Cancer Steering Committee (CSC)
Annual Scientific Symposium
November 14–15, 2023

Day 1: November 14, 2023, 11:00 am – 7:15 pm ET

Opening Session

FNIH Introduction
Dana Connors, Foundation for the National Institutes of Health (FNIH)

Mr. Dana Connors, Director of the Biomarkers Consortium (BC) Cancer Steering Committee (CSC), introduced the Foundation for the National Institutes of Health (FNIH) as an environment to create breakthroughs for patients. Congress founded and launched the non-profit organization in 1996 to support the mission of the National Institutes of Health (NIH). FNIH works closely with the public sector, biopharmaceutical organizations, and foundations to create public-private partnerships that tackle pressing global health challenges through the acceleration of prevention, new therapies, diagnostics, potential cures, and global equity in care. The FNIH BC portfolio bridges work in four therapeutic areas, including inflammation and immunity, metabolic disorders, neuroscience, and cancer.

BC CSC Introduction
Althea Lang, PhD, Foundation for the National Institutes of Health (FNIH)

The BC CSC comprises a diverse range of representatives from NCI, FDA, large and small pharmaceutical companies, biotech companies, and research foundations. The CSC has three co-chairs who represent biotech, pharma, and research.

Several projects in the CSC project development pipeline are close to launching. The committee recently completed a strategic refresh to identify areas of focus for upcoming projects. Liquid biopsy, immune cell engagers, digital pathology, and building on existing projects emerged as the top areas of interest. The current symposium is designed to inform projects in liquid biopsy and digital pathology by bringing together expertise during two dedicated working sessions.

Dr. Lang reminded the group that BC CSC projects are meant to validate and qualify biomarkers and other drug development tools to accelerate new therapeutics. The committee is focused on translational research that helps biomarkers reach the clinic, as well as standardization and community consensus efforts. All BC projects should have a large unmet need, be pre-competitive, have private sector buy-in, include public sector involvement, hold value for all stakeholders, involve the regulatory sector, and define a context of use. A few examples of CSC project

deliverables include drug advancement, guidance documents, and progress through the biomarker qualification process. The committee does not support basic research, competitive activity, or biomarker identification.

The digital pathology and liquid biopsy projects are in the pre-concept development stage. Before moving any projects into project plan development, they will be evaluated based on their suitability; scientific, clinical, and regulatory importance; feasibility; and commercial implications.

Welcome and Overview
Gary Kelloff, MD, National Cancer Institute (NCI)

Dr. Kelloff, CSC Co-Chair, has participated in cancer research for over 40 years at the NCI and has co-chaired on-going efforts under the FNIH BC since 2009 to create public-private partnerships to define biomarker use in cancer drug development and patient management. He introduced the main topics of the symposium.

Session 1: Independent vs. Synergistic Activity of Drugs in Combination

Moderator: Emmett Schmidt, MD, PhD, Merck

Dr. Emmett Schmidt, Vice-President of Clinical Research at Merck and CSC Co-Chair, introduced the terms synergy and independent action in drug development. He described his experience at Merck of partnering with other large and small pharmaceutical companies to test drug combinations. He explained that they tested two ideas: first, that two active drugs together should be more active, and second—referring to the cancer-immunity cycle—that similarly behaving drugs should have synergy. Three years into the testing, only the collaborations testing two active drugs together had progressed into phase 3 trials.

Dr. Schmidt explained that he received a paper written by Dr. Adam Palmer and Dr. Peter Sorger stating that combination therapy can confer benefit via patient-to-patient variability without drug additivity or synergy. The paper matched what he was seeing with his collaborations. A subsequent paper by Dr. Sorger introduced the concept of bet-hedging, in which drugs in combination therapies work by independent action—with patients responding to only one of the two drugs—but both drugs are given to maximize the chance of treating with the correct drug.

Keynote: Drug Independence: Implications for Drug Discovery and Precision Medicine in Oncology
Peter Sorger, PhD, Harvard Medical School

Dr. Sorger introduced the molecular rationale for treating cancer with combination therapies—the idea that a "cold" tumor can be made "hot" with the right combination of drugs. He compared the concept of drug synergy to the idea of synthetic lethality in genetic screening. Independent action, in contrast, is the null model for testing the idea of drug synergy in the clinical setting.

The idea of independent action, also called "Frie independence," is that there are two distinct populations of patients, and each individual patient responds to one of two therapies in a

combination treatment. In a modeling study based on empirical data with ipilimumab and nivolumab, the actual degree of benefit of the drug combination overlaps with the predicted degree of benefit from independent action. This model of independent action means there are multiple chances of achieving above-average responses in a heterogenous population, and giving the combination treatment is a form of bet-hedging. This can provide a very substantial benefit, and to a large extent matches what has been observed in approved clinical trials. In addition, independent action can be tested in patient-derived xenograft (PDX) mouse models, which have confirmed how independent action arises.

Dr. Sorger explained that their idea of independent action was not novel, and the concept had already been put forward in 1961 after the Acute Leukemia Group B conducted a study with sequential and combination treatments. There was no evidence that the two drugs were acting simultaneously, and instead the patients had two chances of receiving a benefit. Work by Dr. Adam Palmer has revealed that this independent action model fits the data for about 80 percent of all combination therapies available. There is, however, a small subset of combinations, particularly with inactive drugs, that are likely true drug interactions at the individual patient level.

No immune checkpoint inhibition trial has shown any evidence of benefit beyond what would be expected by independent action. This is not true, however, across all drug classes. For example, independent action does not explain curative combinations for leukemia and lymphoma, where cures are a consequence of having additive activity.

Recent modeling work by Dr. Palmer looking at all published phase 3 trials with monotherapy data between 2014 and 2018 suggests that:

- It is possible to precisely predict the likely benefit from combining drugs with each other or with other drug classes based on monotherapy data
- Targeting more than one immune checkpoint protein has no benefit for the single patient
- Individual agents must have activity as monotherapies to be effective

These same observations can be seen in pre-clinical cell lines, some of which had been previously reported incorrectly as examples of drug synergy.

Dr. Sorger stated that approved combination therapies are rarely synergistic but instead work by overcoming variability in individual patient response, and the reason bet hedging provides a benefit is due to this inter-individual variation. In biomarker enriched populations, the median response improves, but the variation in response remains the same, which indicates additional unknown features. Clinical trials, however, evaluate the average benefit of combination treatment. Dr. Sorger imagines an approach with individualized measurement of benefit to inform treatment strategy, and he stated that better digital pathology could enable a precision medicine approach to impact patient care.

NCI's Human Tumor Atlas Network (HTAN) program is building a public-facing atlas of spatial data. Dr. Sorger noted a lack of connection between the HTAN community technical accomplishments and digital pathology. Now, a new tool from Rarecyte, called Orion, allows twenty-plex high resolution imaging at the same time as clinical-grade H&E. The resulting overlay provides a clinical opportunity for image-based prognostic and predictive biomarkers, and Dr. Sorger explained that spatial biomarkers are often more effective than genotyping.

In summary, Dr. Sorger reiterated that combination therapies can overcome heterogeneity with no synergy, and that technologies exist today to address variability at the patient and cellular levels.

Discussion

During a brief discussion about synergy and independent action nomenclature, Dr. Sorger emphasized the importance of population-level versus individual-level distinction when evaluating the benefit of drug combinations. He drew a distinction between the descriptions "additive or synergistic" and "better" when describing the benefit of combination treatments at the individual level, explaining that better does not necessarily mean additive or synergistic but instead could be independent action.

Debate: Independent Action versus Synergy

Adam Palmer, PhD, UNC (Pro Independent Action) vs. Michael Curran, PhD, MD Anderson (Pro Synergy)

Dr. Palmer introduced the position that drug synergy is not necessary to make an effective combination therapy, and the common belief that a synergistic interaction is required for efficacy is not supported by evidence, misdirects research, and leads to predictably failed clinical trials. He stated that highly effective drug combinations arise from combining individually effective drugs.

Next, Dr. Palmer defined synergy. Experimentalists define synergy as drug-drug interactions that have more than an additive effect—an outcome greater than the sum of the parts. Clinicians, however, use the term to indicate a combination that has worked well, and this overlap of meanings has led people to believe that a synergistic interaction, as defined by experimentalists, is required for a positive clinical trial result for a drug combination. Dr. Palmer explained that the independent action of highly effective drugs in combination could work better than a different combination of weaker drugs with true synergy.

His analysis across 25 years of FDA-approved combinations revealed that 95% of them were quantifiably additive or less in terms of their progression-free survival distributions, and the most successful combinations were not from synergy but from highly effective individual drugs. In contrast, synergy is most prevalent among weak drugs at sub-inhibitory concentration ranges and is sensitive to genotypes. Therefore, Dr. Palmer argued, synergy is unlikely to manifest reliably in diverse human populations. He expects that combinations of novel agents will drive major survival gains that will be additive in their effect.

Dr. Curran began his discussion of drug synergy by noting that a better understanding of how drugs work and which patients are more likely to benefit from certain combinations can create an opportunity for real therapeutic drug synergy. For example, he explained that immunotherapy began with a very small amount of basic understanding that has now grown to a deeper knowledge about how the drugs work. Dr. Curran highlighted two points:

1. There is now more knowledge about combined mechanisms of action.
2. Drug resistance mechanisms are relevant, and there is a benefit to understanding the mechanisms of resistance to inform combination treatment strategies, regardless of the strength of independent activity.

Dr. Curran explained that the prevention of resistance reaches beyond immuno-oncology and can involve targeted therapies—such as targeted therapies with antibody-drug conjugates—to ward off the major mechanism of escape from the targeted therapy. He agreed that the true synergy he has described is a small subset of cases, but he highlighted that newer drugs are revealing synergies. For example, LAG-3 has almost no individual activity, and in combination with PD-1, the benefit is beyond additive. In addition, there is a potential synergistic rationale for combining checkpoint inhibitors with cancer vaccines, which create the substrate for checkpoint blockade. The cancer vaccines alone do not work well because the T cells are inactivated until checkpoint blockade allows them to function. Dr. Curran cited two trials, one in melanoma and one in advanced HCC, that showed benefit from combination therapies with vaccines that would alone have very little effect.

Discussion

Dr. Schmidt emphasized the importance of precision medicine. He compared bet-hedging to giving multiple antibiotics to a patient with an unknown infection to kill the potential range of pathogens. Dr. Schmidt stated two reasons for talking about the concept of independent action and synergy at a biomarker meeting:

1. There is a need for biomarker refinement to know which patients will benefit from one drug alone without being subjected to the risks of the drug combination.
2. It is critical to find the rare subgroups that do benefit from drug synergy.

Dr. Adam Palmer highlighted that drugs have variable effects in populations and agreed that identifying patient sub-populations is important. He argued that the greatest benefit of adding a drug into a combination will occur in the patient subset in which the drug is most highly active, and not in a patient subset in which drug synergy is anticipated to create the benefit. Dr. Palmer stated that the greatest benefits of drug combinations will occur in subpopulations—whether the combination succeeds by additivity or synergy.

