

Joanne Hackett, IQVIA

Ernest Hawk, The University of Texas MD Anderson Cancer Center

Daniel F. Hayes, University of Michigan Rogel Cancer Center

Robertino Mera-Giler, Freenome

Anna Pugh, Exact Sciences Corporation

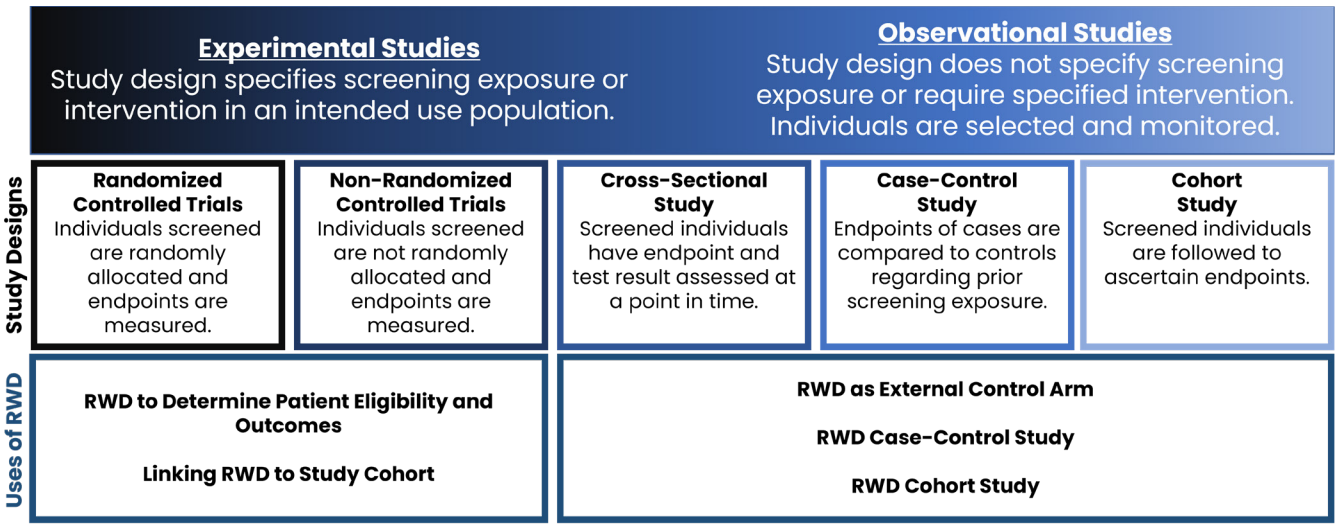
Girish Putcha, Freenome

Natalie Chavez Lau, Freenome

Sam Roosz, Crescendo Health

Sean Tunis, Rubix Health

Figure 1: Types of Study Designs for Screening Tests and Incorporation of RWD



Evidence generation for screening tests can occur through experimental and observational studies, where RWD may be incorporated in a variety of ways, including hybridized methods. This figure, which provides examples for use of RWD, is meant to be directional and not intended to be a comprehensive list of study designs and objectives.

Conducting clinical screening studies, such as RCTs, to generate the appropriate evidence of the clinical validity and utility of MCED screening tests may be logistically challenging. Appropriately powering studies for each cancer type, particularly for rare cancers, requires large enrollment numbers (i.e., on the order of tens of thousands of participants), extensive resourcing, and one or more decades of longitudinal follow-up to demonstrate a cancer-specific mortality benefit for individual cancer types across the large set of cancer types in the intended use population.[10] Additionally, highly-controlled clinical studies with protocol screening and follow-up procedures (including diagnostic procedures) may not reflect the real-world screening, adherence, and clinical practice, which also may evolve over time. To help overcome these challenges with clinical screening studies, real-world data (RWD) may be able to supplement data generated by clinical screening studies to assess MCED screening tests. RWD are data collected during the course of usual patient care and can be used to generate real-world evidence (RWE). For the purposes of this white paper, the group focused on RWD as defined by FDA: Data relating to patient health status and/or the delivery of health care routinely collected from a variety of sources, including electronic health records (EHRs), claims and billing data, and product and disease registries.[11]

The use of RWD to assess MCED screening tests to support regulatory decision-making requires careful forethought to ensure the data collected can address key assessment questions, while also acknowledging and planning for data necessary to support the test's clinical utility, as designing studies that include endpoints addressing both clinical validity and utility can

support multiple purposes (e.g., regulatory decision making, reimbursement decisions, etc.). To provide overarching considerations for generating evidence about MCED screening tests using RWD, Friends of Cancer Research (*Friends*) assembled a multi-stakeholder group of experts including government officials, MCED screening test developers, academic clinicians and researchers, patient advocacy groups, and RWD partners and vendors. We first identified endpoints to consider capturing in RWD and then reviewed opportunities for using RWD study designs to support an understanding of MCED screening test safety and effectiveness. This work complements that of others focused on assessment of MCED screening tests, exploring platform trial designs, and identifying novel endpoints for evidence generation about clinical validity and utility.

Objectives

When captured and analyzed appropriately, RWD can be used to generate RWE to evaluate the safety and effectiveness of a medical product.[11] The group focused on identifying opportunities to generate meaningful RWE to supplement evidence for regulatory decision-making for MCED screening tests, while also considering opportunities for data collection over the continuum of evidence generation. Within the context of current study designs, RWE is likely to serve a supplementary role and be part of the totality of evidence in an initial premarket application. However, as our understanding of these novel tests evolves and the robustness of RWD is better understood, the use of RWD may expand. This should ultimately be informed by conversations between regulators and sponsors.

