FRIENDS of CANCER RESEARCH

Framework for Integrating Change in ctDNA Levels in Advanced Cancer Clinical Trials to Support Meta-analyses for Intermediate Endpoint Validation

Friends of Cancer Research White Paper | 2024

Executive Summary

In oncology clinical trials, using intermediate endpoints that are reasonably likely to predict clinical benefit accelerates access to therapies. For these endpoints to support regulatory decision-making, the U.S. Food and Drug Administration (FDA) expects meta-analyses at the patient- and trial-level that demonstrate associations between the intermediate endpoint and long-term clinical outcomes, such as overall survival (OS; i.e., clinical benefit). Circulating tumor DNA (ctDNA), a biomarker found in the blood, can serve as an indicator of tumor burden, and has shown promise as an intermediate endpoint. Initial findings from multiple clinical trials, including the Friends of Cancer Research (*Friends)* ctMoniTR Project that aggregates patient-level data from several clinical trials, demonstrate that decreases in ctDNA levels while on treatment associate with improved OS. However, evidence is lacking for trial-level meta-analyses due to inconsistencies in approaches across trials including study design, data collection, and ctDNA measurement methods, making it difficult to combine results.

To address this gap, *Friends* assembled a working group of experts, including representatives from the FDA, pharmaceutical companies, diagnostics developers, patient advocate organizations, and academia, to align on key considerations for prospectively designed clinical trials that collect ctDNA in a standardized manner. The considerations focus on advanced non-small cell lung cancer (aNSCLC) treated with immunotherapy (IO) due to the robust data established to date and ongoing drug development in this space. With a standardized approach, these trials may be more appropriate to combine with regards to data quality and coherence into a trial-level meta-analysis to support the use of ctDNA as an intermediate endpoint in oncology drug development.

The working group prioritized several critical recommendations for alignment in terms of study design and data collection; however, additional considerations are also outlined. The most critical recommendations for alignment of study design and data collection are:

- Collect a baseline ctDNA measurement before treatment initiation, preferably on C1D1 before infusion.
- Collect four on treatment ctDNA measurements: three during subsequent treatment cycles (i.e., C2D1, C3D1, and C4D1) and one at 6-months post-treatment initiation; align with RECIST measurements as is feasible.
- Use an assay that is sensitive enough to detect ctDNA in at least 70% of patients at baseline.
- Report data related to ctDNA analysis and measurement in an aligned approach (specific recommendations are included in Table 1 of the white paper).

These recommendations aim to align ctDNA collection and analysis in future clinical trials, supporting validation efforts for using ctDNA as an intermediate endpoint in regulatory decision-making, and ultimately accelerating the delivery of treatments for serious and life-threatening diseases to patients.

Authors

Charu Aggarwal, Abramson Cancer Center, University of Pennsylvania
Valsamo Anagnostou, Johns Hopkins University
Hillary Andrews, Friends of Cancer Research
Zoe June Assaf, Genentech, A Member of the Roche Group
Biswajit Das, Molecular Characterization Laboratory, Frederick National Laboratory
Yu Deng, Genentech, A Member of the Roche Group
Carin Espenschied, Guardant Health
David Fabrizio, Foundation Medicine, Inc.
Keith Flaherty, Mass General Hospital
Leon Freytor, Johnson & Johnson
Minakshi Guha, Takeda
Glenn J. Hanna, Dana-Farber Cancer Institute
Janet Jin, Amgen
John Lee, Daiichi Sankyo
Anne-Marie Martin, GSK
Dimple Modi, Regeneron
Achim Moesta, Regeneron
Kiran Patel, Johnson and Johnson
Reena Philip, U.S. FDA
Kai Qu, Amgen
Rajiv Raja, GSK
Matthew Silver, Takeda
Mark Stern, BMS
Carol Vallett, *Friends'* Advisory Advocate
Diana Merino Vega, AstraZeneca
Paz Vellanki, U.S. FDA
Jian Wang, AstraZeneca
Kathie Winson, Genentech, A Member of the Roche Group
Alexander W. Wyatt, Vancouver Prostate Centre, University of British Columbia and Clinical Cancer Genomics Program and Michael Smith Genome Sciences Centre, BC Cancer
Nevine Zariffa, NMD Group

Table of Contents

Background

Advancements in oncology drug development have significantly improved outcomes for many patients with solid tumor cancers. Given these successes, it can be lengthy and resource-intensive to conduct studies for newer therapies due to the time required for mature survival endpoint readouts, especially for overall survival (OS), which remains the gold standard for evaluating clinical benefit. There is an opportunity to enhance the availability of more treatment options for patients, and thus, a need for additional, novel intermediate endpoints that are reasonably likely to predict clinical benefit, enabling earlier evaluation of efficacy and regulatory decision-making. The Accelerated Approval Pathway can be leveraged for therapies that treat a serious condition and fill an unmet medical need, allowing for U.S. Food and Drug Administration (FDA) approvals based on an intermediate endpoint that is reasonably likely to predict clinical benefit.

While radiographic-based intermediate endpoints exist, such as objective response rate (ORR) and progression-free survival (PFS), there are some challenges with these approaches. In some settings, an objective radiographic baseline measurement, which is a requirement for these approaches, cannot be made (e.g., patients with large pleural effusions or with bone-only metastases). Some cancer types (e.g., metastatic head and neck cancer) are challenging to measure by radiographic measurements and some therapies (e.g., novel treatments with immune-mediated efficacy) may lead to what appears to be progression on imaging but is in fact pseudo-progression. Additionally, guidance from the FDA[1] and recent discussions at an FDA Oncologic Drugs Advisory Committee meeting[2] suggest there is a need for earlier endpoints in the perioperative setting and highlight the challenges with radiographic based endpoints in early-stage disease as surgery often removes any measurable lesions.