Dr. Curran noted three reasons to be optimistic about synergy:

1. There is a deeper understanding of biology combined with AI-aided tools that can identify cancer dependencies and synthetic lethality that translates clinically.
2. Engineering tools have improved. There are drugs now that could not be made before, such as bi-specific drugs.
3. Tools exist to monitor patient response in real time.

Dr. Sorger emphasized that pre-clinical research needs to consider variation in the population. Dr. Curran agreed and stated that tuning animal models to conditions that will show synergy may not accurately predict clinical relevance, and this is particularly applicable to drug synergy in advanced disease.

Dr. Carl Barrett introduced the concept of tumor heterogeneity and reminded the group that biomarkers are heterogenous. He also noted that not all patients have the same degree of response to combinations, and he asked if the depth of response is another feature to consider in the conversation. Dr. Palmer stated that the level of inter-individual variability is the key element of the independent action model, and the population-level response to a combination can improve

even if a patient's individual response is affected by the one best-killing agent for that individual. He added that early-stage trials can have outliers that look promising as a form of synergy, but the larger late-phase trials result in what is expected by additivity, and this can be reliably predicted. Dr. Palmer argued that the rate of failure of phase 3 trials of drug combination could be at least halved by avoiding drug combinations that lack a single-drug benefit.

Dr. Elad Sharon noted his publication with Dr. Jared Foster in *JNCI* that described randomized comparative trials to evaluate combinations of experimental agents.

Panel Discussion: When will precision medicine replace the "bet hedging" used in current combination cancer therapies?
Geoff Oxnard, MD, Loxo Oncology; Emmett Schmidt, MD, PhD, Merck; David Solit, MD, MSK; Baolin Zhang, PhD, Office of Biotechnology Products, FDA

Dr. Schmidt opened the panel discussion and asked what precision medicine will look like in the future.

Dr. Geoff Oxnard reiterated the problem of unperceived heterogeneity. He expressed hope that biomarker development will reduce heterogeneity in the trial population so that sub-populations who do not require combinations, or those who would benefit from drug synergy, can be iteratively identified.

Dr. David Solit argued that the term synergy is not useful for developing rational combination strategies, and he emphasized the importance mechanistically-based combinations that prevent or delay emergence of drug resistance for more durable responses and cures. He highlighted the HIV example of rationally based combinations that transformed the field and led to dramatic improvements, and he stated that technologies that identify bypass mechanisms in individual patients would inform mechanistically-based combinations that can change an acute fatal disease into a chronic disease.

Dr. Baolin Zhang is a Chemistry, Manufacturing, and Controls (CMC) reviewer with FDA's Center for Drug Evaluation and Research. From a drug development point of view, the differentiation between independent action and synergy of drugs is less important than the overall risk-benefit assessment of the biomarker. He noted that the FDA recently published a guidance document called "Guidance for Industry: Benefit-Risk Assessment for New Drug and Biological Products," which outlines basic principles for biomarkers. Dr. Zhang reminded the group that context of use is important for biomarker selection and urged consideration for cancer type, disease stage, available therapies, and how relevant the biomarker is to the therapeutic target.

Dr. Schmidt returned to his opening question about when precision medicine will replace bet-hedging in cancer therapies and discussed biomarkers for combination treatment. He argued that a combination biomarker is a fundamental separation from independent action and that the focus should instead be on the individual treatment components. Based on a study that his group published, the duration of response data for eight out of ten combinations matched what would be predicted for independent action. The two combinations that did not fit that paradigm were ipilimumab and nivolumab for melanoma and the VEGF therapies.

Dr. Oxnard added that bet-hedging is easier than biomarker assessment, and while precision medicine is appropriate for some patients, some combinations are clinically synergistic and adding drugs is warranted in those cases.

Dr. Malcolm Smith stated that combinations with true pre-clinal synergy often require dose reduction in the combination setting due to excessive toxicity from synergistic effects on normal tissues. He posited that this may contribute to failure in the clinic of the synergistic combinations.

- Dr. Curran responded that engineered drugs that are selectively active in the tumor microenvironment are being developed, such as antibody-drug conjugates.
- Dr. Solit agreed that tumor selective drugs may allow for successful combination treatments that do not create additional toxicity.
- Dr. Palmer noted that their model of additivity set the upper bound to efficacy in the majority of assessed FDA approvals, and this may be explained by enhanced toxicity of drugs in combination. The phenomenon of clinical synergy in which a drug combination leads to decreased toxicity, as seen with BRAF and MEK inhibitors in BRAF mutant melanoma, in the absence of anti-tumor synergy was also discussed.

Dr. Zhang stated that the FDA has approved 32 biomarkers, and most of them are the therapeutic target itself. He is on an FDA-NIH working group to develop a guidance document on multi-component biomarkers that combine two biomarkers into a composite biomarker.

Dr. Mike Espey asked how to anticipate acquired resistance and not select out one clone that is in competition with other clones. The panelists responded that understanding tumor heterogeneity in real time to target different clones is not yet possible, and one reason to give combination treatments up front is to prevent resistance from developing while giving one drug before treating with the most effective drug. Dr. Curran highlighted the opportunity to block reproducible metabolic salvage pathways with drugs focused on tumor metabolism.

Next, the panelists discussed when up front combinations are appropriate. Dr. Palmer stated that giving up front combinations with curative intent is appropriate. However, in disease settings where a cure is not achievable, he believes the answer is not as clear as to whether up front combinations or sequential therapy is preferred. He noted that not all patients who receive sequential therapy survive to receive the second line treatment. For diseases in which patients routinely cross to multiple treatment lines, there is a lack of clinical trials testing whether sequential therapy is better or equivalent to up front combinations. Dr. Palmer cited a trial showing that patients with metastatic breast cancer had an equivalent overall survival (OS) benefit and superior quality of life when receiving sequential chemotherapy. Dr. Oxnard wondered if effective pharmacodynamic markers could enable effective switching through therapies to identify the best treatment.

Dr. Javier Perez asked about timing sequential treatments to a measured lack of response combined with an evaluation of resistance mechanism. Dr. Oxnard replied that the existence of those biomarkers would allow precision approaches, but an additive approach is required in the absence of those biomarkers. He emphasized that finding resistance as it emerges is challenging.

Other panelists reminded the group about examples, such as the RAF and MEK inhibitor combination, that work in combination but not sequentially if resistance develops. Dr. Solit argued that inhibiting bypass pathways to these combinations is the next step, and laboratory work could identify the most common bypass mechanism to design a triple drug approach that would work in a subset of patients. However, he expressed concern that a biomarker is not available for this, and he emphasized the need for multiple approaches to reach consensus on the best biomarkers for larger drug combinations. Dr. Zhang stressed the importance of standardization in biomarkers across product class so that all clinicians can follow a protocol.

Dr. Curran suggested categorizing technical biomarkers into three classes:

1. Predictive biomarkers that dictate the choice of treatment
2. Predictive and prognostic biomarkers that provide treatment intensity guidance
3. Prognostic biomarkers that also reveal why a treatment failed

Dr. Alex Snyder asked the panel to identify a focus area within this topic that is appropriate for the FNIH.

- Dr. Zhang replied that precision medicine requires a biomarker, a diagnosis, and more effective therapies combined with standardization for consistency. He argued that collaboration between the pharmaceutical, diagnostic, and regulatory industries is required to develop consistent guidelines for biomarker diagnoses.
- Dr. Solit emphasized the need for varied, decentralized approaches.
- Dr. Oxnard highlighted the importance of digital pathology.
- Dr. Curran imagined a single test to understand the immunology, genetics, and treatment response to guide clinical decision making.
- Dr. Palmer argued that biomarkers will exist for individual drugs within a combination and not for the combination treatment. His advice for drug development is to find the molecular features that define the patient population in which that single drug has activity or resistance, and then use those features for patient enrollment during combination treatment trials.
- Dr. Schmidt emphasized the importance of early disease detection and treatment because surgery alone in some early disease cases can eliminate the need for any drugs, but there is currently no way to stratify risk and determine if a patient needs additional treatment following surgery.

Keynote: Novel Endpoint Development in Oncology
Nicole Gormley, MD, FDA

Dr. Nicole Gormley is the Division Director for the Division of Hematological Malignancies at the FDA as well as the agency's Acting Associate Director for Endpoint Development within the Oncology Center of Excellence.

The International Council for Harmonization (ICH) states that a primary endpoint should provide a valid and reliable measure of a clinically relevant and important treatment benefit. Within the U.S. regulatory approval pathway, drugs can receive regular approval based on demonstration of clinical benefit or an effect on an established surrogate. Drugs can receive accelerated approval for treatment of serious or life-threatening illness based on a surrogate endpoint reasonably likely to

predict clinical benefit or an intermediate clinical endpoint. Accelerated approval may require post-approval trials.

Dr. Gormley noted several cancer endpoint considerations:

- OS is objective and measures safety and efficacy but takes longer.
- Progression free survival can be assessed earlier but may not correlate with OS.
- Response rate may not correlate with OS.
- MRD may be impacted by assay performance.

Biomarkers have multiple potential uses. The clinical and regulatory uses of biomarkers often have different purposes, and the trials and data needed to support the biomarkers in each of these settings is different. Clinical uses include screening, monitoring, and guiding treatment decisions. Regulatory uses during clinical trial design include patient stratification, patient selection or enrichment, risk-based treatment assignment, and biomarker use as a surrogate endpoint.

The factors that are considered in the regulatory assessment of a biomarker include the risk introduced by use of the biomarker, the biological rationale, assay considerations, and the ability to assess its strength of association with the proposed clinical outcome. The two fundamental criteria during biomarker assessment are biomarker clinical validity and assay analytical validity.

Next, Dr. Gormley presented two ways to validate surrogate endpoints for regulatory use, including prentice criteria and meta-analytical methods. When using a meta-analysis, the inclusion of more trials as well as trials with a range of treatment effects increases the rigor and accuracy of surrogacy assessment. However, there are times when use of a surrogate may not be appropriate.

Recently, the FDA published two drug-development guidance documents: a document for the use of MRD in hematologic malignancies and a document for the use of ctDNA in early-stage solid tumors. Dr. Gormley provided an update on the current use of MRD, including specific examples of how MRD is being used to support accelerated approvals, when MRD results have been included in prescribing information, and ongoing efforts in various diseases to formally evaluate MRD. She noted that some reasons MRD data may be excluded from FDA labels include analytical test validation issues, assay performance issues, and trial design and data collection issues. The BELLINI trial and trials of various PI3K inhibitors highlight the possibility of concerning discordance between early efficacy endpoints, such as MRD, and OS.

The FDA co-sponsored a workshop with AACR and the American Statistical Association to discuss best practices of trial design, analysis, and interpretation of OS in oncology clinical trials. The workshop takeaway was that a rigorous plan for assessment of OS as a safety endpoint is important when it is not a primary or secondary endpoint, and the results of the OS analysis may affect the appropriate regulatory pathway.