The group's objectives were to:

- Identify potential endpoints (including performance metrics and clinical outcomes) that could be captured from RWD sources to assess the clinical validity and utility of MCED screening tests,
- Characterize opportunities and challenges associated with using RWD to support assessment of MCED screening tests, and
- Highlight key considerations for using RWD to generate RWE to support assessment of MCED screening tests.

Every assay may have unique characteristics that are not covered by this document. Further, MCED screening technology is an evolving area, and as evidence continues to build, the optimal approach for assessment of MCED screening tests may also evolve and adapt. MCED test developers are strongly encouraged to submit a pre-submission to FDA to discuss the details about their specific test.

RWE Generation for Assessment of MCED Screening Tests

MCED screening tests use various technologies to detect cancer signals, therefore evaluation approaches may differ both across MCED screening tests and when compared to current screening tests. Some MCED screening tests provide a likelihood score for the tissue of origin (TOO), sometimes referred to as the cancer signal origin (CSO), while other MCED screening tests prompt clinical follow up of positive test results using imaging modalities like whole body PET-CT to identify the TOO. Additionally, analytic approaches to determine safety and effectiveness in multiple cancers are different from a focus on a single cancer, as seen with

currently available screening tests. These differences suggest a need for a review of the current regulatory and evidence development paradigms to assess clinical validity and utility to inform potential solutions.

RWD studies may offer some logistical advantages over clinical screening study designs. RWD studies have the potential to provide data over a lengthy follow-up period to generate evidence about long-term outcomes encompassing a large number of subjects in the intended use population, including those with rare cancers. RWD also provides information reflective of the real-world population setting about diagnoses, screening frequencies, biopsy compliance, and treatment patterns, including as these may evolve over time. However, RWD is subject to its own limitations due to the observational setting and the generation of RWD for administrative and billing, rather than research purposes. These limitations can lead to issues with non-random missing data, mismeasured data, and selection bias. Despite these limitations, RWD represents an opportunity to explore and propose additional, pragmatic solutions to assess MCED screening tests.

While clinical screening studies continue to be a key component and the foundational source of evidence for in vitro devices, there may be opportunities for RWD studies to inform regulatory decisions for MCED screening tests. Previously published FDA guidance notes that RWD of sufficient quality may potentially be used to inform or support a particular regulatory decision for medical devices and diagnostics, with the specific use determined by the specific type of technology[11], including use of RWD as:

- Generating hypotheses to be tested in a prospective clinical study,
- A historical control, a prior in a Bayesian trial, or as one source of data in a hierarchical model or a hybrid data synthesis,
- A concurrent control group, or as a mechanism for collecting data in a setting where a registry or some other systematic data collection mechanism exists,
- Evidence to identify, demonstrate, or support the clinical validity of a biomarker,
- Evidence to support FDA approval or authorization,
- Support for a petition for reclassification of a medical device,
- Evidence for expanding the label to include additional indications for use or evidence to update the labeling to include new information on safety and effectiveness,
- Public health surveillance efforts,
- To conduct post-approval studies that are imposed as a condition of device approval or to potentially preclude the need for postmarket surveillance studies ordered under section 522 of the FD&C Act,
- In certain circumstances, for use in generating summary reports of Medical Device Reports (MDRs), and
- To provide postmarket data in lieu of some premarket data.[11]

Key Questions for Assessment of MCED Screening Tests

To assess MCED screening tests throughout the product life cycle (e.g., premarket, post-market data collection, benefit-risk determinations), the working group identified key questions to frame necessary evidence generation, endpoints, and data:

1. **Performance Characteristics:** How well does the MCED screening test detect cancer? How early does the test detect cancer?

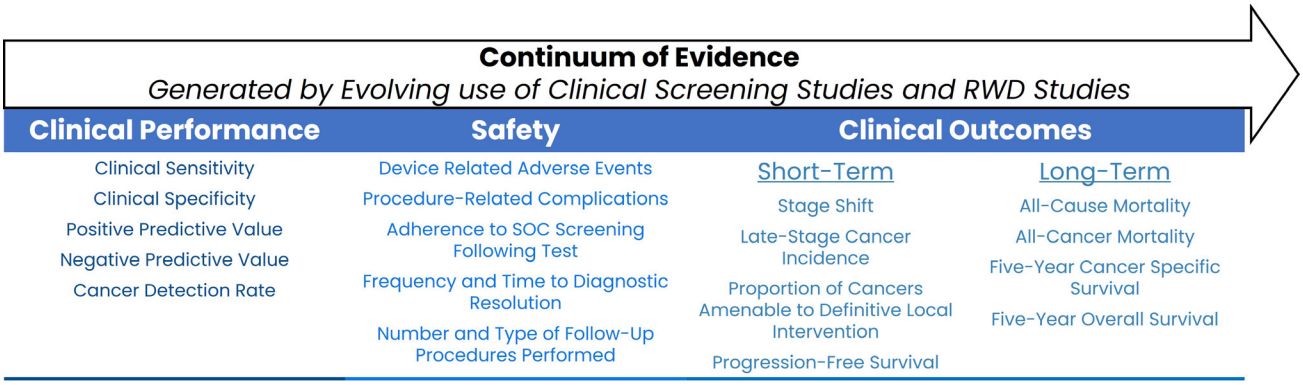
2. **Safety:** What are the health burdens/harms of MCED screening tests, including the diagnostic confirmation process?

3. **Clinical Outcomes and Utility:** How does an MCED screening test impact cancer outcomes?

Continuum of Evidence Generation

The working group mapped out these key questions in the context of the continuum of evidence generation, which can be supported by data from both prospective clinical screening studies and RWD sources (**Figure 2**). To help answer these questions, we identified a list of possible endpoints to consider.