In early- and late-stage settings, an objective and standardized intermediate endpoint that can predict long-term clinical benefit is needed to overcome these limitations of radiographic imaging and support efficacy evaluation in a timely manner. On-treatment change in circulating tumor DNA (ctDNA) levels from baseline can capture response at a molecular level and could potentially be used as an intermediate endpoint. Many sponsors recognize the value of ctDNA and leverage early change in on-treatment ctDNA to predict clinical benefit and inform internal Go/No-Go decisions. A more coordinated effort to have a consistent and unified approach to define molecular response based on ctDNA and for the analysis of such endpoints could support the development of ctDNA endpoints in regulatory decision-making.

To qualify novel intermediate endpoints, FDA guidance recommends meta-analyses of randomized controlled trials at both the individual patient- and trial-levels.[1, 3] The Friends of Cancer Research (*Friends*) ctMoniTR Project combines data from multiple clinical trials in the metastatic setting to assess associations between change in ctDNA levels and OS and PFS at the patient-level. These retrospective patient-level analyses have demonstrated that a decrease in ctDNA is associated with improved PFS and OS. The focus herein will be on patients with advanced solid tumors, as these

reflect the bulk of aggregate data analyses conducted to date. Where there are parallels in late-stage that are relevant for early-stage disease, the same approaches could be considered or adapted, as appropriate.

Scope and Approach to Assessing Change in ctDNA to Date

Friends coordinated a working group with representatives from pharmaceutical companies, diagnostics developers, FDA, academia, and patient advocacy groups to align on recommendations for standardized, harmonized, and robust data collection to include in prospectively designed trials that can support meta-analyses. The primary focus is to set the stage for collaborative evidence collection that assesses change in ctDNA levels and associations with OS, supporting the use of change in ctDNA levels as an intermediate endpoint in regulatory decision-making. (While we recognize that the evidence developed to support using ctDNA as an intermediate endpoint could support approaches for using ctDNA to inform clinical practice, the proposed scope of work is not intended to evaluate the use of ctDNA to guide treatment decisions for individual patients.)

Current approaches to assessing change in ctDNA levels and associations with OS set the stage for which data need to be collected. To date, ctDNA is measured early in clinical trials, with many trials including a baseline blood collection before treatment starts and an on-treatment measurement usually taken 3-12 weeks after treatment initiation.[4, 5] There are a variety of assays to measure ctDNA, including next generation sequencing (NGS) and digital droplet PCR (ddPCR) assays. Often, the variant allele frequency (VAF) for all variants included at each timepoint is determined (e.g., mean or maximum VAF) and used to calculate a percent change in ctDNA from baseline to on-treatment. In some cases, a single variant is tracked (e.g., in oncogene driven cancers), and increasingly, various measures of tumor fraction are used to measure ctDNA.[6–9] Some studies have shown that results are similar regardless of whether multiple genes from a panel test are considered or just the gene of interest.[10]

The following sections provide recommendations for incorporating ctDNA into prospectively designed clinical trials. As a use case, we developed initial recommendations based on observations from advanced non-small cell lung cancer (aNSCLC) treated with immunotherapy (IO), due to the robust data established to date and ongoing drug development in this space. The most critical recommendations for alignment in terms of study design and data collection are prioritized, however, additional considerations are also outlined. Robustly designed trials that have incorporated these recommendations may support meta-analyses for validating the use of change in ctDNA levels as an intermediate endpoint. We also outline initial thoughts for how to approach a meta-analysis to support ctDNA as an intermediate endpoint using these prospectively collected data in the Appendix.

Criteria for a Molecular Response Measurement

A key aspect of each of the following sections is considering which data should be collected and reported for inclusion in a meta-analysis. In addition to outlining these data throughout the following sections, Table 1 provides recommendations for which datapoints should be reported.

Table 1. Suggested data to collect and report when running prospective trials that incorporate ctDNA.

Reportable	Category	Description	Required or Recommended
Plasma volume	Per sample	The volume (mL) of plasma input into assay workflow	Required
cfDNA extracted	Per sample	The amount of total cfDNA (nanograms) extracted from plasma	Required
cfDNA used	Per sample	The amount of cfDNA (nanograms) input into assay workflow	Required
ctDNA detected	Per sample	Binary yes/no of whether ctDNA is detected in the sample	Required
Mutation frequency in ctDNA	Per sample	Amount of ctDNA (VAF) per variant measured	Required*
Measure of ctDNA level	Per sample	Continuous metric (e.g., between 0 and 100%) summarizing the per sample fraction of ctDNA derived from tumor, which could be the mean, max, or median VAF of detected somatic mutations, or alternative estimates of tumor fraction (multi-omic, methylation, etc.)	Required
Timing of baseline measurement	Per patient	Days relative to treatment initiation	Required
Timing of on-treatment measurement	Per sample	Days relative to treatment initiation for each sample	Required
Percent of patients with detected baseline ctDNA	Baseline characteristics	How many patients have baseline ctDNA detected, including pre-specified level of ctDNA	Required
Prior therapy	Per patient	Number of lines of prior therapy, types of therapy, and time since the last therapy (i.e., days from prior line ending to treatment initiation)	Recommended
Concurrent therapy	Per patient	Report any medications/ surgery/ radiation that the patient receives beyond systemic chemotherapy	Recommended
Molecular response	Per patient	Yes/no of whether the patient has a molecular response	Recommended
Molecular response approach	General	Describe how the molecular response was defined including when the response was assessed - consider including the following thresholds: 50% reduction, 90% reduction, 100% reduction (i.e., clearance)	Recommended