Next, Dr. Gormley introduced Project Endpoint, a multidisciplinary group that aims to enhance use of early endpoints and foster engagement with stakeholders committed to the advancement of endpoints in oncology drug development.

Biomarkers have the potential to expedite drug development, but there are existing uncertainties. Dr. Gormley emphasized that biomarker assessments and novel endpoints in clinical trials should be discussed with the FDA.

Discussion
During a brief discussion on establishing MRD as a secondary endpoint in Multiple Myeloma, Dr. Gormley highlighted a few lessons learned so far:

- International collaboration and consensus on a standardized level of assessment is important.
- A repository can help move the field forward.
- MRD in myeloma is not a surrogate, but Dr. Gormley is hopeful it can be used as an intermediate clinical endpoint when compiled with OS data.
- Even in the precursor disease setting, there needs to be confidence in early endpoint assessment reliability.
- Relying on OS as the primary endpoint is challenging, and this drove the consideration of OS as a safety endpoint when in combination with other earlier endpoints.
- Some biological and clinical uses of biomarkers are separate from regulatory uses.

Dr. Sudhir Srivastava asked about the use of reduction in late-stage disease as an endpoint for screening trials. Dr. Gormley replied that any endpoint, regardless of trial type, must have data to show a clinically meaningful therapeutic benefit, even in early-disease settings.

Working Session 1: Digital and Computational Pathology

Moderator: Brandon Gallas, PhD, FDA

Dr. Brandon Gallas introduced the working session on digital and computational pathology. He explained that the focus is on datasets, specifically image data that begins with a glass slide containing H&E staining.

Digital and Computational Pathology Past, Present and Future

Carl Barrett, PhD, UNC

Dr. Barrett summarized the last two centuries of cancer diagnosis, beginning with the first H&E staining in the late 1800s to molecular pathology in the early 2000s to computational pathology today. Only a small percent of testing laboratories in the Unites States and Europe use digital pathology. Therefore, there is a lot of potential, but low access and implementation. There has been recent growth in digital pathology publications and interest.

Next, Dr. Barrett shared examples of digital pathology being used to develop diagnostics for research in a non-regulatory way. He divided them into three categories by technology and described the benefits of digital pathology for each, including:

- H&E slides – Machine learning with image analysis improves quantification efforts.
- Immunohistochemistry (IHC) – Image analysis informs cell proliferation rates, PDL1 and HER2 scoring, IO response, and quantitative scoring efforts.

- Multiplex Immuno-fluorescence (mIF) – Multiple markers enable powerful quantitative methods. In addition, mIF can be combined with H&E and spatial transcriptomics in AI research, which has value for the pathologist and the patient.

Advantages of digital pathology for drug discovery include discovering drug targets, prognostic features, and predictive markers. For example, collagen fiber orientation disorder from H&E images is prognostic for early-stage breast cancer. Dr. Barrett highlighted the importance of "measuring" versus "reading" in pathology, noting the relevance of HER2 expression in breast cancer.

Challenges in digital pathology and AI research include:

- The large size and variable quality of data sets
- A lack of cohort diversity
- Challenges with data sharing
- The need for established standards
- AI that is interpretable
- Health record and digital pathology unification
- Regulatory issues
- Community access

There are ongoing efforts by NCI, FDA, and advocacy organizations to address these challenges.

Discussion

Dr. Srivastava asked if synthetic data is acceptable to the FDA. Dr. Barrett replied that independent data sets with rich cohorts and diverse data is essential, but having access to the data sets is also critical. Dr. Srivastava suggested that synthetic data could help overcome the problems of missing data.

Introduction to the Pathology Innovation Collaborative Community (PIcc): Challenges and Needs in Digital Pathology

Joe Lennerz, MD, PhD, Massachusetts General Hospital, Harvard Medical School

Dr. Joe Lennerz introduced the challenges of moving digital pathology into the clinical space. Currently, a cancer diagnosis is made with H&E slides and a microscope. Therefore, digital pathology is a fundamental change in the practice of medicine for all of cancer. The workflow of digitizing pathology data is complex, and the clinical impact can only be realized with a focus on patient needs and regulatory-science-grade tools.

The Pathology Innovation Collaborative Community (PIcc) is a free collaborative community that brings together multiple stakeholders in a precompetitive space to tackle complex problems in digital pathology jointly. The group meets monthly, and membership is currently over 700 people. Dr. Lennerz stated that each stakeholder has a preferred way to solve complex digital pathology problems, but the field needs more collaboration. The remaining speakers of the working session were selected to represent certain vantage points to help explain why digital pathology would benefit from a collective effort.

MedPerf: An open benchmarking platform for medical AI, enabling healthcare stakeholders to assess the performance of AI models in an efficient and human-supervised approach
Alex Karargyris, PhD, MLCommons

Dr. Karargyris is a representative of MLCommons, a non-profit global community started in 2018 with a focus on benchmarking for speed and best practices in machine learning. The organization is structured into working groups, one of which is a medical working group focused on best practices for machine learning in medical AI. For the past two years, they have worked on MedPerf, an open-source community platform for federated medical AI benchmarking.

Real-world AI validation is important to ensure relevance for real-world settings. This requires diverse data and data sharing, which is difficult due to technical and regulatory reasons as well as privacy concerns. MedPerf evaluates AI models on diverse real-world data. It does this based on federated evaluation of AI, which is driven by stakeholders, runs on real-world data, remotely deploys models, and alleviates data privacy concerns. The approach integrates human accountability throughout the process.

The design of MedPerf includes a benchmark committee composed of regulatory bodies, patient advocacy groups, clinicians, and data/model owners that defines benchmarks in clinically impactful areas. Some of those who can benefit from MedPerf include AI researchers, healthcare providers, patients, and regulators.

Use cases for MedPerf include:

- Supporting 71 hospitals for federated learning and an evaluation experiment
- Benchmarking 40 models across 31 sites for brain tumor segmentation
- Supporting neuro-oncology federation at over 90 hospitals
- Partnering to add model privacy technology in a clinical trial

Dr. Karargyris listed four high-level factors for a successful benchmark, including:

1. Clinical relevance
2. Data harmonization
3. Diverse and representative data access
4. Reference standards and end user quality controls

In summary, MLCommons offers technical and benchmarking expertise with neutral governance. The group looks to support impactful industry-wide benchmarks that are scalable and sustainable.

Collecting, curating, and sharing regulatory-ready data to assess the performance of pathologists and AI/ML models
Brandon Gallas, PhD, FDA

Several barriers to digital pathology exist, including:

- Pathologist variability in creating reference standards and utilizing AI outputs
- Too few statisticians supporting the field
- Need for FDA cleared devices to produce the data
- Limited availability of data

Dr. Gallas argued that regulators can proactively engage with all stakeholders in the community to provide input in a pre-competitive space. He highlighted the MedPerf model as an example where regulators can contribute to a benchmark committee to help regulatory-ready data emerge from community efforts. Dr. Gallas highlighted what he expects in submissions of Software as a Medical Device, such as indications for use, technological characteristics, imaging modalities, database information, reference standards, and assessments. Reference standards and assessments are both impacted by pathologist variability. Dr. Gallas published a Multi-reader, Multi-case (MRMC) analysis software that completes the statistical analysis needed to account for pathologist and case variability.

To determine when an AI/ML model is sufficient, Dr. Gallas argued that expert-to-expert agreement can serve as the performance baseline for software-to-expert agreement. This led to the High-Throughput Truthing (HTT) Project, which aims to create a dataset of pathologist annotations to validate AI models analyzing digital scans of pathology slides. HTT deliverables and tools include data-collections methods and platforms, pathologist training, methods to validate a quantitative algorithm, and ultimately an FDA-qualified dataset.

As an FDA employee, Dr. Gallas has encountered challenges in managing the HTT project. He is seeking a multi-stakeholder partner who can manage the submission to the Medical Device Development Tools (MDDT) program in FDA's Center for Devices. The partnering organization would receive the HTT tools and deliverables, the initial feedback from MDDT, existing data, and the team's expertise and experience. Dr. Gallas re-emphasized the importance of regulators helping to contribute regulatory-ready data.

Bridge of Trust – Closing Gaps between Patient Care and Innovation
Matthew O. Leavitt, MD, LUMEA & DDx Foundation

Standardization at the level of the lab as well as the clinic is critical for digital pathology to become useful. Market forces can drive standardization, such as tools that make it easier for a clinician to collect tissues in a standardized way. He argued the same concept applies to the laboratory setting, and changing processes must make life easier and less expensive for adoption of standardization processes.

Next, Dr. Leavitt described their virtual histology practice model. In this model, digital histology centers (DHCs) consolidate specimens and pathology image data from numerous hospitals and laboratories. This enables small hospitals to have a pathologist on staff at the center and immediate access to digital pathology. For example, PathNet includes 3 DHCs that serve hundreds of clinical sites across 35 states. There are over 2 million imaged prostate samples linked to patient metadata, and relationships with the local providers are maintained for ongoing follow-up. This is an aggregation of real-world data from diverse and underserved communities that otherwise would lack access to standardized digital pathology processes.

The Digital Diagnostics (DDx) Foundation is a non-profit organization that bridges gaps between patients, technology companies, and healthcare providers. It aims to provide innovators with the data and biomaterials needed to advance digital medicine while protecting patients. Data from

DHCs is curated, anonymized, and aggregated in a biobank that can share the anonymized real-world data with third parties through a secure blockchain network. Through this distributed biorepository, individual patients can be identified for new therapies or clinical trials and notified by their existing care providers.

A Multi-Stakeholder Regulatory Science Consortium to Create a Sustainable Ecosystem for Data Driven Diagnostics
Joe Lennerz, MD, PhD, Massachusetts General Hospital, Harvard Medical School

Dr. Lennerz explained that a collective effort is required to navigate the complexity of digital pathology toward a regulatory-ready solution, including a robust platform with real-world regulatory-ready data. During the 2023 PIcc annual meeting, representatives from diverse professional societies narrowed their interest to the topic of a regulatory pathway for real-world data collection and evaluation methodologies for AI driven devices.

Dr. Lennerz emphasized that essential stakeholders have come together to develop this platform, but they lack the required resources. He stated that FNIH would be an ideal partner for this endeavor, citing FNIH's demonstrated commitment to collaborative initiatives, its proven track record in navigating complex challenges, and a shared vision for advancing the field through collaborative efforts. A real-world regulatory-ready dataset would have enhanced regulatory confidence and submission process, interoperability and consistency, and accelerated adoption.

The international component of PIcc developed recommendations for compiling test datasets for evaluating AI solutions in pathology. The paper outlines considerations from a diagnostic perspective and identifies the appropriate regulatory pathway. They also published initial interactions with the FDA developing a validation dataset as a medical device development tool.