Figure 2: The Continuum of Evidence Generation and Proposed Endpoints to Help Answer Key Questions



Evidence can be generated to assess MCED screening tests by clinical screening studies using traditional data capture methods and/or RWD sources, with the types of studies evolving over time and purpose of evidence generation. Possible endpoints to consider for generating evidence are highlighted along the continuum of clinical validity and utility. The specific endpoints necessary to establish clinical validity and utility may vary depending on the technology. The safety and clinical outcomes endpoints require the use of the test in patient management.

Intended Use Considerations for Evidence Generation

Evidence generation should be conducted in the intended use population. Designing the study plan and identifying appropriate endpoints will be influenced by the intended use of the test, as well as the interval of MCED test screening, with considerations including:

- The TOO component of the test. While MCED screening tests detect cancer-associated signals generally, regulatory expectations are for the TOO to be identified, either with TOO

ascertainment built within the assay capabilities and followed by diagnostic confirmation, or by a follow-up methodology (e.g., PET-CT) after a cancer signal is detected.

- The intended use population on the label. The MCED screening test's intended use population defined on the assay label may differ among tests, including age cut-offs, specific types of cancers detected, point of use in the clinical care pathway (e.g., complement of the test to SOC screening tests), and risk profile of individuals eligible for the test. For example, the test may be intended for only high-risk populations, including those with a genetic predisposition, occupational or environmental exposure, a history of cancer, or specific lifestyle factors.

Potential Endpoints to Evaluate MCED Screening Tests

To understand what data are needed to assess MCED screening tests in the various phases of evidence generation, the working group defined possible endpoints that assess clinical validity and clinical utility in the context of the key questions that were asked. Clinical validity is the ability of the test to accurately identify cancer, as well as identify TOO, while clinical utility is the likelihood that patients managed in accordance with test results will demonstrate improved health outcomes, such as a reduction in late-stage cancer diagnoses and mortality.[12] Many of these endpoints encompass both clinical validity and utility. Analytical validity, which confirms that the test accurately measures the target analytes in the blood, is assumed to have been established as part of product development and is not included in the scope of this work. As indicated above, specific endpoints may vary depending on the intended use population of the MCED screening test. **Appendix Table 1** provides aligned definitions for each of the proposed endpoints and is not meant to be a comprehensive list of endpoints.

Performance Characteristics: How well does the MCED screening test detect cancer? How early does the test detect cancer?

It is critical to determine that an MCED screening test detects cancer, including at an earlier stage than it would otherwise be clinically diagnosed. The evidence should show that the MCED screening test returns a positive result in individuals who have cancer (sensitivity), while also providing a negative result for individuals who do not have cancer (specificity).

Cancers vary in preclinical latency, and MCED screening tests will vary in sensitivity per cancer, and across stages, based on a variety of factors unique to each test. Therefore, performance should be reported both in the aggregated form for all cancer detection, as well as for individual cancer types, with performance stratified by stage. In general, screening test sensitivity and specificity are initially assessed via retrospective evaluations, such as case-control studies, in which pathologically confirmed cases and suspected non-cases are examined.[13,14]

The observed sensitivity and specificity in the intended use population will depend on the diagnostic accuracy of the confirmatory diagnostic test (e.g., PET-CT), which may differ for different cancers. Therefore, it is important to document the evaluation workflow, and develop methods to address the imperfect accuracy of the confirmatory tests.

In the prospective screening setting, sensitivity and specificity are challenging to assess. One method to establish true sensitivity and specificity would be to require full body imaging and pathological confirmation of all individuals for all cancers included in the test, but this method would be impractical because of an undue burden for patients. An approximation of screening test sensitivity can be given by the ratio of screen-detected cancers to the sum of screen- and interval-detected cancers at a given point in time.[15–17] This estimate of screening test sensitivity may be affected by multiple factors including overdiagnosis, preclinical latency, previous screening history, and the time interval chosen. This estimate may deviate from the true sensitivity and will not necessarily match estimates obtained from already-diagnosed cases.[18,19]

Just as an MCED screening test may exhibit variability in its ability to detect different cancers, there may also be variability in TOO accuracy. The same is true for TOO assessment by PET-CT, in which the accuracy differs for different cancer types. Therefore, performance should be reported in aggregate form for all cancers, and on a per-cancer basis based on TOO assessment.

Additional measures of diagnostic performance under prospective screening are the positive and negative predictive values (PPV and NPV). A high PPV implies a low rate of unnecessary confirmation tests or biopsies, but alone is not a reliable indicator of the likely benefit of the test. Time elapsed without a confirmed cancer diagnosis after a negative test result could be used to determine the NPV; in this case a long interval without cancer following a negative test would indicate that the test was a true negative. The necessary monitoring time for individuals with a negative test result will depend on the cancer type, its given natural history, the effectiveness of the related diagnostic workup, and the interval for any established SOC screening for the cancer type.

Performance Characteristics Endpoints Include (Calculated based on detection of cancer signal and cancer signal detection +TOO):

- *Clinical Sensitivity*
- *Clinical Specificity*
- *Positive Predictive Value*
- *Negative Predictive Value*
- *Cancer Detection Rate*

Safety: What are the health burdens/harms of MCED screening tests, including the diagnostic confirmation process?

Although not specific to MCED screening tests, FDA has released general guidance that details considerations for the assessment of probable benefit and risks/harms of a device, including the risks of adverse events directly related to the test as well as those related to the follow-up diagnostic procedures after a positive test result. Adverse events include both physical and psychological negative occurrences. Additional evidence generated from patient reported outcomes regarding quality of life and anxiety may support an understanding of these adverse events.[20] Although not specific to MCED screening tests, based on this guidance, the timing of an assessment of safety should include the interval from administration of the MCED screening test until the determination of cancer status is complete.