*For assays that measure genetic alterations in multiple genes

Assay Characteristics

Measuring ctDNA involves assays that assess various genes or other somatic features from a liquid biopsy (here we focus on plasma, but cerebral spinal fluid, urine, and saliva are other examples). The poolability of molecular data in a meta-analysis will depend on the similarity of assays with respect to performance metrics. As feasible, we recommend that an appropriate set of reference materials is used to demonstrate comparability across multiple assays. When selecting an assay for a clinical trial, sponsors should consider sensitivity and specificity at a particular limit of detection (LoD) and clinical cut-off, as well as other assay performance measures, the number and types of genes or somatic features assessed, and approaches to clonal hematopoiesis (CH) variant removal. There are many assays currently in use that detect and quantify ctDNA and technology continues to evolve.

Table 2 outlines proposed minimum requirements for assays to ensure there is transparency in how the diagnostic is used and below we discuss some key technical considerations.

Minimum Requirements for Assay Analytical Validation

Various factors can influence assay performance including pre-analytical variables (e.g., the volume of plasma collected), the bioinformatics pipeline, and inter-assay variability (e.g., depth and breadth of genomic coverage). When selecting a ctDNA assay, it is critical that the assay follows current recommendations for analytical and clinical validation. BLOODPAC proposed a set of analytical and pre-analytical validation protocols for assessing NGS ctDNA platforms.[11, 12] We recommend diagnostic developers use these or similar protocols to ensure analytical and clinical accuracy and reliability and that clinical trial sponsors report the approach used. We also recommend the cut-off is pre-specified, and the same assay and algorithm be used for the entire trial, including the associated cutoffs.

Considerations Regarding CH Removal

CH variants are somatic mutations that originate from expansions in hematopoietic progenitor cells.[13, 14] ctDNA is measured as a part of total circulating free DNA (cfDNA), which includes CH variants that can pose a challenge when trying to identify tumor related content or quantify ctDNA levels. It is critically important to be accurate when removing CH variants as they may alter interpretation of changes in ctDNA levels. To account for CH-related mutations, diagnostic developers currently employ one of three main approaches for their assays:[15]

1. An algorithmic approach to removing CH variants that is part of the bioinformatic pipeline, which involves removing genetic mutations commonly found in hematopoietic cells and may leverage other information available from the assay. A challenge with this approach is the possibility of removing variants of interest or not appropriately removing the CH variants given that alterations in some genes (i.e., *TP53, ATM*) may be CH or tumor-derived leading to incorrect CH calls.
2. Tumor informed or bespoke approaches consider the variants found in the sequenced tumor tissue to distinguish ctDNA variants in cfDNA. Apart from limited tissue availability, a

challenge with this approach is that it requires tumor tissue to not only be accessible and removed surgically or through a biopsy but also requires waiting for tumor sequencing to select the appropriate ctDNA variants for measuring/tracking, which may or may not be available in real-time. The analysis is limited to variants present in the tumor tissue specimen at baseline, which comes from a single lesion that may not be representative of genetic alterations at other sites.

3. Peripheral blood mononuclear cells (PBMC) removal approaches use PBMCs collected from blood samples to filter out CH or germline variants. Challenges with this approach include the cost of running the samples twice and sensitivity limitations.

We recommend diagnostic companies explicitly state how they identify and/or exclude CH in their assay and specifically report CH-specific false positive rate. Reporting the probability of detection based on sample-level and allele-level coverage is important for all variants, tumor-derived and CH.

Assay Sensitivity

In cancer, it is assumed that patients with a sizeable, proliferative tumor have ctDNA in their bloodstream prior to any therapy, reflecting the burden of disease, However, ctDNA detection is impacted by both biological factors, such as tumor location, vascularization and aggressiveness, as well as technical factors, especially assay sensitivity and plasma collection volume.

For prospective trials assessing aNSCLC treated with IO, assays should be sufficiently sensitive such that most patients in the planned trial will have 'detected' ctDNA at baseline. Approximately 70-85% of patients with aNSCLC have detected baseline ctDNA when using an assay with a LoD ~0.1% VAF (1000 ppm),[16, 17] a range that should be considered when selecting an appropriate assay for use. This approximation of detection is a lower range, as more sensitive assays would result in a greater number of patients with detected ctDNA. We recommend that sponsors report their predetermined ctDNA detection level cutoff and the rate of ctDNA detection at baseline.

ctMoniTR findings in patients with aNSCLC treated with anti-PD(L)1 demonstrate that when using an assay with a LoD of as low as 0.3% VAF (3000 ppm), a 50% or 90% decrease in ctDNA is associated with improved OS. Additional data are emerging and will determine the level of sensitivity for other treatment types and settings, including early-stage disease.

Emerging Technology

To date, much of the work assessing associations between change in ctDNA levels and OS has focused on measuring ctDNA levels by assessing tumor-derived variants (i.e., changes to the genome sequence). There are a variety of emerging approaches for quantifying ctDNA that do not rely only on sequence variants, including assessing changes in cfDNA methylation and cfDNA fragment size distributions as well as physical properties of cfDNA fragments (i.e., the cell free DNA fragmentome). As appropriate, characteristics included herein should be reported for these emerging technologies so that their potential utility relative to currently established approaches can be understood.