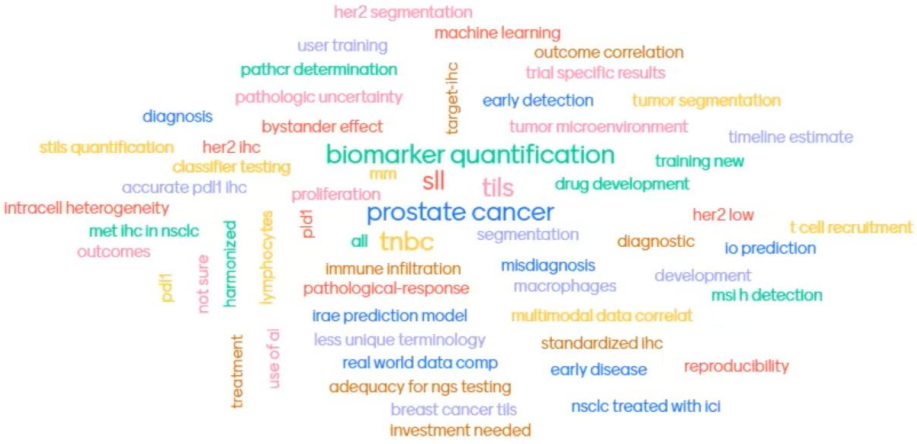
Dr. Lennerz proposed three initial use cases for the platform, including H&E based TILs quantification, HER2 IHC quantification, and tumor segmentation.

In summary, the goal is to form a central hub organization, create and submit a first dataset for regulatory qualification, and implement a defined roadmap for additional use cases.

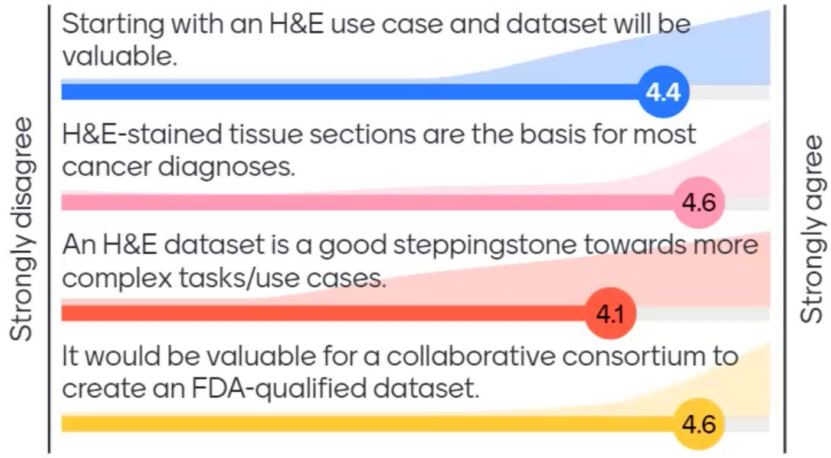
Interactive Data Gathering

Participants were asked a series of questions related to building a regulatory-ready data set of digital pathology slides and patient metadata sourced and curated from a diverse set of real-world health providers. The following questions were asked:

For the proposed regulatory-ready dataset, what are some use cases you would like to see treated?

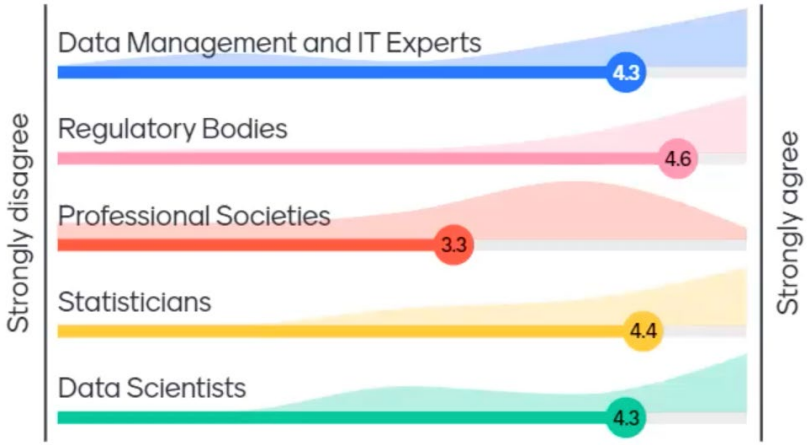


Please rank your opinion about the following statements:



Who should be involved in this data gathering effort?

Who should be involved in this data sourcing effort?



Who else should be involved in this effort?

Panel Discussion on Digital Pathology
Moderator: Joe Lennerz, MD, PhD, Massachusetts General & Harvard Medical School
Panelists: Carl Barrett, PhD, UNC; Kim Blenman, PhD, MS, Yale; Brandon Gallas, PhD, FDA; Alex Kalof, MD, Association of Directors of Anatomic and Surgical Pathology; Alex Karargyris, PhD, MLCommons; Laura Lasiter, PhD, AstraZeneca; Matthew O. Leavitt, MD, LUMEA & DDx Foundation

Dr. Lennerz asked each new panelist to briefly describe their roles and contributions in the field of digital pathology.

Dr. Kalof is Vice Chair for the University of Vermont Health Network. Digital pathology has been identified as a necessity for their network to deal with pathology services under pressure from high volumes, staff shortages, and financial pressures. Their network has no digital services, and they lack funds for a capital investment. Dr. Kalof feels fortunate to work with the DDx foundation.

Dr. Blenman is an assistant professor of Medical Oncology in the Yale School of Medicine and an assistant professor of Computer Science in the Yale School of Engineering and Applied Science. She researchers predictive and prognostic tools for treatment emergent adverse events, processes to verify and validate AI algorithms for use in clinical care, and equity in health and STEM. Dr.

Blenman is working with the FDA to create a reference standard for using annotations to evaluate Software as a Medical Device.

Dr. Lasiter is a virologist by training and joined the panel to share her collective professional experience working for the Senate, Friends of Cancer Research, PhRMA, and AstraZeneca. She emphasized the importance of shared learning gained from pre-competitive, collaborative partnerships to accelerate novel therapies and treatments.

Dr. Oxnard asked how to use images from patients with and without pathological complete response to implement an experiment that assesses if the centers can use digital pathology technology to assess a biomarker rigorously.

- Dr. Gallas replied that this approach needs to be built with standardized assessment methods and standardized data collection protocols.
- Dr. Barrett added that the hypothetical use case should be broadened from complete pathological response to partial pathological response, which is not clear how to score quantitatively.
- Dr. Leavitt agreed that standardization is critical and what works in a single setting may not apply elsewhere.

Dr. Espey asked if it is beneficial to consider longitudinal sampling beyond initial diagnostic samples and if there are exemplars to focus on first.

- Dr. Barrett appreciated the idea of longitudinal sampling but noted it is difficult in the metastatic setting. He suggested longitudinal sampling in the preclinical setting would help inform an understanding of cancer evolution, and there are collections of those samples.
- Dr. Leavitt described a case in which a company evaluating MSI-high prostate cancers developed an algorithm able to identify MSI-high cancers in a Detroit community setting, and some of those patients progressed to metastatic disease. He urged caution about moving too quickly and emphasized the need to work collectively to advance these technologies.
- Dr. Gallas replied that he is focused on exemplars and leveraging a diverse set of health providers to capture patient diversity and characterize key metadata. His goal is to establish qualified datasets and help bring in additional stakeholders to interact with the medical device development tool program.
- Dr. Blenman added that the consortium process allows others to observe what is required and avoid missing data.

The panelists discussed how digital pathology can benefit clinical trials and drug development.

- Dr. Barrett explained AI and digital pathology could empower an understanding of the biological processes and provide predictive markers for drug developers. In addition, he noted drug development is moving into early diseases, and a combination of computational pathology with longitudinal ctDNA measurements could be a powerful technology.
- Dr. Lasiter replied that oncology is moving more into a chronic disease state, and industry is beginning to focus on how to predict which patients will have a recurrence and design drugs to prevent that recurrence.
- Dr. Kalof emphasized the importance of accurate diagnoses and current pressures for sub-specialization. Digital pathology allows access to sub-specialists and rapid communication.

- Dr. Blenman noted that many biomarkers and tools are qualitative or semi-quantitative in nature, and digital pathology and computational pathology will enable the following:
    - Setting thresholds
    - Answering questions faster
    - Uncovering complex patterns that humans cannot detect
    - Reducing or limiting biases in clinical trials
    - Facilitating equity
- Dr. Leavitt stated that digital pathology will demand consolidation of tissues and histology given financial constraints and limited infrastructure of small hospitals. With consolidation, the screening process for identifying patients and accelerating acquisition of patients into clinical trials could begin in the digital hubs with relationships with the clinics.

The panelists discussed the challenge of obtaining patient samples from clinical centers.

- Dr. Leavitt agreed many institutions understand the value of patient data. Small community hospitals have too little data to be of value individually, but there is value in the aggregation of data from every demographic. The foundation is focused on modernizing the ability to capture patient data under one umbrella in a way that respects patient privacy.
- Dr. Karargyris added that the technology is available to support the governance and participation of small hospitals in the democratization of AI.

The panelists discussed the incorporation of outcomes data to validate novel phenotypes from a digital pathology endeavor.

- Dr. Barrett noted that algorithm outputs can be associated with patient outcomes, but that does not always replicate in another dataset. He argued that the federated data provides a timely avenue for algorithm validation.
- Dr. Gallas stated that patient outcome is the gold standard, and a dynamic dataset with automated assessments can begin to address biomarkers that the pathologist cannot see. He emphasized that their project is not prescriptive on data set characteristics to include, but instead they are bringing pieces of a proposal forward to create a reference standard.

Dr. Lennerz summarized the panel discussion. Individual institutions are incorporating digital approaches, but they are using individual solutions that are diverse and not interoperable. The creation of a real-world, regulatory-ready data collection effort could preemptively solve this problem of interoperability among individual entities.

Day 2: Wednesday November 15, 2023, 8 am – 5:30 pm

Opening Remarks
Dr. Geoff Oxnard welcomed attendees and emphasized the goal to discuss collaborative, pre-competitive research that advances biomarkers for precision medicine in drug development.

Session 2: Early Disease – Detection and Treatment

Moderator: Gary Kelloff, PhD, NCI

Dr. Kelloff introduced the Day 2 agenda, highlighting the benefits of precision medicine and the opportunity to diagnose biopsies in early disease to obtain information about drug targets and the potential for drug activity. Up to 90% of cancer deaths are from metastatic disease, but precancers can exist for a long time emphasizing the importance of biomarkers for cancer prevention.

Dr. Kelloff introduced Dr. James Doroshow, Director of NCI's Division of Cancer Treatment and Diagnosis.

Keynote: Developing Combination Therapies for Rare Tumors
James Doroshow, MD, NCI

Dr. Doroshow introduced NCI's Patient-Derived Models Repository (PDMR) and the rare patient-derived xenograft (PDX) tumor models. He noted that a subsequent project stemming from this work was an in vivo drug screening to learn about response assessments for xenografts.

The NCI-PDMR started about 7 years ago with a goal to establish a repository for individuals without access to their own PDX models library. There are now about 800 models that are clinically annotated and available at low cost, with genetic information available online. In addition, 2D cells lines and autologous cancer-associated fibroblasts (CAFs) were established.