There are many facets of MCED screening tests' safety that will factor into the benefit-risk assessment. The first safety concern is how a positive MCED screening test result impacts an individual's health care journey due to follow-up procedures to establish a definitive diagnosis. It will be important to analyze the number and type of follow-up procedures performed, any complications, and the frequency and time to diagnostic resolution. Lack of a diagnostic resolution following a positive MCED test could lead to adverse effects on an individual's quality of life. Theoretically, if MCED screening tests detect cancers that would otherwise go undetected, in the short-term, more surgeries and procedures may occur leading to more safety concerns; however, over a longer term, the net safety profile may improve since the individual may avoid complications and costs associated with diagnosis of (and treatment for) their cancer at later stages. Stratifying the safety outcomes by cancer type will also be important, as the benefit-risk profiles for the diagnostic resolution will vary across cancer types.

The second safety concern is how MCED screening might impact SOC screening. Tests currently in development are expected to have multiple intended uses, including complementing SOC screening. For such tests, whether individuals tested adhere to SOC screening may inform their impact and implications, so SOC screening among individuals who have such an MCED screening test should be recorded to determine if there are changes.

Safety Endpoints Include:

- *Device Related Adverse Events (Physical and Psychological)*
- *Procedure-Related Complications (Physical and Psychological)*
- *Adherence to SOC Screening Following Test*
- *Frequency of Confirmation Diagnostic Tests and Time to Diagnostic Resolution*
- *Number and Type of Follow-Up Procedures Performed*

Clinical Outcomes and Utility: How does an MCED screening test impact clinical cancer outcomes?

To demonstrate that MCED screening tests improve clinical outcomes, evidence must show the test detects cancers that are otherwise undetected before symptoms appear (i.e., at earlier stages) and reduces morbidity and mortality associated with cancer and its treatment. It is important to evaluate endpoints that measure both short-and long-term outcomes.

Short-term endpoints ascertained shortly after the determination of cancer status can support evidence that the test detects cancer earlier and may ultimately translate into improvements in cancer-specific morbidity or mortality (e.g., reductions in late-stage cancer diagnosis). Defining early-stage cancer will likely be cancer type specific but can be considered to mean cancers generally amenable to local intervention for curative intent, whereas late-stage cancers usually cannot be cured via localized treatments. Stage shift is a possible surrogate for the impact of an MCED screening test on disease mortality. There is a concern that increasing the proportion of early-stage cancers may lead to overdiagnosis without any effect on late-stage detection[21]; therefore, a decrease in incidence of late-stage disease is more informative to support an understanding of the likely implications of stage shift.

Long-term endpoints require data capture over multiple years, such as survival and mortality, and are necessary because short-term endpoints may not translate into longer term reductions in morbidity and mortality. Disease-specific mortality is the primary measure of clinical utility for cancer screening trials and the most reliable indicator of whether a cancer screening test reduces deaths from cancer.[22] Other long-term and short-term endpoints can complement this primary endpoint but can be difficult to interpret on their own. All-cause mortality has been discussed as an endpoint in single-cancer screening studies but may not be sensitive enough to discern screening benefit.[23] Survival endpoints can also be difficult to interpret due to lead-time and length bias.[24] The limitations and potential biases impacting interpretability of each of these endpoints should be carefully noted. As with the clinical validity endpoints described above, assessing the clinical utility by cancer type, in addition to all-cancer, will be important.

Clinical Outcomes Endpoints Include:

- *Short-Term Endpoints*
  - *Stage Shift*
  - *Late-Stage Cancer Incidence*
  - *Proportion of Cancers Amenable to Definitive Local Intervention*
  - *Progression-Free Survival*
- *Long-Term Endpoints*
  - *All-Cancer Mortality*
  - *All-Cause Mortality*
  - *Five-Year Cancer Specific Survival*
  - *Five-Year Overall Survival*

Proposed Data Elements Necessary to Generate Evidence

Generating evidence for the endpoints described above will require rigorous data capture with appropriate ontologies and validated definitions, including specific data elements (suggestions included in **Table 1**). It will also be critical to identify and capture the selection factors that characterize the individual receiving an MCED screening test to understand the representativeness of this population and generate a list of potential confounding or selection variables for comparative and causal studies. Further, it will be important to capture any comorbidities or risk factors for cancer that the individuals have, as comorbidities may influence long-term outcomes, and risk factors for cancer can provide additional information about the risk profile of the population receiving the MCED screening test. Previous FDA guidance, not specific to MCED screening tests, for selecting the study population recommends including individuals across the entire range of disease states, with relevant confounding medical conditions, and across different demographic groups to prevent bias in estimates of test performance.[25] An obstacle to unbiased clinical utility analyses includes potential differences in the post-diagnosis treatment pathways for those who receive the test and those who do not. Therefore, collecting treatment information to understand treatment pathways following diagnosis will be valuable.

Table 1: Suggested Data Elements to Consider for Evidence Generation to Support Assessment of MCED Screening Tests

Category	Data Elements	Endpoint Category
Patient Characteristics	• Age at Time of Test • Gender • Race/Ethnicity • Socioeconomic Status • Insurance Status • Access to Care for Diagnosis and Treatment • Comorbidities	Demographics/Intended Use
Cancer Risk Factors	• Family History • Smoking History • Alcohol Use • Obesity, Diet, and Exercise • Genetic Predisposition • Prior Cancer History (Cancer Type, Diagnosis Date, Previous Treatments) • Other Risk Factors	Demographics/Intended Use
MCED Screening Test Administration	• Reason for Test Administration • Test Administered • Test Result (Positive/Negative and TOO) • Adverse Events with Administration	Clinical Performance, Safety
SOC Screening	• Adherence to Appropriate SOC Screening Methods • SOC Screening Results	Clinical Performance, Safety