Table 2. Minimum reporting requirements for assays to ensure there is transparency in how the diagnostic is used and support considerations regarding poolability in the meta-analysis.

Characteristic	What should be reported	Recommendations
Limit of detection (LoD)	• LoD as reported by the diagnostics company • Report whether this is LoD50 or LoD95 • Reportable range for values below LoD • Approach to defining the LoD	• The LoD of the assay as reported by the assay developer should be 0.3% VAF or lower • Higher sensitivity assays may be worth exploring

CH Removal Approach	• Approach to removing CH variants:
	  ○ Algorithm/machine learning- based removal
	  ○ Tumor informed
	  ○ PBMC-analysis-based removal
	• CH-specific false positive rate at the sample level and variant level
		• PBMC removal or tumor informed approaches are preferred
		• If using algorithmic removal, the assays should clearly report limitations or uncertainties

Assay Characteristics	• Assay version (to account for potential modifications over time) • Number of genes and alterations measured and gene names • Detection threshold (cut-off) and approach to determine detected vs. non-detected ctDNA	• Sponsors may choose to conduct all assays at trial completion to avoid time-drift of assay methodology which could add noise (or worse confounding factors) to the trial specific data set
Performance Data	• Limit of the Blank (LoB) • Precision • Accuracy • Assay sensitivity and specificity • Pre-analytical approach, including which guidelines were followed • Standardized protocols for sample collection, storage, processing, and handling	• Assay sensitivity should ensure that most patients in the planned trial have 'detected' ctDNA at baseline based on historical data • Pre-analytical assessments should follow established guidelines

Timing of Sample Collection for ctDNA Assessment

One of the most critical areas for alignment regarding the ability to combine data from various prospectively designed clinical trials is the timing of blood sample collection for ctDNA analysis. Table 3 prioritizes recommendations for timing of sample collection.

Baseline Sample Collection

It is critical that sponsors collect a baseline ctDNA measurement before treatment initiation. Ideally, this collection should occur on the same day as the first cycle of therapy (i.e., cycle 1 day 1; C1D1) before infusion. However, some flexibility may be warranted as some patients may visit the healthcare system for laboratory work before their first treatment. When considering appropriate flexibility, sponsors should avoid using the ctDNA assessment from the screening assessment as the baseline because there may be differences in these values.[18] The aligned approach from ctMoniTR was to consider samples collected up to 14 days before treatment initiation as the baseline sample.[19]

On-treatment Sample Collection

For the on-treatment sample collections, many studies collect samples 3-12 weeks after treatment initiation.[5] The ctMoniTR project assessed on-treatment ctDNA up to 10 weeks from treatment initiation as the 1st on treatment measurement. The project combined data from multiple collection time points within that time window and saw associations with outcomes, suggesting there could be some flexibility on which specific week the samples are collected early in treatment. Ideally, samples would be collected when the patient is present for other reasons such as labs, scans, or infusions. To continue building evidence to compare and contrast radiographic response, sample collections near scans for radiographic response assessment could be helpful. Since many IO infusions occur on similar schedules (i.e., once every 3 or 4 weeks), we highly recommend sample collections for ctDNA assessment occur prior to drug administration on infusion day. Along with this, we recommend ensuring that at least one on-treatment measurement occurs between 2-10 weeks post treatment initiation.

Frequency of Sample Collection

Whether a “confirmation of response” is necessary for molecular response assessment is a question of interest for which we do not currently have sufficient data. A confirmation is required for radiographic imaging progression in RECIST guidelines[20] and is recommended for biochemical disease progression using prostate specific antigen (PSA) where the 'confirmed' category requires two consecutive measurements to agree on response or non-response. Few studies have assessed plasma during multiple on-treatment timepoints, which makes it challenging to provide recommendations on the dynamics of ctDNA. Clinical trialists are challenged to simplify trials and patients may have clinical progression or toxicity due to treatment, so while collecting samples over continued cycles is ideal, it may not be feasible or practical. To support identifying the most appropriate timing for ctDNA collection, we highly recommend at least 3 subsequent on-treatment

samples are collected (i.e., C2D1, C3D1, and C4D1). This would ideally be on or around the same time as radiological assessment, as is feasible.

Durability of Response

While demonstrating early associations of change in ctDNA levels with outcomes would be the primary goal of a meta-analysis (Appendix), understanding the durability of the molecular response, or how long a decrease in ctDNA or clearance of ctDNA lasts, is also important. It is likely unfeasible for clinical trialists to collect samples for ctDNA assessment every cycle for the entire trial, rather, a single aligned timepoint later in the trial may be more beneficial, for example, using 6-, 9-, or 12-months post-treatment initiation, similar to what was considered when establishing MRD as an intermediate endpoint in multiple myeloma.[21] We recommend sponsors prioritize including a 6-month post-baseline sample collection, as some literature has shown durability of response and this measurement aligns with when PFS6 is assessed, an endpoint often used in studies focused on aNSCLC treated with IO.[22] If feasible, consider also including a 1-year assessment[23] and a measurement at the time of progression.