Many of the incoming tumors for the repository were rare tumors. This presented an opportunity to develop methodologies for creating PDX models of understudied cancers and use the well-characterized models to develop a systematic in vivo therapeutic screening effort.

NCI-PDMR developed 39 PDX rare cancer models that were based on a high need for effective therapies, a lack of new therapies available in the last decade, and the existence of a patient population in NCI clinics. An expert group recommended a series of combinations of novel agents.

The rare tumors PDX study methodology included three phases:

1. Testing novel therapeutic combinations in small treatment cohorts with quality control following every tumor passage
2. Repeating the study with single agent arms when a response was observed in combination
3. Performing a full efficacy study with biomarker exploration in combinations that had additive or synergistic effects

A total of 57 agent combinations are being tested, and testing should be complete within the next four to six months.

Next, Dr. Doroshow discussed how to define an in vivo response and suggested using multiple parameters. First, he presented a binning example with qualitative response criteria. Second, he listed potential quantitative assessments, including regression, area under the curve, and time to volume quadrupling. The deconvoluted data from a paclitaxel-nilotinib combination to treat different cancers was shown to highlight examples of an unpredicted combination effect. Several other combinations also showed a combination effect using the response assessment criteria. There can be a large amount of heterogeneity of response across the same tumor type.

The Rapid Analysis and Response Evaluation of Combination Anti-Neoplastic Agents in Rare Tumors (RARE CANCER) study leverages robust preclinical combination data for agents with phase 2 dosing information to inform a clinical approach. Patients with eligible histologic subtypes can join these rare cancer clinical trials with successive enrollment in study cohorts as the cancer progresses. If there are early signs of efficacy, the patient can move into one of the NCI therapeutic networks to take advantage of what is observed. The primary objective is to evaluate the proportion of patients with advanced rare cancers who have objective responses to treatment with study agents. Patients on the nilotinib and paclitaxel study have already had responses, even in patients who are definitively Taxol resistant. There is a continuing need to have more data to apply to patients with rare cancers to understand if preclinical evidence can usefully guide entry into clinical trials.

Discussion

Dr. Kelloff asked about the status of Combo-MATCH. Dr. Doroshow replied it is not a rare tumor trial but instead a series of phase 2 trials open at 300 sites with others in the process of opening. The first four trials have screened 60 patients, and four to six trials are opening soon.

Dr. Barrett asked if liquid biopsies can be incorporated into PDX models. Dr. Doroshow replied certain histologies in mouse models develop circulating tumor cells (CTCs), and they are trying to understand how the CTCs correlate with the in vivo xenograft response. For some histologies, there are thousands of CTCs, which allows for interrogation without killing the mice.

Dr. Bruce Johnson asked what types of tumors withstand the shipping process and if some tumor types do better than others. Dr. Doroshow replied that colon cancer from either primary or metastatic sites will successfully create a PDX at least 50% of the time. Pancreatic cancer and melanoma have rates of about 30 to 40%, and sarcomas perform better than expected. Renal cancer and ER-positive breast cancers are challenging, with pancreatic cancer being among the most challenging to create a xenograft model.

Dr. Espey asked how the lack of the immune system in the PDX models impacts mechanism of action of drug combinations. Dr. Doroshow replied that there is a collaboration between PDXnet and the repository to take advantage of the matching autologous CAFs with the PDX models, and there potential to use technology that can convert autologous CAFs into immune cells.

Dr. Palmer asked if any histologies with common features emerged as having a common response to certain drug mechanisms. Dr. Doroshow replied that they meet every week to address the genomics of responses and to find biomarkers, but do not yet have an answer.

Neoadjuvant Immunotherapy: Leveraging the Immune System to Treat Early-Stage Disease
Bruce Johnson, MD, DFCI

The Lung Cancer Mutation Consortium, supported by the American Recovery and Reinvestment Act under a NCI Grand Opportunity Program, was formed in 2008-2009 to determine the frequency of oncogenic drivers in 1000 patients with advanced adenocarcinoma lung cancer. Of the 1000 patients, 733 people were genotyped for ten genes through multiplex testing. The study identified a distribution of mutations, but at the time no targeted agents were approved for any of the genomic changes. Now, most of these mutations have targeted agents, and patients with an oncogenic driver who are treated with a targeted agent live about a year longer than those who were not treated with a targeted agent.

In 2017, the neoadjuvant atezolizumab study in patients with locally advanced non-small cell lung cancer (NSCLC) was launched as an open label, single-arm, phase 2 study of neoadjuvant followed by an optional adjuvant atezolizumab. The study met its primary endpoint of a 20% major pathological response rate after neoadjuvant atezolizumab in patients who had surgery, excluding those with EGFR or ALK alterations. A subset of patients had a complete response, which is almost never observed with chemotherapy and radiation. For patients who received both treatments and went on to surgical resection, their survival outcomes were substantially better than previously observed with stage 3 disease. A visual representation of the data showed the relationship between maximum tumor response, genomic changes, and tumor mutation burden. Specifically, those with KRAS and STK11 mutations experienced the worst outcomes. This finding was again observed in another cohort with single agent checkpoint inhibition treatment for NSCLC.

Dr. Johnson described an IMMUNOME flow cytometry study to identify patient sample subsets associated with maximal tumor response. The study found that a higher prevalence of non-natural killer (NK) cells expressing specific immunoregulatory receptors in pre-treatment peripheral blood was associated with major pathological response, and a higher prevalence of NK and NK-like T cells with inhibitory receptors was inversely associated with major pathological response. This offers clues about how to incorporate immunomodulatory molecules.

Lastly, he discussed the Checkmate-816 trial, stating its findings were one of the most rapid uptakes in changing practice. The trial showed adding nivolumab to chemotherapy followed by surgery and adjuvant therapy leads to pathological complete response in 24% of patients compared to 2% of patients with chemotherapy alone and an OS hazard ratio of .6 in favor of combination. Similar outcomes have been seen with the use of pembrolizumab in the neo-adjuvant space.

In summary, these examples in locally advanced NSCLC showed that genomics are helpful in deciding who is likely to respond to treatment, revealed potential targets to improve treatments, and showed the benefit of combining chemotherapy with checkpoint inhibitor therapy.

<u>Discussion</u>
Dr. Johnson was asked how the NK cell metrics in the periphery contribute to overall pathological response in the tumor microenvironment. Dr. Johnson replied that the cancer moonshot is assessing the impact of the tumor microenvironment using single-cell sequencing as well as spatial analysis in patients receiving immunotherapy before and after generating drug resistance. He noted a

specific study assessing the similarities and differences between melanomas and lung cancers in response to immunotherapy.

Dr. Oxnard asked about the extent to which pathologists can determine complete versus major versus partial pathological response. Dr. Johnson stated that The International Association for the Study of Lung Cancer has a consortium that assembles pathologists for central review of responses to establish agreement on cases.

Dr. Oxnard asked if there are cases in which single agent immunotherapy will be used in the neoadjuvant setting. Dr. Johnson replied that picking out patients with high PD-L1 values and excluding those with specific genetic variants could potentially select out a group for treatment with single agent immunotherapy.

Dr. Bhadrasain Vikram asked if adjuvant checkpoint inhibitor made a difference. Dr. Johnson replied that those who received adjuvant checkpoint inhibitors fared better but had to be well enough to receive treatment.

Dr. Palmer asked for Dr. Johnson's opinion about using the measure of PD-L1 ≥ 1%. Dr. Johnson replied that those who have < 1% PD-L1 do not have much of a benefit in OS. If patients have any form of contraindication for checkpoint inhibition, Dr. Johnson does not offer it in the adjuvant setting.

Panel Discussion on Early Disease Detection and Treatment
Nicholas Dracopoli, PhD, DELFI; Jennifer Litton, MD, MDACC; Rob Iannone, MD, Jazz; Elad Sharon, MD, DFCI

Dr. Kelloff invited each panelist to introduce their specialty and give a brief presentation.

Dr. Dracopoli, who co-founded Delphi diagnostics, introduced cell-free DNA fragmentation as an approach for the early detection of cancer and the monitoring of treatment response in patients. DNA fragmentation patterns on DNA in circulation are determined by basal chromatin organization. The cell tightly controls open and closed chromatin to regulate gene expression. DNA in a closed configuration is protected from degradation, which leads to DNA fragments in circulation. There are enormous differences between normal cells and malignant cells in this fragmentation pattern. These fragmentation profiles are apparent without the need for targeted panels, which decreases assay cost. Dr. Dracopoli shared the fragmentation profile differences between non-cancer and cancer and explained that the profiles are used to create a DNA Evaluation of Fragment Length (DELFI) score that can distinguish between non-cancer and cancer cases. Early access to the tool is now available through the FirstLook Lung program.

Dr. Iannone described two programs at Jazz Pharmaceuticals that focus on earlier intervention. First, he described neoadjuvant targeted therapy to de-escalate intensity and reduce toxicity in early-stage breast cancer, which is a highly curative disease but with very toxic side effects with current treatments. Jazz pharmaceuticals is developing zanidatamab, a bispecific monoclonal antibody directed to different epitopes of the HER2 receptor. With a biomarker and the appropriate targeted therapy, there is an opportunity to reduce toxicity. A trial is underway at MD Anderson.

They have also partnered with I-SPY 2.2 with a goal to introduce novel targeted therapies in which there is a biomarker to reach pathologic complete response or minimal residual cancer burden with less intense therapy. Next, Dr. Iannone highlighted areas of opportunity to intervene earlier than is the current standard of care even beyond early-stage disease, such as with extensive-stage small cell lung cancer (SCLC). He described an ongoing phase 3 switch maintenance trial in patients with SCLC. In response to a question about the use of liquid biopsy to identify biomarkers in these trials, Dr. Iannone welcomed collaboration to identify the best way to move forward with that.

Dr. Litton, a breast medical oncologist and vice president of clinical research at MD Anderson, introduced the NeoSTEEP committee. The committee formed to standardize definitions for efficacy endpoints in neoadjuvant breast cancer trials. She briefly reviewed the committee's recommendations and highlighted the benefit of including non-binary endpoints such as residual cancer burden. In addition, the committee defined the time to event survival endpoints and emphasized the inclusion of event-driven endpoints. Dr. Litton summarized additional considerations, including:

- Upfront decision-making on when to look for responses, and consistency across sites in these decisions.
- The recommendation to use residual cancer burden as a secondary endpoint in neoadjuvant breast cancer trials, rather than a registrational endpoint.
- Consideration for nodal disease.
- Consistency in planned biopsies.
- Other endpoint considerations for hormone receptor positive breast cancer.