Clinical Confirmation	• Imaging Recommended and Performed • Biopsy Recommended and Performed • Other Procedures Recommended and Performed for Definitive Diagnosis • Results from Definitive Diagnostic Procedures (Cancer Present/Absent) • Time to Diagnostic Resolution • Adverse Events with Confirmatory Procedures • Other Cancer(s) Detected that were Not Tested for or were a Negative Result using the MCED Test	Clinical Performance, Safety
Cancer Characteristics	• TOO • Stage • Histology • Subtype • Method of Detection (if cancer is not detected by the MCED screening test, such as clinical findings, symptoms, etc.)	Clinical Performance, Clinical Outcomes
Cancer Treatment	• Relevant Treatments for Cancer, Including Doses and Duration	Safety, Clinical Outcomes
Clinical Outcomes	• Living Status (Dead/Alive) • Duration of Clinical Follow-Up • Cause of Death, if applicable • Progression or Metastasis (and time) • Disease-Free Survival • Morbidity	Clinical Outcomes

Adapted and modified.[26]

RWD Study Design Informed by Characteristics of Cancer Types

One key factor to consider regarding the benefit-risk profile is the characteristics of the specific cancers reported by the MCED screening test. While MCED screening tests detect the presence of a cancer signal in general, it will be important to consider specific cancer characteristics including the incidence and availability of SOC screening modalities, and aggressiveness (or natural history) of the cancer types being evaluated.

Incidence
Cancers with low incidence may be difficult to assess using non-RWD clinical screening studies, and therefore are more likely to have limited evidence to understand the benefit-risk profile for these cancers. RWD enables an analysis of the performance of MCED screening tests on a scale (tens to hundreds of thousands) that is difficult to achieve in a time- and cost-effective manner with traditional clinical screening studies, allowing for evidence generation for rare cancers. Moreover, RWD is also valuable to use in studies that evaluate test performance for cancers that have moderately high incidence and may require large numbers to sufficiently power the analysis.

Existence of Recommended Standard of Care Screening
Cancers with SOC screening recommendations may have more standardized pathways that allow for aligned RWD capture compared to those without such recommendations. For example, cancers with existing United States Preventive Services Task Force (USPSTF) A or B recommendations (breast, cervical, colorectal, and lung cancer) have standard diagnostic pathways which can be captured routinely in RWD. As a result, the follow-up to clinically confirm cancer is more aligned across settings, which can help standardize collection and assessment of RWD for endpoint measurements. In cancers without SOC screening, there may be higher variability in the workup for diagnostic confirmation, creating challenges for the use of RWD to ascertain cancer diagnoses.

Natural History
Variations in the natural history, or aggressiveness, of different cancer types may also affect data capture and evidence generation. Indolent cancers grow slowly and rarely metastasize or contribute to cancer-related death, resulting in better clinical outcomes. Conversely, highly aggressive cancers usually form, grow, or spread quickly, generally resulting in worse morbidity and mortality outcomes.[1] Therefore, the time frame for RWD data captures will be influenced by the natural history and aggressiveness of the cancer (e.g., time to ascertain false negative).

Types of RWD Study Design for MCED Screening Test Assessment

RWD may be incorporated into study designs in a variety of ways, with varying levels of reliance on the RWD in the overall study design, and may include hybrid methods incorporating RWD with traditional study data (**Figure 1**).[27] At one end of the spectrum, traditional RCTs may use RWD elements, such as selected outcomes identified using EHR or claims data. In the middle are trials in clinical practice settings that may be RCTs with pragmatic designs or single arm studies using a RWD external control arm. Lastly, studies may be designed to collect data following a

'usual care' model that is not mandated by study protocol, captured completely through RWD, either with data collection designed prospectively or using existing data infrastructure. One potential strategy to improve data quality, consistency, and completeness is to prospectively design data capture, such as the use of a registry specifically designed for assessment of MCED screening tests. Determining the best study design to support the assessment of a specific MCED screening test will require discussions between the test developer and FDA. Examples of possible use cases for RWD are highlighted in **Table 2**, illustrating the advantages and challenges associated with use of RWD. The possible use cases provided are suggestions and should not be viewed as prescriptive.

The value of RWD depends on the data quality, consistency, and completeness. FDA has previously and generally outlined how to determine whether RWD is fit-for-purpose (not for MCED screening tests).[27] FDA does not endorse a particular RWD source but assesses the relevance and reliability of the source and its elements for appropriate use. If RWE is generated from multiple RWD sources, each RWD source must be evaluated individually as well as in aggregate to determine appropriate use.[11]

Further guidance may be helpful to clarify the appropriate RWD sources, types of data important to capture, and considerations for capture specific to MCED screening tests. For example, evaluation of clinical performance measures in the RWD setting may be subject to selection bias, as patients who receive an MCED test may be systematically different from those who do not, in terms of patient characteristics and disease risk. Further, patients who select MCED testing, and those who receive a positive versus negative test result, may receive different follow-up imaging tests and treatments than those who do not. Accounting for these differences while determining appropriate comparison groups and study designs will be critical. Further, as RWD sources have increased in availability and accessibility, the comparative effectiveness community has generated a host of analytic methods designed to address these challenges to be able to validly draw inferences about the risk and benefit of interventions based on RWD.[28–30]

Table 2: Possible Use Cases of RWD for Generating Evidence for MCED Screening Tests