Table 3. Prioritized recommendations for timing of sample collection for ctDNA assessment.

ctDNA Sample	Definition	High priority	Lower priority
Baseline	Sample measurements before treatment initiation	Collect baseline ctDNA on C1D1 before infusion	Measurements up to 14 days before C1D1 can be considered
Molecular Response Assessment	Samples collected after treatment initiation but before progression	Collect samples at the same time as infusions for the subsequent 3 cycles (i.e., C2D1, C3D1, and C4D1)	Continue collecting samples every infusion through progression
Durability	Samples collected after a period of time to assess durability of response	Collect a 6-month post-treatment initiation sample	Collect a sample at progression and at 1-year post treatment initiation

Patient Inclusion Considerations

When performing the meta-analysis, there may be baseline characteristics (e.g., specific clinical prognostic factors) that should be included in the analysis to assess their impact on the predictive nature of ctDNA. Sometimes, patients with non-measurable disease at baseline by radiographic assessment using RECIST guidelines are excluded from clinical trials. A parallel scenario is non-detected ctDNA at baseline, which may be due to either the limited sensitivity of the assay or true non-detectable ctDNA in the plasma sample. Either way, various reports demonstrate that non-detected ctDNA at baseline is a prognostic biomarker.[8, 17, 24, 25] In the ctMoniTR Project, most patients with non-detected ctDNA at baseline also have non-detected ctDNA on treatment, which makes it challenging to know whether the resulting associations with outcomes are related to the patient's response to treatment. The serial non-detection can also be due to assay limitations (i.e., the

patient's tumor does not have a mutation in the gene panel), the assay's LoD may not support the detection of the mutation, or the volume of plasma was too low (i.e. by chance, insufficient tumor DNA fragments were in the small sample).

Many prospective clinical trials currently under development will consider ctDNA as an exploratory endpoint and inclusion/ exclusion criteria will be tailored to evaluating the primary endpoint (e.g., measurable disease by RECIST assessment). Excluding patients with non-detected ctDNA at baseline could lead to bias and there may be patients who go on to have detected ctDNA on treatment. Additionally, the time it takes for ctDNA results to return may be too long for many patients to wait to start a trial if detected baseline ctDNA were an inclusion criterion. As such, we recommend including patients with non-detected ctDNA in the clinical trial. The meta-analysis plan should include an approach to how these patients' data will be considered (e.g., as a stratification factor).

Prior anti-cancer therapies may impact baseline ctDNA values. When describing patients' baseline measures, it is important to report the history of prior therapy and the time since previous line of therapy. A minimum washout period before ctDNA analysis is unclear but should be accounted for and has the potential to be analyzed in meta-analyses.

Calculating Molecular Response

Approaches to defining molecular response are evolving and will be finalized prior to undertaking the formal validation meta-analysis. As such, several characteristics might be required for defining response and sponsors will be asked to ensure all necessary data are prospectively collected.

While it is agreed that change in ctDNA from a baseline to on-treatment should define a molecular response, the approach to calculating change is yet to be determined. Most commonly, a change is calculated as a percent change, which in the context of an aggregate analysis, accounts for assay differences making it more poolable. Currently, a greater than 50% or 90% decrease in ctDNA levels or ctDNA clearance is used to determine a molecular responder.[5, 26, 27] One concern with this approach is that patients with small VAFs at baseline (i.e., <1.0%) may have large percent changes as VAF values become exponentially smaller, but these changes may not translate to biologically or clinically relevant differences. Another potential concern is with the reliability of the numerical results at low VAF. As such, a proposed method for calculating molecular response is to use absolute change–though it is unclear how to apply this strategy in the context of an aggregate analysis without comparability across assays. A third proposed method is to consider clearance of ctDNA (i.e., ctDNA that becomes non-detected on treatment). Again, assay differences, including variations in sensitivity, may influence results and contribute to differences in detection levels and this may overly limit the population of patients who qualify as molecular responders.

We recommend that sponsors use a percent change for studies that assess aNSCLC treated with IO given past work, but if a considerable number of patients have ctDNA clearance on treatment, a clearance cutoff could be considered in the context of a clearly documented assay LoD. The meta-

analysis should consider >50% decrease, >90% decrease, and clearance as three approaches to calculate a molecular responder. The primary endpoint for validation will be defined prior to conducting the meta-analysis. For each patient, it is important to record the precise volume of input plasma used for cfDNA extraction, the total cfDNA extracted from the plasma, the tumor fraction estimate, a measure of the error range/ confidence interval around the estimate, and the amount of cfDNA input into DNA sequencing library preparation. These values can support recalculation of ctDNA output and ctDNA change metrics as needed to ensure consistency.

Comparing Treatment Groups Using Molecular Response

While there is robust evidence demonstrating associations between change in ctDNA levels and outcomes, few studies focused on comparing two trial arms to determine superiority using ctDNA data. While this is something that can be further explored in meta-analyses, it is important to consider what data should be collected in prospective analyses to ensure meaningful results.

There are two main approaches for determining whether one group has a better molecular response over another: 1) defining a cutoff (e.g., >50% decrease) as a "molecular responder" then calculating a molecular response rate (i.e., percentage of patients who are molecular responders) and/or 2) determining the depth of response and comparing whether one arm has a "deeper" response compared to another (i.e., greater reduction in ctDNA levels). Duration of response is another important element for evaluating molecular response, which measures how long ctDNA levels remain reduced. While FDA does not consider durable clinical benefit a meaningful endpoint on its own, it remains a key component of RECIST-based evaluations. Therefore, we recommend the definition of molecular response focus on change in ctDNA at earlier timepoints; however, understanding the length of response is important and, as described above, collecting later blood measurements (i.e., 6 months) as well as at progression is recommended. Considered as a package, this enables both the primary goal of the validation of an early endpoint based molecular response and the additional secondary goal of assessment of the evolution of ctDNA mid- to long-term post initiation of treatment.