Dr. Sharon is the Clinical and Translational Director at Dana-Farber Cancer Institute following 16 years in NCI's Division of Cancer Treatment and Diagnosis. He focused on neoadjuvant trial design with immune checkpoint inhibitors. He showed event-free survival was markedly improved with neoadjuvant and adjuvant pembrolizumab treatment compared to adjuvant pembrolizumab alone in resectable melanoma in the S1801 trial. He highlighted use of event-free survival measured from the time of randomization as opposed to relapse-free survival measured from the time of surgery and noted that this design creates bias against the neoadjuvant arm. Dr. Sharon explained that they were interested in neoadjuvant work because they wanted to look at pathological complete response, which is not well established across oncology, and he showed that recurrence free survival could be segmented by pathologic response.

Next, Dr. Sharon shared two examples of neoadjuvant immunotherapy treatment of mismatch repair (MMR) deficient colon cancer, which is particularly sensitive to immunotherapy. For example, in a trial of nivolumab and relatimab in the neoadjuvant setting, 100% of patients had a pathologic response and 79% of patients had a pathologic complete response. He pointed out the 100% response rate in patients with MMR deficient locally advanced rectal cancer treated with neoadjuvant PD-1 inhibitor, highlighting the sufficiency of single agent treatment for that cancer.

In summary, Dr. Sharon concluded:

- Neoadjuvant trials require clinically meaningful endpoints, such as EFS and OS.
- Pathologic complete response can be explored as an endpoint with input from disease-focused pathologists and validation from established endpoints.
- Randomization is essential to determine contribution of effect in combination therapy.

- Neoadjuvant therapy can be a platform to inform biomarker discovery and drug development.

Dr. Schmidt asked if DNA measures are viable to identify patients who do not need additional adjuvant treatments. Dr. Sharon replied that a benefit of NCI's NCTN is that the samples are available to retrospectively validate those questions, and the trials can cross-validate each other. However, collected blood is a scarce resource and must be cautiously utilized.

Dr. Johnson asked panelists to comment on the relationship between the initial tumor volume and its impact on the pathological complete response.

- Dr. Litton replied the answer is subtype specific in breast cancer. In HER2-positive and triple-negative cancers, tumor size does not affect the pCR. In hormone receptor positive cancer, neoadjuvant therapy has not affected the endpoint in terms of tumor size.
- Dr. Sharon replied that he does not have an answer, but it is a great question, and the data may be available. He does not have an impression that it is a factor.

The panelists were asked if they have considered combining liquid biopsy approaches to boost sensitivity in early disease settings and how well each of the liquid biopsy approaches correlate within a patient population.

Dr. Dracopoli replied they are exploring those issues now, partly to lower screening cost with liquid biopsy. Methylation and fragmentation likely reflect the same biological event of gene regulation, and he does not expect much synergy. He anticipates the most orthogonal approaches will not be assays that are dependent on cell death given the limit of what is released into circulation during early disease. Instead, he sees promise in host-based immune response markers, which have the potential for signal amplification independent of the amount of tumor material released into circulation. Exosomes that are released by living cells are another orthogonal approach. Assessment of mitochondrial DNA amount has not yet revealed anything in early studies.

CSC Project Highlight: MGUS, MGIP, & Promise Study
Irene Ghobrial, MD, DFCI

Dr. Ghobrial introduced the topic of early detection and interception in multiple myeloma (MM). A precursor condition for MM called Monoclonal Gammopathy of Unknown Significance (MGUS) can be detected in the blood and can progress to MM many years later. The chance of progressing to MM from MGUS increases approximately 1% per year. The next stage of progression is called smoldering myeloma, and those patients have a chance of progression that increases by 10% per year.

About ten years ago, a study revealed that treating smoldering myeloma with lenalidomide plus dexamethasone improves OS. This opened the door to ask if patients should be screened for these precursor diseases to intercept MM. In addition, many new MM therapies have recently been approved, but patients are not cured because they are treated only after end organ damage. Dr. Ghobrial introduced three areas of focus:

1. Screening early for high-risk individuals.

2. Risk stratification using genomic and immune biomarkers for more precise risk assessment.
3. Early interception with more efficacious therapies, or earlier use of late-stage therapies.

The PROMISE study is screening 30,000 individuals at high-risk of developing MM to better understand the mechanisms of disease progression and the intrinsic and extrinsic factors that influence the progression. To date, over 12,500 participants have been enrolled from the U.S. and South Africa, with additional samples collected from Mass General Brigham Biobank. With the use of mass spectrometry, initial results have revealed early B cell changes and immune system alterations that occur with age. They also show that as many as 13% of the high-risk population who are older than 50 years have MGUS.

Next, Dr. Ghobrial discussed genomic alterations as biomarkers of disease progression to identify who will progress to MM. Her study found several major somatic alterations associated with progression to MM, particularly MYC alterations. Dr. Ghobrial presented several findings and opportunities from subsequent studies:

- Whole genome sequencing and single-cell sequencing of circulating tumor cells can be used to characterize clonal evolution from precursor disease.
- Malignant clones can be identified as early as MGUS.
- Single-cell RNA-sequencing of bone marrow samples can be used to reveal the differences between the malignant and normal plasma cells in precursor conditions.
- There are alterations in the immune microenvironment as early as MGUS.
- A reactive signature in the immune system prior to therapy is predictive of therapy response.
- A longer PFS was observed in those with post-therapy normalization of the immune system.

Now, the Immuno-PRISM study is comparing bispecific antibodies to lenalidomide plus dexamethasone, and the CAR-PRISM study is testing CAR-T therapy for smoldering myeloma.

Discussion

Dr. Kelloff asked if insurers are now paying for smoldering myeloma treatment. Dr. Ghobrial replied that it is an open question of who should be treated as standard of care.

Dr. Blenman asked about the incidence of MGUS and smoldering myeloma in the Asian, American Indian, Hispanic, and Pacific Islander populations. Dr. Ghobrial replied that the incidence is not known in those populations and that large, diverse population studies are needed to understand the incidence rates.

Dr. Johnson asked Dr. Ghobrial about how their trial will address geographical implications and diverse environmental exposures. Dr. Ghobrial replied that PROMISE is a nation-wide study that has opened at several sites with diverse populations. They are specifically trying to understand geo-spatial factors, such as pollution and obesity.

Dr. Oxnard asked if combination therapies that failed in late-stage disease should be repeated in early cancer given the differences in biology. Dr. Schmidt replied no, and Dr. Palmer added that

combination treatments could perform well in early stage if each agent has good single agent activity.

Dr. Palmer asked if the mass spectrometry (MS) assay can sensitively quantify the prevalence of pre-myeloma cells over time to determine if the development of disease is postponed or eradicated with treatment of early-stage disease. Dr. Ghobrial replied they assess MRD by adaptive and MS, and data is beginning to show that it is possible to use MS to measure MRD during patient follow up. However, there is no long-term follow-up data at this point.

Dr. Schmidt commended Dr. Ghobrial for considering the implications of treating a patient in precursor disease without knowing if clones will progress to frank disease. Dr. Ghobrial agreed and noted that none of the trials she discussed were for early MGUS.

Mr. Connors stated that this work has been approved by the CSC and EC for funding. He highlighted three opportunities for organizations to become involved with the CSC work, including:

1. The MGUS project presented by Dr. Ghobrial.
2. Digital pathology implementation and standardization.
3. Projects in liquid biopsy.

Working Session 2: Contemporary Applications of Liquid Biopsy

Moderator: Mickey Williams, PhD, Frederick National Lab

Dr. Williams opened the working session and introduced keynote speaker, Dr. Oxnard.

Keynote: ctDNA Analysis as a Tool for Precisions Drug Development
Geoff Oxnard, MD, Loxo

Dr. Oxnard began his keynote by highlighting the differences between clinical care biomarkers and drug development biomarkers. Clinical care biomarkers focus on patient diagnosis, therapy, and prognosis, while drug development biomarkers focus on drug potency, efficacy, and clinical trial design. Dr. Oxnard described three applications for ctDNA biomarkers in drug development.

First, ctDNA can be used for drug resistance profiling in the development of new drugs and drug combinations. He cautioned, however, that negative liquid biopsy also can indicate inadequate tumor DNA, and not all ctDNA mutations stem from tumor DNA.

Second, ctDNA is a highly dynamic analyte that can be used for response analysis to identify if an emergent drug is working within weeks. This can help identify an effective drug dose with real-time measurements during first-in-human studies. There is increasing standardization in how to use ctDNA response analysis in monitoring patient response.

Third, ctDNA-based trial enrollment is allowing liquid biopsy-positive patients to be part of a registrational cohort. However, plasma-negative participants may not be a true negative and therefore need additional discernment. In addition, the variability in ctDNA shedding across

cancers and patients has implications for the detection of biomarkers in ctDNA, and selecting for high levels of ctDNA may enrich for patients with a less favorable prognosis.

Dr. Oxnard described several challenges and future opportunities in liquid biopsy, including:

- Hematopoietic mutations can cause false positives in liquid biopsies.
- MRD-based drug development has encountered major challenges.
- There are other fluid types that can provide ctDNA for drug development investigations.
- There is a need for surrogates of gene expression across cancer types.

Discussion

During a brief question and answer session, Dr. Oxnard made the following points:

- Cancers can acquire alterations that ctDNA testing will detect, but use of ctDNA testing for this purpose requires the availability of a therapy for the liquid biopsy to have an impact.
- Precision testing and precision therapies are nuanced, but ctDNA testing can contribute to the nimbleness of sequential therapies versus bet-hedging.
- Liquid biopsy standards for benchmarking would enable progress.
- Canine trials and veterinary liquid biopsy approaches are a creative approach, but the quality of those technologies may be different than those used in human investigations.
- The use of ctDNA for multi-cancer detection is too vague for drug development with the current design of the technology. However, bridging the gap from a detection test to a profiling test could support drug development.

Liquid Biopsy Today

Moderator: Mickey Williams, PhD, NCI

Liquid Biopsy Today: FDA Perspective

Paz Vellanki, MD, PhD, FDA

Dr. Paz Vellanki is a medical oncologist as well as a thoracic and head and neck cancer team lead at FDA. She focused on uses of liquid biopsy for drug development and highlighted four uses of ctDNA, including:

1. ctDNA for patient selection – Liquid biopsy can identify a molecular alteration, detect early-stage disease, or monitor clonal evolution in the metastatic setting.

2. ctDNA MRD in clinical trials for early-stage disease – The FDA has published industry guidance on use of ctDNA for early-stage solid tumor drug development, which in part discusses trial designs for early-stage tumors and ctDNA MRD for patient enrichment. Trial designs include patient stratification by ctDNA status, enrichment with enrollment based on ctDNA positivity, and adaptive enrichment in which the ctDNA negative arm is closed at an interim analysis.

3. ctDNA as a marker of treatment response for drug development – In early-stage disease, there is interest in looking at ctDNA molecular response in correlation with pathological complete response after neo-adjuvant therapy. In metastatic disease, there is potential to complement RECIST assessments for stable disease and assess pseudoprogression versus

real progression. In all stages, ctDNA can be used for dose optimization, to determine therapy duration, and for escalation and de-escalation of therapy.

4. ctDNA as a potential early endpoint – Additional robust data, such as correlation with long term outcomes, is needed to develop ctDNA as an endpoint for regulatory use. Meta-analysis could be used to validate ctDNA as an early endpoint.