Use Case	Description	Advantages of RWD	Challenges to Use of RWD
RWD External Control Arm	The control arm is fully comprised of RWD, reflecting the intended use population but without the use of MCED screening tests	• Reduces the need for patients participating in control arms, for which they may stand to gain no potential clinical benefit • Potential to reduce cohort size necessary to demonstrate clinical validity and utility • Potential to establish a platform study with the same concurrent comparator across MCED screening tests • Potential to achieve adequate comparison by using advanced matching methodology and causal modeling	• Historic data may be less suitable than concurrent collection due to variability in cancer incidence, SOC screening adherence, and exposure to risk factors over time • Characteristics of RWD cohort may be quite different than those of study cohort, creating challenges for propensity matching • A nonrandomized control may introduce bias into detection rate comparison
Participant-Consented RWD Collection	Approach that supports patients in exercising their rights to access their own data to contribute to the study (via record requests or APIs)	• Data gathered from SOC of study participants can serve as salvage pathway to adjudicate outcomes for patients who otherwise would be lost to follow-up • For some cancer types it may be sufficient as primary means of adjudicating study outcomes (e.g., through ascertaining cancer diagnoses in EHR or claims data)	• Requires consent and involvement of the patient, making it more suited to prospective than retrospective studies • Logistically and technically complex workflows

Linking of De-identified Data to Study Cohort	Data from aggregated de-identified sources is linked to study cohort to expand available data for analysis	• Opportunity to more deeply profile study population • In some cases can address gaps in data through linking to external sources • May be performed retroactively in some instances	• In most cases, available comprehensive data will only overlap with a small subset of the studied population. Larger overlap may be possible, but at the expense of comprehensiveness (e.g., a participant may be found in the external data, but the dataset lacks the relevant info) • Subject to potential data quality issues such as misclassification bias
Post-Marketing RWD Studies	Study that aggregates RWD for patients who have received a commercially available MCED test	• Ability to assess the impact of various factors that might not reach statistical significance in pivotal study • Potential to support clinical validation for expanded set of cancer types, including rare cancers • Creates opportunity to collect healthcare resource utilization data to support considerations for guideline inclusion and reimbursement • Creates opportunity to amass information on signals currently of unknown clinical significance	Similar limitations to study types outlined above depending on approach used

Conclusions

This white paper helps identify possible endpoints for assessing MCED screening test performance and characterizes opportunities for capturing these endpoints from RWD. RWE generated from the application of MCED screening tests in the real-world intended use population may help supplement data generated in non-RWD clinical screening studies and may be used to inform regulatory decisions. It is critical that the MCED test developer and regulator align on a plan for the types of data and evidence generation necessary to support regulatory decision-making about the MCED screening test, including approval and post-approval studies to update or expand the label or provide additional supportive evidence.

In addition to this working group, there are many ongoing efforts surrounding the development, assessment, and use of MCED screening tests to support robust characterization of the safety and effectiveness of MCED screening tests while facilitating development and continued innovation for these technologies in a timely manner. Discussions with our working group highlighted the need for alignment on terminology used in MCED screening test development, validation, and evaluation, and an effort by BloodPAC is underway to develop a lexicon for the field. Further work is also needed for the design of clinical screening studies evaluating the clinical validity and utility of MCED screening tests. The National Cancer Institute (NCI) is evaluating the landscape of study designs and seeking to potentially launch a multi-arm, multi-stage, pivotal RCT to evaluate multiple MCED screening tests in the years ahead.[31]

While we have suggested possible data elements and endpoints to capture in RWD, work is still needed to federate data into a common model for ease of future analyses. Another area that needs attention is the use of machine learning and artificial intelligence by many of the MCED screening tests to determine cancer status. RWE can enable real-world learning and evaluation of these technologies as they enter clinical practice, helping to achieve the goal of a learning healthcare system. There may be a role for RWE in periodic (post-market) re-evaluation of MCED screening tests utilizing machine learning to assess the real-world performance of initial and future versions of these tests.

We hope this document supports efforts to collect robust, consistent, and relevant data from various studies and helps optimize evidence generation to facilitate development of MCED screening tests and integration into clinical care.

Disclaimer: This paper reflects discussions that occurred among stakeholder groups, including governmental agencies, on various topics. The topics covered in the paper, including recommendations, therefore, are intended to capture key discussion points. The paper should not be interpreted to reflect alignment on the different topics by the participants, and the recommendations provided should not be used in lieu of FDA published guidance or direct conversations with the Agency about a specific development program. This paper should not be construed to represent views or policies of the US Federal Government.

Glossary

Cancer: A disease in which cells grow and proliferate uncontrollably, not including pre-cancerous lesions.

Clinical Screening Studies: Prospectively designed studies in the intended use population using traditional data capture methods (e.g., electronic data capture, case report forms, patient reported outcomes).

Clinical Utility: The likelihood that use of a test will change the management of patients and, by doing so, improve health outcomes, including, for example, safety, morbidity, quality of life, resource utilization, or survival and mortality.[12]

Clinical Validity: The accuracy with which a test identifies, measures, or predicts the presence or absence of a clinical condition in a patient (e.g., the likelihood that someone with a positive test actually has the specified cancer).[12]

Early-Stage Cancer: Specific TNM stage will vary depending on cancer type but is generally a localized cancer amenable to local intervention for curative intent.

Late-Stage Cancer: Specific TNM stage will vary depending on cancer type but is generally a cancer that has metastasized, and is not amenable to localized intervention.

Liquid Biopsy: The detection of biomarkers using only a blood or fluid sample.

Multi-Cancer Early Detection (MCED) Screening Test: Assays using different technologies to detect cancer-associated biomarkers, such as circulating tumor cells, tumor DNA, and other analytes, to screen for cancers in a defined patient population.

Real-World Data (RWD): Data relating to patient health status and/or the delivery of health care routinely collected from a variety of sources, including electronic health records (EHRs), claims and billing data, and product and disease registries.[11]

Real-World Evidence (RWE): The clinical evidence regarding the usage, and potential benefits or risks, of a medical product derived from analysis of RWD.[11]

Tissue of Origin (TOO): The tissue source of the primary cancer (e.g., breast, lung, etc.); also, sometimes referred to as Cancer Signal Origin (CSO).