Conclusions

We have outlined a variety of considerations for data collection in prospective clinical trials assessing aNSCLC treated with IO and incorporating ctDNA. The most critical recommendations for alignment of study design and data collection are:

- Collect a baseline ctDNA measurement before treatment initiation, preferably on C1D1 before infusion.
- Collect four on treatment measurements: three during subsequent treatment cycles (i.e., C2D1, C3D1, and C4D1) and one at 6-months post-treatment initiation; align with RECIST measurements as is feasible.
- Use an assay that is sensitive enough to detect ctDNA in at least 70% of patients at baseline.

- Report all relevant information as outlined in Table 1.

As sponsors plan and execute clinical trials assessing aNSCLC treated with IO, there is an opportunity to prospectively incorporate ctDNA with an aligned approach. We focused on aNSCLC treated with IO, however, principles outlined here could be considered for other advanced cancer types and treatment modalities. Trials focused on aNSCLC treated with IO, if using an aligned approach, can support the development and implementation of a meta-analysis plan that can assess how change in ctDNA levels associate with OS in an aggregate trial-level manner. These analyses can lay the groundwork for using ctDNA as an intermediate endpoint to ensure more rapid availability of safe and effective drugs to patients with cancer.

Citations

1. Use of Circulating Tumor DNA for Early Stage Solid Tumor Drug Development Guidance for Industry. Available from: https://www.fda.gov/media/158072/download
2. Meeting of the Oncologic Drugs Advisory Committee - 07/25/2024 | FDA. Available from: https://www.fda.gov/advisory-committees/updated-agenda-information-july-25-2024-meeting-oncologic-drugs-advisory-committee-meeting
3. Vellanki PJ, Ghosh S, Pathak A, et al: Regulatory Implications of ctDNA in Immuno-oncology for Solid Tumors. J Immunother Cancer 11:5344, 2023. Available from: https://pubmed.ncbi.nlm.nih.gov/36796877/
4. Stewart MD, Anagnostou V: Liquid Biopsies Coming of Age: Biology, Emerging Technologies, and Clinical Translation- An Introduction to the JITC Expert Opinion Special Review Series on Liquid Biopsies. J Immunother Cancer 11, 2023. Available from: https://pubmed.ncbi.nlm.nih.gov/36657817/
5. Thompson JC, Scholes DG, Carpenter EL, et al: Molecular Response Assessment Using Circulating Tumor DNA (ctDNA) in Advanced Solid Tumors. Br J Cancer 129:1893–1902, 2023. Available from: https://pubmed.ncbi.nlm.nih.gov/37789101/
6. Stutheit-Zhao EY, Sanz-Garcia E, Liu Z, et al: Early Changes in Tumor-Naive Cell-Free Methylomes and Fragmentomes Predict Outcomes in Pembrolizumab-Treated Solid Tumors. Cancer Discov 14:1048–1063, 2024. Available from: https://pubmed.ncbi.nlm.nih.gov/38393391/
7. Adalsteinsson VA, Ha G, Freeman SS, et al: Scalable Whole-Exome Sequencing of Cell-Free DNA Reveals High Concordance with Metastatic Tumors. Nat Commun 8, 2017. Available from: https://pubmed.ncbi.nlm.nih.gov/29109393/
8. Fonseca NM, Maurice-Dror C, Herberts C, et al: Prediction of Plasma ctDNA Fraction and Prognostic Implications of Liquid Biopsy in Advanced Prostate Cancer. Nat Commun 15, 2024. Available from: https://pubmed.ncbi.nlm.nih.gov/38418825/
9. Tolmeijer SH, Boerrigter E, Sumiyoshi T, et al: Early On-treatment Changes in Circulating Tumor DNA Fraction and Response to Enzalutamide or Abiraterone in Metastatic Castration-Resistant Prostate Cancer. Clin Cancer Res 29:2835–2844, 2023. Available from: https://pubmed.ncbi.nlm.nih.gov/36996325/
10. Soo RA, Martini JF, van der Wekken AJ, et al: Early Circulating Tumor DNA Dynamics and Efficacy of Lorlatinib in Patients with Treatment-Naive, Advanced, ALK-Positive NSCLC. J Thorac Oncol 18:1568–1580, 2023. Available from: https://pubmed.ncbi.nlm.nih.gov/37295609/
11. Lockwood CM, Merker JD, Bain E, et al: Towards Preanalytical Best Practices for Liquid Biopsy Studies: A BLOODPAC Landscape Analysis. Clin Pharmacol Ther, 2024. Available from: https://pubmed.ncbi.nlm.nih.gov/39164947/
12. Godsey JH, Silvestro A, Barrett JC, et al: Generic Protocols for the Analytical Validation of Next-Generation Sequencing-Based ctDNA Assays: A Joint Consensus Recommendation of the BloodPAC's Analytical Variables Working Group. Clin Chem 66:1156–1166, 2020. Available from: https://pubmed.ncbi.nlm.nih.gov/32870995/
13. Abbosh C, Swanton C, Birkbak NJ: Clonal Haematopoiesis: A Source of Biological Noise in Cell-Free DNA Analyses. Ann Oncol 30:358–359, 2019. Available from: https://pubmed.ncbi.nlm.nih.gov/30649226/

14. Liu J, Chen X, Wang J, et al: Biological Background of the Genomic Variations of cf-DNA in Healthy Individuals. Ann Oncol 30:464–470, 2019. Available from: https://pubmed.ncbi.nlm.nih.gov/30475948/