There are several ctDNA testing considerations related to sampling and assays. Sampling considerations include following standardized protocols, pre-specifying timing of ctDNA testing, and considering variation in ctDNA shedding. Assay considerations include holistic validation, assay cutoffs, and predictive value requirements based on escalation versus de-escalation trials.

Dr. Vellanki encouraged the committee to think about how ctDNA can be most impactful, what data are already available, and how data can be harmonized in prospective trials. She listed several important considerations for harmonizing data, including ctDNA assessment timing, assay harmonization, minimum performance specifications, and reference material development.

Discussion
In response to a question about the venue to discuss ctDNA use in drug development, Dr. Vellanki noted FDA is actively engaged in collaborative efforts, including with Friends of Cancer Research. FDA plans to host round table discussions and would like to include NCI in those discussions.

Following a question about the ctMoniTR Project that evaluated whether changes in ctDNA reflect response to cancer treatment, Dr. Vellanki reiterated that FDA is not ready to use ctDNA as a regulatory input, and a brief conversation concluded that ctMoniTR was a promising proof of principle study and first step.

Development of FNIH Quality Control Materials (QCM) for ctDNA
Christopher Karlovich, PhD, Frederick National Lab

Dr. Karlovich introduced the FNIH quality control materials for ctDNA project. The project stemmed from a 2019 study showing discordance among four commercial ctDNA vendors when they analyzed a variant allele fraction (VAF) below 1% in replicate plasma samples. A standardized set of quality control materials at known VAF would facilitate comparisons between assays. The project was divided into three phases:

- Phase 1: Performance evaluation to determine suitability of a lab's assay performance
- Phase 2: Functional characterization to demonstrate performance comparable to ctDNA
- Phase 3: A real-world clinical pilot in 11 commercial and academic laboratories

In Phase 1, three manufacturers provided a pool of common variants at varying allele frequencies in synthetic plasma. The Molecular Characterization Laboratory (MoCha) performed cfDNA extraction, aliquoted the samples, and gave them to NIST, MoCha, AstraZeneca, and Belfer/DFCI. Among those laboratories, there were three NGS assays and 2 ddPCR assays. Phase 1 provided insights into performance of both QCM and assays. Dr. Karlovich highlighted several findings, including discordance among manufacturers and an assay with large replicate imprecision.

In Phase 2, pharma donors supplied clinical samples to compare FNIH QCM performance. There were clear differences in performance between the clinical samples and QCM dilutions when using Hybrid capture NGS versus Amplicon NGS, and these same biases were observed in Phase 3 of the study. When the dilutions were near the limit of detection, three QCMs performed about as expected, but one majorly deviated from expectation.

Several lessons were learned during the study, including:

- Studies like this would be challenging for rare clinical specimens.
- Use of cancer patient blood as a source of cfDNA diluent would not have been feasible.
- Laboratories and manufacturers were sensitive to negative results.
- Engaging the FDA was helpful and led to study design modifications.

Discussion

In response to a question about specimen handling following shipment to each laboratory, Dr. Karlovich replied that the synthetic specimens contain encapsulated or preserved cfDNA. For the clinical specimens, they tried to track and control for those handling variables. He added that a sizing profile can identify if there has been degradation and genomic DNA contamination.

Measurable Residual Disease (MRD) in Acute Myeloid Leukemia (AML)

Christopher Hourigan, MD, PhD, NHLBI

The MRD in AML project is an active FNIH project in its second year. The project has two academic co-principal investigators and 20 industry partners. AML is a rare, fatal blood cancer diagnosed in approximately 20,000 Americans every year. Combination chemotherapy is effective in achieving an initial complete response, and having no evidence of MRD is associated with better survival. However, the genetic heterogeneity of AML presents challenges for finding the optimal MRD monitoring strategy, which could change the course of treatment. Potential use cases for AML MRD testing include quantification of therapy efficacy, early relapse detection, therapeutic assignment, patient selection for trials, and as a surrogate endpoint for regulatory approval. The European LeukemiaNet (ELN) produces evidence-based guidelines for clinical use of AML MRD testing.

The persistence of the most common mutations in AML prior to transplant is strongly associated with increased relapse and decreased OS after a transplant. An MRD test for negativity before stem cell transplant would improve transplant outcomes, but there are barriers to widespread clinical adoption of the technology.

The FNIH MRD in AML project aims to establish and validate new methods for detecting MRD in AML, including a library of reference standards and evidence of clinical utility. The project is organized into four groups that work in parallel, including:

1. Standards – Enables methods comparisons and multi-site testing
2. Methods – Generates evidence for specific contexts of use
3. Retrospective – Focuses on molecular methodologies
4. Prospective and Regulatory

Not all targets for measuring residual disease stratify patients. The group is working to identify clinically appropriate targets and develop generalizable principals on the best way to detect them. Specifically, they are testing the hypothesis that patient personalized testing is superior to fixed content assays and developing a framework for harmonized AML MRD testing in a prospective, nationwide study of 1,000 patients. In addition, MRD testing may provide information about quantity and quality of residual subclinical leukemia and inform subsequent therapies. This approach may be tested in the NCI myeloMATCH trial. Dr. Hourigan emphasized moving the field forward will require collaboration for robust standards, harmonized validated assays, and evidence of clinical utility.

Discussion

Dr. Mike Espey asked if the FNIH MRD in AML project results would inform treating AML in infants. Dr. Hourigan replied that the pattern of mutations is different, but there is some overlap. The structural variants are amenable to this type of analysis, and a measure in pediatrics is planned for the next iteration of the project.

Dr. Oxnard asked if there are people who have MRD but do not have recurrent AML. Dr. Hourigan replied sometimes bone marrow transplants work, but there are no predictive biomarkers. The next stage for measure is collaborating with immunologists to identify why a transplant is successful.

Asked about the goal of the initiative, Dr. Hourigan replied it is to be comprehensive in capturing all variants as well as standardizing the variants that have already been identified.

Liquid Biopsy Tomorrow

Exploring the Potential of Epigenomic Liquid Biopsies to Impact Cancer Care
Jacob Berchuck, MD, Dana Farber Cancer Institute

Dr. Jacob Berchuck is a practicing medical oncologist and translational researcher who studies novel epigenomic liquid biopsy tools that can be leveraged to impact patient care. He began with an introduction to the clinical relevance of epigenomic profiling. Epigenomics are the way in which cells regulate which genes are turned on and off in a coordinated manner, and epigenomic analysis through liquid biopsies provides a non-invasive method to look at factors that regulate gene expression and ultimately protein expression. Epigenetic modifications that can be measured in cfDNA include DNA methylation, nucleosome positioning, histone modification, fragmentation patterns, and chromatin accessibility. Dr. Berchuck briefly described the benefits and limitations of each epigenetic modification for liquid biopsy.

DNA methylation is tissue- and tumor-specific and can be used for multi-cancer early detection and monitoring MRD. DNA methylation can also detect clinically actionable tumor subtypes. Strengths of DNA methylation include high stability and clear clinical utility, but it has a weaker association with gene expression and is less dynamic than other epigenomic features of cfDNA.

Histone modifications provide a read out for gene activation and gene repression. For example, cfDNA histone profiling can detect drug target expression and dynamic changes in tumor biology.

Chromatin accessibility relies on the concept that the transcription start site of genes will be in an open configuration when being transcribed and closed when repressed. Open DNA is unprotected and disappears when it enters the blood stream, while DNA wrapped around a nucleosome will remain protected. Therefore, a drop in signal in cfDNA indicates an open configuration and a non-invasive surrogate of gene expression. Fragmentomics relies on this concept of DNA protection, and fragmentation profiles associate with tumor cell of origin and gene expression.

As new drugs that target cancer-specific cell surface proteins are developed, there is an opportunity to develop non-invasive epigenomic cfDNA-based biomarkers to enrich for patients who will benefit from these emerging targeted therapies. Several questions remain, including which circulating analytes are suitable for non-invasive assessment of gene expression, and what are the test characteristics of the commercial products for measuring gene expression.

Discussion

In response to a question about the ability for epigenetic-based measures to differentiate between HER2 1+, 2+, and 3+ when there are variations in tumor shed, Dr. Berchuck replied that the tools are still being developed to account for the extent of tumor shed, but he believes it will be possible.

In response to a question about the feasibility of detecting a small loss in signal based on chromatin accessibility, Dr. Berchuck replied that the optimal tool for those measurements is an open question, but preliminary data suggests integrating multiple epigenetic features is helpful in that they each provide non-overlapping information. However, he agreed looking for loss of signal with low tumor shed will be challenging.

Dr. Espey asked if comorbidities in cancer cofound a good signal. Dr. Berchuck replied that they are taking a biologically informed approach by identifying signal and background in white blood cells and plasma from healthy controls and limiting their analysis to areas that avoid that background. He noted they need to explore that area more to avoid false positives.

From Clinical Outcome to Clinical Assay: A Study of Blood-Based, 3D Genomic Biomarkers for the Checkpoint Inhibitor Response Test (CiRT)

Sasha Akoulitchev, PhD, Oxford Biodynamics

Dr. Akoulitchev introduced the concept that DNA packaging forms a specific three-dimensional (3D) architecture that serves as a footprint for active genome network regulation. When a cell divides, it produces the same 3D structures. Dr. Akoulitchev stated that the 3D genome represents the molecular regulatory network underlying cellular phenotypes or clinical outcomes.

Exosome signaling and extracellular vesicles spread epigenetic information from primary sites. This is an active mechanism, with about 10 billion vesicles from different cells in 1 mL of blood. Cells that accept the vesicle signaling switch their epigenetic status and 3D architecture to synchronize with the primary cell. Dr. Akoulitchev is interested in reading the systemic profile of 3D architecture in relation to clinical outcomes.

Next, Dr. Akoulitchev described the basis for the EpiSwitch 3D Platform. There are about 1.2 million anchoring sites in the genome that interact to form a 3D architecture in reaction to signaling

and phenotypic changes. Agilent arrays allow monitoring of the top one million interactions that might occur in a particular phenotype. creating a footprint representing the molecular complexity for a clinical condition. These footprints can be compared in those who respond or do not respond to a drug therapy. Through a FNIH Partnership for Accelerating Cancer Therapies (PACT) award, this modality was translated into a clinical test called the Checkpoint inhibitor Response Test (CiRT^{TM}). Over 700 patients have been tested, and several case studies were shown in which patients responded to checkpoint inhibition as predicted by the test. The logistics of the EpiSwitch CiRT includes sampling, extraction, amplification, analysis, and the production of a clinical report that predicts a high or low probability of response to immune checkpoint inhibition therapy.

Dr. Akoulitchev emphasized that CiRT is not a profile of the tumor but instead a read out of the dialogue between the host system and the tumor. Several trials predicting response to immune checkpoint inhibitors using EpiSwitch have been completed and made public. There are many other clinical applications for EpiSwitch in Oxford Biodynamic's product pipeline.