References

1. Noone AM, Howlader N, Krapcho M et al., ed. SEER Cancer Statistics Review, 1975-2018 - SEER Statistics. National Cancer Institute Accessed November 30, 2021. https://seer.cancer.gov/csr/1975_2018/
2. Hackshaw A, Cohen SS, Reichert H, Kansal AR, Chung KC, Ofman JJ. Estimating the population health impact of a multi-cancer early detection genomic blood test to complement existing screening in the US and UK. Br J Cancer. 2021;125(10). doi:10.1038/s41416-021-01498-4
3. Chudgar NP, Bucciarelli PR, Jeffries EM, et al. Results of the National Lung Cancer Screening Trial: Where Are We Now? Thorac Surg Clin. 2015;25(2):145. doi:10.1016/J.THORSURG.2014.11.002
4. Shapiro S, Venet W, Strax P, Venet L, Roeser R. Ten- to Fourteen-Year Effect of Screening on Breast Cancer Mortality. JNCI J Natl Cancer Inst. 1982;69(2):349-355. doi:10.1093/JNCI/69.2.349
5. Schröder FH, Hugosson J, Roobol MJ, et al. The European Randomized Study of Screening for Prostate Cancer – Prostate Cancer Mortality at 13 Years of Follow-up. Lancet. 2014;384(9959):2027. doi:10.1016/S0140-6736(14)60525-0
6. Chiu HM, Chen SLS, Yen AMF, et al. Effectiveness of fecal immunochemical testing in reducing colorectal cancer mortality from the One Million Taiwanese Screening Program. Cancer. 2015;121(18):3221-3229. doi:10.1002/CNCR.29462
7. Putcha G, Gutierrez A, Skates S. Multicancer Screening: One Size Does Not Fit All. JCO Precis Oncol. 2021;(5):574-576. doi:10.1200/PO.20.00488
8. SUMMARY OF SAFETY AND EFFECTIVENESS DATA (SSED): Stool DNA-Based Colorectal Cancer Screening Test. Accessed February 24, 2022. https://www.accessdata.fda.gov/cdrh_docs/pdf13/P130017b.pdf
9. SUMMARY OF SAFETY AND EFFECTIVENESS DATA (SSED): Human Papillomavirus DNA detection kit. Accessed February 24, 2022. https://www.accessdata.fda.gov/cdrh_docs/pdf10/P100020b.pdf
10. Prorok PC, Andriole GL, Bresalier RS, et al. Design of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. Control Clin Trials. 2000;21(6 Suppl). doi:10.1016/S0197-2456(00)00098-2
11. CBER, CDRH. Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices Guidance for Industry.; 2016. Accessed May 18, 2021. https://www.fda.gov/media/99447/download
12. NCI Dictionary of Genetics Terms - National Cancer Institute. Accessed February 15, 2022. https://www.cancer.gov/publications/dictionaries/genetics-dictionary
13. 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. 2020;31(6):745-759. doi:10.1016/J.ANNONC.2020.02.011/ATTACHMENT/6577D31A-6339-4982-95C7-B370542D40E5/MMC4.PDF
14. Cohen JD, Li L, Wang Y, et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. Science. 2018;359(6378):926-930. doi:10.1126/SCIENCE.AAR3247
15. Toyoda Y, Nakayama T, Kusunoki Y, Iso H, Suzuki T. Sensitivity and specificity of lung cancer screening using chest low-dose computed tomography. Br J Cancer. 2008;98(10):1602-1607. doi:10.1038/SJ.BJC.6604351
16. Hakama M, Auvinen A, Day NE, Miller AB. Sensitivity in cancer screening. J Med Screen. 2007;14(4):174-177. doi:10.1258/096914107782912077
17. Lehman CD, Arao RF, Sprague BL, et al. National Performance Benchmarks for Modern Screening Digital Mammography: Update from the Breast Cancer Surveillance Consortium. Radiology. 2017;283(1):49-58. doi:10.1148/RADIOL.2016161174

18. Ten Haaf K, Van Rosmalen J, De Koning HJ. Lung cancer detectability by test, histology, stage and gender: estimates from the NLST and the PLCO trials. Cancer Epidemiol Biomarkers Prev. 2015;24(1):154. doi:10.1158/1055-9965.EPI-14-0745
19. Hamashima C. Overdiagnosis of gastric cancer by endoscopic screening. World J Gastrointest Endosc. 2017;9(2):55. doi:10.4253/WJGE.V9.I2.55
20. CDRH. Factors to Consider When Making Benefit-Risk Determinations in Medical Device Premarket Approval and De Novo Classifications Guidance for Industry and Food and Drug Administration Staff Preface Public Comment. Accessed January 21, 2022. https://www.fda.gov/media/99769/download
21. Esserman LJ, Thompson IM, Reid B, et al. Addressing overdiagnosis and overtreatment in cancer: a prescription for change. Lancet Oncol. 2014;15(6):e234. doi:10.1016/S1470-2045(13)70598-9
22. What Cancer Screening Statistics Really Tell Us - National Cancer Institute. Accessed March 10, 2022. https://www.cancer.gov/about-cancer/screening/research/what-screening-statistics-mean
23. Weiss NS, Koepsell TD. Re: All-cause mortality in randomized trials of cancer screening. J Natl Cancer Inst. 2002;94(11). doi:10.1093/JNCI/94.11.864
24. Duffy SW, Nagtegaal ID, Wallis M, et al. Correcting for Lead Time and Length Bias in Estimating the Effect of Screen Detection on Cancer Survival. Am J Epidemiol. 2008;168(1):98-104. doi:10.1093/AJE/KWN120
25. CDRH. Guidance for Industry and FDA Staff - Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests. Published online 2007. Accessed February 12, 2022. https://www.fda.gov/media/71147/download
26. Etzioni R, Gulati R, Weiss NS. Multicancer Early Detection: Learning From the Past to Meet the Future. JNCI J Natl Cancer Inst. 2021;(0):168. doi:10.1093/JNCI/DJAB168
27. Concato J. FDA Clinical Investigator Training Course Real-World Evidence. Published online 2021.
28. Hernán MA. Methods of Public Health Research — Strengthening Causal Inference from Observational Data. N Engl J Med. 2021;385(15):1345-1348. doi:10.1056/NEJMP2113319/SUPPL_FILE/NEJMP2113319_DISCLOSURES.PDF
29. Berger ML, Sox H, Willke RJ, et al. Good practices for real-world data studies of treatment and/or comparative effectiveness: Recommendations from the joint ISPOR-ISPE Special Task Force on real-world evidence in health care decision making. Pharmacoepidemiol Drug Saf. 2017;26(9):1033-1039. doi:10.1002/PDS.4297
30. Haneuse S, Daniels M. A General Framework for Considering Selection Bias in EHR-Based Studies: What Data Are Observed and Why? EGEMS (Washington, DC). 2016;4(1):16. doi:10.13063/2327-9214.1203
31. NOT-CA-22-033: Request for Information (RFI): Seeking Input from Multi-Cancer Early Detection Test Developers on Readiness for Participation in an NCI-Sponsored Clinical Utility Randomized Controlled Screening Trial. Accessed January 24, 2022. https://grants.nih.gov/grants/guide/notice-files/NOT-CA-22-033.html