15. Chan HT, Chin YM, Nakamura Y, et al: Clonal Hematopoiesis in Liquid Biopsy: From Biological Noise to Valuable Clinical Implications. Cancers (Basel) 12:1–18, 2020. Available from: https://pubmed.ncbi.nlm.nih.gov/32823942/

16. Assaf ZJF, Zou W, Fine AD, et al: A Longitudinal Circulating Tumor DNA-based Model Associated with Survival in Metastatic Non-Small-Cell Lung Cancer. Nature Medicine 29:859–868, 2023. Available from: https://www.nature.com/articles/s41591-023-02226-6

17. Gray JE, Ahn MJ, Oxnard GR, et al: Early Clearance of Plasma Epidermal Growth Factor Receptor Mutations as a Predictor of Outcome on Osimertinib in Advanced Non-Small Cell Lung Cancer; Exploratory Analysis from AURA3 and FLAURA. Clin Cancer Res 29:3340–3351, 2023. Available from: https://pubmed.ncbi.nlm.nih.gov/37379430/

18. Vega DM, Markovets A, Abbosh C, et al: Abstract 1052: Background Variability in Plasma ctDNA Levels in Patients with Advanced Lung Cancer in the Absence of Treatment. Cancer Res 83:1052–1052, 2023. Available from: https://aacrjournals.org/cancerres/article/83/7_Supplement/1052/721446/Abstract-1052-Background-variability-in-plasma

19. Vega DM, Nishimura KK, Zariffa N, et al: Changes in Circulating Tumor DNA Reflect Clinical Benefit Across Multiple Studies of Patients with Non-Small-Cell Lung Cancer Treated with Immune Checkpoint Inhibitors. JCO Precis Oncol 6:e2100372, 2022 Available from: http://www.ncbi.nlm.nih.gov/pubmed/35952319

20. Eisenhauer EA, Therasse P, Bogaerts J, et al: New Response Evaluation Criteria in Solid Tumours: Revised RECIST Guideline (Version 1.1). Eur J Cancer 45:228–247, 2009. Available from: https://pubmed.ncbi.nlm.nih.gov/19097774/

21. Ola Landgren C, and Devlin S: Evaluating Minimal Residual Disease as an Intermediate Clinical Endpoint for Multiple Myeloma: The EVIDENCE Meta-Analysis Presentation at the April 12, 2024, Meeting of the Oncologic Drugs Advisory Committee Meeting. Available from: https://www.fda.gov/media/177737/download

22. Kok PS, Cho D, Yoon WH, et al: Validation of Progression-Free Survival Rate at 6 Months and Objective Response for Estimating Overall Survival in Immune Checkpoint Inhibitor Trials: A Systematic Review and Meta-analysis. JAMA Netw Open 3:e2011809, 2020. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC7489825/

23. Gandhi L, Rodríguez-Abreu D, Gadgeel S, et al: Pembrolizumab plus Chemotherapy in Metastatic Non-Small-Cell Lung Cancer. N Engl J Med 378:2078–2092, 2018. Available from: https://pubmed.ncbi.nlm.nih.gov/29658856/

24. Reichert ZR, Morgan TM, Li G, et al: Prognostic Value of Plasma Circulating Tumor DNA Fraction Across Four Common Cancer Types: A Real-World Outcomes Study. Ann Oncol 34:111–120, 2023. Available from: https://pubmed.ncbi.nlm.nih.gov/36208697/

25. Zhang Q, Luo J, Wu S, et al: Prognostic and Predictive Impact of Circulating Tumor DNA in Patients with Advanced Cancers Treated with Immune Checkpoint Blockade. Cancer Discov 10:1842–1853, 2020. Available from: https://pubmed.ncbi.nlm.nih.gov/32816849/

26. Sivapalan L, Murray JC, Canzoniero JV, et al: Liquid Biopsy Approaches to Capture Tumor Evolution and Clinical Outcomes During Cancer Immunotherapy. J Immunother Cancer 11, 2023. Available from: https://pubmed.ncbi.nlm.nih.gov/36657818/

27. Anagnostou V, Ho C, Nicholas G, et al: ctDNA Response After Pembrolizumab in Non-Small Cell Lung Cancer: Phase 2 Adaptive Trial Results. Nature Medicine 29:2559-2569, 2023. Available from: https://www.nature.com/articles/s41591-023-02598-9

Appendix: Proposed Meta-Analysis Strategy for Validation of ctDNA as an Intermediate Endpoint in aNSCLC Treated with Immunotherapy and Summary of Additional Substudies to Consider

Background

As we develop recommendations for data collection in prospectively designed trials, it is critical that the approach to conduct the meta-analysis be considered concurrently. Although a fully detailed meta-analysis protocol is outside of the scope of this white paper, we felt it was important to provide considerations. Finalizing the analysis plan will require extensive feedback from different statisticians, including statisticians from FDA's Center for Drug Evaluation and Research and from regulators to ensure that the results would be sufficient for the validation of ctDNA as an intermediate endpoint. Herein, we provide considerations for a meta-analysis that could be used to aggregate randomized controlled trials that assess immunotherapy treatment in aNSCLC based on ctDNA incorporated to future trials following the recommendations provided in the white paper.