Discussion

In response to a question about the availability of hyper-progression samples, Dr. Akoulitchev replied they are interested in hyper-progressive samples, and have a prototype of the test. They are working with PACT, plus there is a subsidiary linked to the the Malaysian Health System with annotated samples.

Panel Discussion: **Objectives for CSC Liquid Biopsy Project in Development**

Sasha Akoulitchev, PhD, Oxford Biodynamics; Jacob Berchuck, MD, DFCI; Craig Cummings, PhD, Genentech; Zhiyong He, PhD, NIST; Christopher Hourigan, MD, PhD, NHLBI; Christopher Karlovich, PhD, Frederick National Lab; Diana Merino Vega, PhD, AstraZeneca

Dr. Williams invited the three new panelists to introduce themselves.

Dr. He recently joined NIST where they are developing reference materials for cfDNA methylation measurements as well as matched tumor and normal cell lines. Dr. Merino Vega is the Director of Translational Medicine, Cancer Biomarker Development, at AstraZeneca where she oversees the implementation and strategy of the use of ctDNA liquid biopsies across the portfolio. She has also worked with the BLOODPAC consortium and an MRD strategic working group. Dr. Cummings leads a computational biology group at Genentech. He represented the interests of the FNIH CSC Emerging Biomarker Technology Working Group (EBTx).

Dr. Cummings emphasized the importance of working together in a pre-competitive consortium. He introduced potential applications for blood-based gene expression assays in drug development, including molecular subtyping, identifying relapse and resistance mechanisms, reading out target gene expression, and assessing the state of the immune system. The EBTx working group is proposing the formation of a pre-competitive consortium to perform a robust and unbiased evaluation of multiple assays and vendors. The next step includes identifying investigators who are interested in designing and conducting this study, completing a full project plan, and recruiting industry partners. This working session is intended to inform development of this project.

The panelists discussed the challenge of knowing what the tumor fraction is and the potential to standardize the methodology of determining tumor fraction.

- Dr. Vega commented on harmonization efforts across tumor fraction measures. Referring to the ctMoniTR Project, she noted some studies used ddPCR while others used NGS, and new technologies that evaluate methylation signatures are introducing additional variables. She wondered if that would limit interpretation as studies are pooled for metanalysis.
- Dr. Oxnard expressed his doubt that harmonization is the effort to pursue, and suggested there should be a reference truth to serve as the benchmark for all efforts.
- Dr. Karlovich noted they have not found a way to accurately calculate tumor fraction below 3%, and the task is complicated. Some groups are using the maximum somatic allele frequency or mutant molecules per mL as a surrogate of the true tumor fraction. Dr. Karlovich suggested using Dr. He's matched cell lines to create dilutions as a truth in order to find an algorithm that performs best against that benchmark. He also suggested sourcing clinical samples with clinical outcome data to verify if a surrogate approach works.
- Dr. Williams agreed having a reference truth for tumor fractions in a well-developed reference set would be a powerful contribution from NIST. Dr. He replied they would need help and collaborators for access to hospitals and patients.
- Dr. Barrett agreed developing and validating reference samples for measuring tumor fractions would be pre-competitive, low cost, and have important applications. He emphasized the impact of a pharma-FNIH consortium for validating new technologies but highlighted the need for a common dataset and rules of engagement.
- Dr. Hourigan added that an open dialogue and model of trust is critical for the consortium, and that biotech companies themselves can benefit from these benchmarking bake-offs because they can make tangible changes to their products based on the findings. Dr. Vega added that rather than a gold standard assessment, it is possible to evaluate variability across assays instead.

Mr. Connors highlighted four points for the broader conversation:

1. Deciding what not to do can help target limited resources.
2. Moving tools that have emerged from the ctDNA project on to the next step could provide a parameter of focus.
3. Thinking about time as a parameter highlights the importance of developing materials over time and asking what will be learned over the next five years.
4. Coordinating globally with EMA and other regulatory agencies is important.

Panelists discussed the concept of ground truth and the best approach to establish ground truth:

- Dr. Palmer noted that two assay approaches can seem highly inconsistent, but assay agreement does not indicate ground truth. Instead, he suggested the best assay is the assay that separates the Kaplan Meyer curves.
- Dr. Berchuck agreed and wondered if associating assays with outcomes by using cohorts with annotated clinical outcomes following a targeted therapy could be a better reference approach.

- Dr. Karlovich indicated that ground truth exists in the tumor, but there is heterogeneity. He questioned how to know whether a predicted gene expression profile from an epigenomic assay represents the ground truth given intratumor tumor heterogeneity and tumor heterogeneity between primary and metastatic sites.
- Dr. Berchuck suggested using molecular imaging to read degree of marker heterogeneity.
- Dr. Cummings suggested multi-site analysis of primary tumors but noted the approach is expensive and laborious.
- Dr. Karlovich asked how quantitative the assay needs to be. Dr. Berchuck replied one approach is using tumor specific proteins, such as PSMA, for a dichotomous ground truth approach.
- Dr. Oxnard returned to the idea of constructing ground truth samples using dilutions with matched cell lines to have an established quantification of tumor content and known gene expression to test assays. He suggested this may be the most feasible experiment on tight timelines.
- Dr. Barrett agreed that tumor heterogeneity is a major problem for drug development and cancer care, and he emphasized that there is a lot to be learned from comparing technologies using the same samples. Tying in digital pathology, Dr. Barrett highlighted the value in computational studies of tissue that serves as the standard.
- Dr. Williams suggested pharma companies will have blood collected from targeted HER2 studies with annotated outcomes data, and that could serve as a good use case for assessing multiple assays alongside the development of computational algorithms from the tissue.

Panelists discussed whether the goal is to solve for tumor gene expression levels or tumor protein levels, and if those read outs could be discordant:

- Dr. Berchuck defines genes expression in this use case as the active expression of a gene that results in a predictive biomarker of response to a targeted therapy. He noted that studies have not been completed to determine if the non-invasive surrogates of gene expression correlate with RNA and protein levels.
- Dr. Hourigan encouraged the committee to have the end in mind to identify the appropriate sample sets. He noted he has observed discordance between protein level expression and single-cell partitioning of gene expression in the AML cancer model.

The panelists discussed if the goal for the CSC liquid biopsy project should be focused on early disease with low tumor burden or in the advanced disease setting with higher tumor burden. Dr. Berchuck suggested the advanced disease setting where there are targeted therapies in development and high tumor shed. He added that defining the limit of detection for these gene expression inference technologies could be an important outcome of a CSC effort in this space.

Dr. Vega defined the end goal as the ability to select a population likely to respond to a personalized therapy, and she suggested using all tools necessary to identify that population. Rather than have liquid biopsy replace IHC, she suggested that one may work better than the other in specific cases, and the project should fit within the global goal to develop precise biomarkers that identify a

population who are likely to respond to precision medicine. Dr. He encouraged the committee to also consider mRNA as the read out for gene expression in liquid biopsy.

As therapy moves into earlier stages of disease, another use case for liquid biopsy is to identify patients who have a high likelihood of relapse following conventional treatments. Panelists discussed identifying MRD in ctDNA with liquid biopsy.

- Dr. Karlovich noted that currently the negative predictive value is not sensitive enough. He expressed his support for a bake-off of different approaches, and emphasized the need for reference materials that work with each of the different types of approaches. He added that histology will also be a factor, as different cancers will have different footprints.
- Dr. Oxnard argued the MRD space lacks the feasibility to establish a ground truth and therefore designing reference materials for MRD does not meet the criteria of being an achievable goal.
- Dr. Vellanki commented that the FDA guidance that they wrote is for early-stage solid tumor drug development because there is a perceived need in that space. She highlighted three questions for the committee to consider:
    1. Where is the need?
    2. What is feasible?
    3. Where should the committee focus its efforts?

Dr. Barrett asked what problem the committee is trying to solve and listed several potential issues of focus, including MRD, surrogate endpoints, and gene expression. He emphasized that the discussion spanned different use cases but also different technologies, including liquid biopsy of ctDNA, digital pathology, proteomics, and single-cell sequencing. Identifying the focus and partners to bring in still needs to be defined.

- Dr. Cummings replied that the focus is not MRD, as other groups are already focused there. He explained this project is about a new set of technology that will provide a read out on what is happening in the tumor, as opposed to whether cancer is present or not.
- Dr. Hourigan suggested a circulating tumor, such as any type of leukemia, as an easy positive-control use case for profiling actual expression and collecting cfDNA fractions.
- Dr. Berchuck suggested selection of a particular protein (e.g., PSMA, HER2) for a use case with a controlled dilution and the incorporation of a molecular imaging modality. The focused question could provide opportunity for a bake-off to observe results of several assays that use distinct methods for inferring gene expression in the setting of a good enough truth.
- Dr. Vega highlighted the possibility of becoming paralyzed with options and asked if there is any objection to focusing on a specific cancer, like prostate cancer.
- Dr. Oxnard replied a project in heme malignancies is intriguing and potentially a feasible project, but he is not ready to say that prostate cancer should be the focus because of the unknown surrounding timeframes and reference truth. He emphasized the most important point is designing an experiment with a reference truth and a corresponding assay.

The panelists discussed the two new assays presented during the "Liquid Biopsies Tomorrow" working session. Specifically, they were asked to comment on whether the two assays are competing, complementary, or different:

- Dr. Akoulitchev confirmed that EpiSwitch and the epigenetic modifications discussed by Dr. Berchuck do not measure the same features. He emphasized the importance of measurements that can accurately predict clinical outcomes with no ambiguity, and that many molecular features may not contribute to outcomes.
- Dr. Barrett asked if EpiSwitch assesses tumor intrinsic changes or changes within the biosystem. Dr. Akoulitchev replied they indirectly pick up the re-setting of the immune cells involved in infiltration. In most cases, but not always, innate and adaptive immune cells that come to the site of the pathology are the frontline of cells that they use as a readout.
- Dr. Barrett expressed the benefit of exploring the space beyond the CTCs, whether immune effects or physiology of the system. He indicated NCI has an interest in systems biology.
- Dr. Berchuck noted the methylation, histone modification, fragmentomics, and chromatin accessibility measures are tumor-derived fragments. While they are a surrogate for gene expression, they are tumor intrinsic features that are dynamic. He highlighted the potential to prioritize targeted therapy over another systemic therapy based on a readout of tumor intrinsic biology at a specific point in time.
- Dr. Barrett asked if it is possible to have signals in the epigenetic modification measurements that represent changes in the immune cells subtracting from the tumor intrinsic signal.
- Dr. Berchuck agreed there is opportunity in that idea. He noted they have been careful to take a biologically-informed approach for their proof of principial and believes it possible to design an appropriate experiment for a bake-off effort with a focused question and some relative ground truth about expression. He added the field is still a few years away from having the computational skills for AI-based discovery using genome-wide data.
- Dr. Barrett suggested rare cancers as an alternative indication for the liquid biopsy effort.

Concluding Remarks

Dr. Kelloff thanked all participants for joining the symposium and expressed his excitement for the upcoming work over the next year.