Appendix Table 1: Defining Endpoints for Evaluating MCED Screening Tests

Category	Endpoint	Definition
Clinical Performance	Clinical Sensitivity **Calculated based on detection of cancer and cancer detection+TOO, both stratified by stage*	All-Cancer Sensitivity: The proportion of subjects with clinically confirmed cancer in whom the MCED screening test was positive. Cancer-Specific Sensitivity: The proportion of subjects with a specific clinically confirmed cancer type (e.g., breast cancer) in whom the MCED screening test accurately identified that specific cancer type.
	Clinical Specificity ** Calculated based on detection of cancer and cancer detection+TOO*	All-Cancer Specificity: The proportion of subjects without clinically confirmed cancer of any type in whom the MCED screening test was negative.
	Positive Predictive Value (PPV) ** Calculated based on detection of cancer and cancer detection+TOO, both stratified by stage (if applicable)*	All-Cancer PPV: The proportion of MCED screening test positive subjects who have any clinically confirmed cancer. Cancer-Specific PPV: The proportion of MCED screening test positive subjects with the TOO accurately identified (e.g., breast cancer) who have clinically confirmed cancer of that type.
	Negative Predictive Value (NPV)	The proportion of MCED screening test negative subjects who do not have cancer of any type.
	Cancer Detection Rate	The proportion of cancers detected by the MCED screening test out of the cancers expected in the population monitored over a defined period of time (requiring a control arm or acceptable external reference cohort).

Safety	Device Related Adverse Events (AEs) *Serious vs. Non-Serious Events*	Any untoward medical occurrence (physical or psychological) directly before, during, or directly after the MCED screening test is administered that is directly attributable to the test. Serious events include: events that may have been or were attributed to the use of the device and produce an injury or illness that is life-threatening, results in permanent impairment or damage to the body, or requires medical or surgical intervention to prevent permanent harm to the body.[20]
	Procedure-Related Complications **Stratified by TP vs. FP, Cancer Type, and Serious vs. Non-Serious Events*	Any harm to screened individuals that is not directly attributable to the test itself but relates to any untoward medical occurrence (physical or psychological) after the MCED screening test is administered until definitive diagnosis (i.e., determination of cancer status).
	Adherence to SOC Screening Following Test **Stratified by Test-Positive and Negative*	The proportion of subjects who have all their USPSTF A or B recommended cancer screening tests (e.g., mammogram, colonoscopy, low-dose chest CT, and cervical screening) completed within the recommended period.
	Frequency of Confirmation Diagnostic Tests and Time to Diagnostic Resolution	The time between receiving a positive MCED screening test result and determination of both the presence or absence of cancer, and specific cancer type.
	Number and Type of Follow-Up Procedures Performed **Stratified by TP vs. FP, Cancer Type, and Invasive vs. Non-Invasive Events*	The number and type of medical procedures performed after the MCED screening test is administered that support the definitive diagnosis.
Clinical Outcomes	**Short-Term Outcomes**	
	Stage Shift **Stratified by Cancer Type*	An increase in the proportion of cancers detected in early- versus late-stage disease with and without the MCED test, with a concurrent decrease in the proportion detected in late-stage disease.
	Late-Stage Cancer Incidence **Stratified by Cancer Type*	The number of new cancer cases diagnosed at a late stage per 100,000 people per year.
	Proportion of Cancers Amenable to Definitive Local Intervention **Stratified by Cancer Type*	The proportion of cancers diagnosed in the specified population where definitive, curative, local intervention is clinically feasible.

Clinical Outcomes	Progression-Free Survival **Stratified by Cancer Type*	The length of time from diagnosis of cancer to first clinical evidence of disease progression.
	Long-Term Outcomes	
	All-Cause Mortality	The total number of deaths occurring in the population, regardless of the cause of death, in a specified time period.
	All-Cancer Mortality **Stratified by Cancer Type*	The total number of deaths occurring in the population due to cancer in a specified time period.
	5-Year Cancer Specific Survival **Stratified by Cancer Type*	The probability of surviving cancer in the absence of other causes of death in a 5-year period.
	5-Year Overall Survival **Stratified by Cancer Type*	The percentage of patients alive in the population five years after their cancer diagnosis.