This concept analysis plan provides considerations for potential statistical methods for trial- and patient-level metanalyses to validate change in ctDNA levels as an intermediate endpoint. The proposal serves to validate the trial design considerations discussed in the white paper, as well as identify the key considerations to establish an analysis plan. As a guiding principle for the primary goal, we consider the adoption of the simplest binary scenario for molecular response (MR). The primary analysis in this concept analysis plan would use a percent change cutoff at a single early timepoint for assessing associations with overall survival. The collaboration participants and regulators will agree on both the cutoff and the timepoint prior to conducting the analysis. Secondary objectives will include other cutoffs and timepoints. An initial list of sensitivity analyses is listed and will be prioritized as part of the final analysis plan. Additional substudies are also provided in summary form. These may be promoted to secondary analyses as part of the finalization process.

Trials to consider for the meta-analysis must be randomized controlled trials that meet assay and clinical specifications described throughout the white paper. Studies should be included whether or not they show a treatment effect on overall survival. Trial selection should be transparent and unbiased (i.e., based on trial quality, relevance, and consistency rather than outcome driven).

Criteria to Establish for Study Inclusion in the Analysis

- Minimum number of patients per arm
- Minimum number of patients per arm have a MR
- Minimum number of pairwise comparisons to support the study level analysis
- If survival follow-up is ongoing at the time of data cut-off, determine a minimum degree of maturity

Data Collection

ctDNA Timing

- Baseline measurement (relative to treatment initiation, days)
- Each on-treatment measurement (relative to treatment initiation, days)

Patient Characteristics

- Age
- Sex
- Race
- Smoking status
- Stage (advanced stave IV vs else)
- ECOG Performance Status
- Number of prior lines of therapy
- Histology
- PD-L1 expression
- Others to be pre-defined

Clinical Characteristics

- Overall survival
- Progression free survival
- Confirmed (Yes/No)
- BICR used (Yes/No)
- Radiographic measurements throughout the study (i.e., RECIST categories, sum of diameter calculations, timing for RECIST measurements relative to treatment initiation in days)

Assay Characteristics

- Limit of detection (LoD)
- Percent of patients with detected baseline ctDNA
- Clonal hematopoiesis (CH) removal approach
- Sample volume
    - Serum at blood draw
    - Input volume for ctDNA assay
- Performance Parameters
    - Limit of the blank (if applicable)
    - Precision
    - Accuracy
    - Sensitivity/ Specificity
    - Pre-analytical approach including guidelines followed

Descriptive Analyses

Various tabular and graphical summaries to describe:

- Study design features: sample size, arms under study, patient characteristics, median duration of follow up
- Assay description
    - Limit of detection
    - Percent of patients with detected baseline ctDNA
    - Sample volume
- ctDNA data completeness
- ctDNA distribution at baseline and primary timepoint
- OS summary statistics
- PFS and RECIST summary statistics

Primary Endpoints for Evaluation of ctDNA as an Intermediate Endpoint

OS is the clinical outcome of interest analyzed as a time to event variable. KM estimate at 2 or 3 years will be a secondary or sensitivity analysis. Molecular response will be defined prior to the conduct of the analysis as described above and is denoted as MR below.

Primary Analysis for Individual Patient Level Assessment

Degree of separation between MR and nMR relative to OS in a Cox Proportional hazard model accounting for all relevant patient-level covariates, MR (y/n), treatment, treatment by MR interactions (depending on the treatments included in the studies) and a stratification term for study. The primary endpoint will be the HR for MR relative to nMR, the confidence interval and the p-value serve to assess the strength of evidence. A specific threshold should be developed ahead of time (e.g., HR at least 0.7 or better and 95% CI excludes 1).

Secondary Analyses for Individual Patient Level Assessment

Secondary analyses will explore the relationship between various cutpoints and time points of molecular response and overall survival, while sensitivity analyses will assess the robustness of these findings by evaluating different patient subgroups, assay types, and MR thresholds.

Primary Analysis for Trial Level Assessment

Weighted linear regression model of log HR OS (test treatment vs. control in each study) vs. Log odds ratio of MR to nMR

HR OS based on proportional Cox hazard with adjustment for covariates as described above.

The regression will be weighted by the inverse variances of the log odds ratio for log OR MR.

The linear regression may include additional terms for covariates such as study and paired types of treatment-control pairs.

Study-level association metrics of R^2 and associated confidence interval will be calculated. Criteria will be pre-specified (e.g., R^2 at least 0.7 and the lower end of the confidence interval is above 0.5).

Secondary Analyses for Trial Level Assessment

Secondary analyses will explore the relationship between various definitions and time points of molecular response and overall survival, while sensitivity analyses will assess the robustness of these findings by additional statistical model (e.g., weighting by study size) and as above, evaluating different patient subgroups, assay types, and variant allele frequency thresholds.

Substudies

Additional substudies may be valuable as part of a supportive package for the main analysis to validate ctDNA as an intermediate endpoint. An initial list is provided below:

Aspect Evaluated	Description
Baseline ctDNA as a prognostic factor	Include assessment of stratification in tertiles or quintiles of outcomes by baseline values of ctDNA.
Time course of ctDNA: depth and duration of response	Enables evaluation of response over time, which may be key to future development of more refined tools/definition in aNSCLC and in other settings. If sufficient data are provided, we can further evaluate any 'lead time' in ctDNA to identify when molecular progression occurs.
Reproduce the validation for PFS to OS assessment and/or an earlier RECIST based assessment of response	Serves as context setting for the results achieved with ctDNA.
Value of ctDNA beyond other intermediate endpoints such as PFS or ORR	Addresses how ctDNA can be positioned alongside other intermediate endpoints (e.g., Is molecular response more or less predictive than PFS or other definitions of RECIST relative to OS? Does MR have a role in further stratifying patients with stable disease?).