

2023 SCIENTIFIC REPORT

Regulatory Advancements for Patients



FRIENDS
of CANCER
RESEARCH

3 INTRODUCTION

15 PATIENT-FOCUSED DRUG DEVELOPMENT: ADVANCING PATIENT-CENTERED TECHNOLOGIES AND TRIAL DESIGNS

- 16 An Analysis of Dosing-Related Postmarketing Requirements for Novel Oncology Drugs Approved by the U.S. Food and Drug Administration, 2012-2022
- 21 Interpreting Data from Dose-Finding Studies in Early Phase Oncology Trials to Determine the Optimal Dose
- 37 [A Study Design to Harmonize Patient-Reported Outcomes Across Data Sets*](#)

45 REAL-WORLD EVIDENCE: LEVERAGING RWD FOR INSIGHTS ON REAL-WORLD RESPONSE

- 46 Real-World Response Endpoints in Patients with mNSCLC Treated with Chemotherapy Across Real-World Datasets
- 47 Considerations for Leveraging Real-World Endpoints in Oncology Drug Development

55 INNOVATIVE DRUG DEVELOPMENT: INSIGHTS FOR ADVANCING ONCOLOGY TRIALS AND THERAPIES

- 56 Maximizing Data from Academic-Led Studies for Regulatory Decision-Making
- 64 Accelerating the Development of Engineered Cellular Therapies
- 80 Novel Approach to Accelerate Lung Cancer Research: Lung-MAP and the Potential of Public-Private Partnerships
- 84 Incorporating Pragmatic Elements in Trial Designs to Enhance Oncology Randomized Clinical Trials
- 101 Impact of COVID-19 Pandemic Mitigation strategies on industry and NCI Cancer Treatment Trials
- 102 [Sponsor Perspectives on the Impact of the COVID-19 Pandemic on Interventional Cancer Clinical Trial Protocols and Data Quality*](#)

115 COMPLEX BIOMARKERS: ALIGNING BEST PRACTICES TO SUPPORT FUTURE UTILIZATION

- 116 An Analysis of 13 Independently Performed Assays to Measure Homologous Recombination Deficiency Using 90 Freshly Extracted High Grade Serous Ovarian Tumors: Findings from the Friends of Cancer Research HRD Harmonization Project
- 117 Supporting the Application of Computational Pathology in Oncology
- 144 Establishing Evidence: New Advancements Using ctDNA
- 146 Changes in ctDNA Levels as an Early Indicator of Outcomes in Advanced NSCLC Treated with TKI: Initial Findings from a Retrospective Aggregate Analysis of 8 Clinical Trials

**The journal's top downloaded publications of 2023*

Introduction

In 2023, Friends of Cancer Research (*Friends*) played a key role in shaping and informing the landscape of oncology drug development and regulatory policy to bring advancements in treatment to patients through collaborative and innovative initiatives. Serving as a bridge for scientists, advocates, experts, and patients, *Friends* leverages partnerships and comprehensive research efforts to address critical challenges impacting oncology drug development and patient care.

Friends accomplished several significant milestones in 2023, including data readouts from the ctDNA to Monitor Treatment Response (ctMoniTR) Project evaluating the use of ctDNA as an early endpoint in oncology drug development and the Real-world Evidence (RWE) Pilot Projects (highlighted in our Project Spotlight on page 8) exploring the use of endpoints captured in real-world data. *Friends*' commitment to generating novel data to support regulatory policy is exemplified through these efforts and our other research partnerships, including the Digital Pathology Project and Homologous Recombination Deficiency (HRD) Harmonization Project.

The data developed from these partnerships, along with the outputs of our working groups, roundtables, and policy research, constitute the core content of this Scientific Report and are helping generate novel insights and support ongoing policy discussions. This report aims to serve as a resource for stakeholders in drug development, regulatory policy, and advocacy, by offering insights, solutions, and evidence-based strategies developed through collaborative research.

The 2023 Scientific Report captures the full text of our white papers and publications focused on several themes:

1 PATIENT-FOCUSED DRUG DEVELOPMENT
Advancing Patient-Centered Technologies and Trial Designs

2 REAL-WORLD EVIDENCE
Leveraging RWD For Insights on Real-World Response

3 INNOVATIVE DRUG DEVELOPMENT
Insights for Advancing Oncology Trials and Therapies

4 COMPLEX BIOMARKERS
Aligning Best Practices to Support Future Utilization

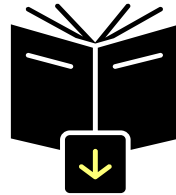
2023 By the Numbers



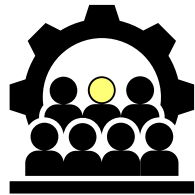
5,500+
meeting
attendees



roundtables
& public
meetings



9,000+
publication
downloads



500+
working group
participants

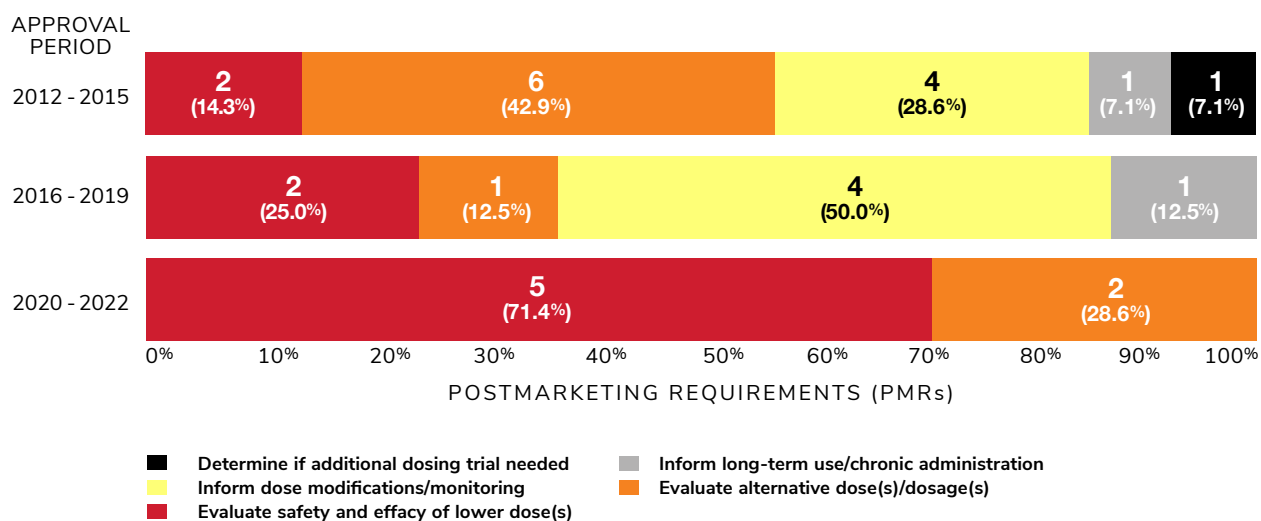


white papers,
abstracts, posters,
& publications

Patient-Focused Drug Development: Advancing Patient-Centered Technologies and Trial Designs

The FDA Oncology Center of Excellence continues to reiterate the importance of patient-centered approaches to dose optimization through its Project Optimus initiative and the release of a draft guidance document entitled “Optimizing the Dosage of Human Prescription Drugs and Biological Products for the Treatment of Oncologic Diseases.” Assessing dosing throughout oncology drug development is a complex process that requires consideration of a range of data including pharmacokinetic, pharmacodynamic, safety, efficacy, and tolerability data.

In 2023, recognizing this complexity and the need for patient-centered approaches to dose optimization aligned with FDA guidance, Friends worked with stakeholders to assess current challenges and propose patient-centered solutions for dose optimization. An analysis of dosing-related postmarketing activities in the last decade published in *Clinical Cancer Research* identified opportunities to refine approaches to premarket dosing studies and support a timely selection of the optimal dose. In addition, Friends’ 2023 Annual Meeting featured a white paper and panel discussion that elaborated on opportunities to refine approaches to premarket dosing studies, including how to consider tolerability (i.e., the extent to which the side effects of a treatment affect the ability or desire of a patient to adhere to the dose or intensity of therapy) as part of the totality of evidence generated through early phase dose-finding studies. Further, discussions at the Annual Meeting highlighted how, in the age of electronic data capture, digital tools, such as mobile app-based data platforms for capturing electronic patient-reported outcomes (ePROs), can be leveraged for critical insights into how a patient is feeling and functioning between clinical visits to inform tolerability.



Dosing variation PMRs by type of information over time. FDA issued more PMRs directing sponsors to evaluate a dose lower than the one approved in the last 3 years (n = 5, 2020–2022) compared with the preceding 8-year period (n = 4, 2012–2019).

Collins G, McKelvey B, Andrews HS, Allen JD, Stewart MD. An Analysis of Dosing-Related Postmarketing Requirements for Novel Oncology Drugs Approved by the U.S. Food and Drug Administration, 2012-2022. *Clin Cancer Res*. Published online December 12, 2023. doi:10.1158/1078-0432.CCR-23-2268

ePRO tools can capture data in real-time, enable earlier detection of treatment-related adverse events (AE), and ultimately support improved management and patient outcomes. In 2023, *Friends* collaborated with stakeholders to align on a framework for assessing data generated by different ePRO tools used in real-world care and identifying key real-world data elements to help assess whether ePRO data supports improved patient outcomes. *Friends'* work over the past year will help ensure optimal use of ePRO tools in clinical care and clinical development, including for assessing patient experiences to inform tolerability, and ultimately support timely, patient-centered dose optimization.

GUIDANCE DOCUMENTS

- **Patient-Focused Drug Development: Incorporating Clinical Outcome Assessments Into Endpoints for Regulatory Decision-Making**, Draft Guidance, April 4, 2023
- **Submitting Patient-Reported Outcome Data in Cancer Clinical Trials**, Final Guidance, November 6, 2023
- **Optimizing the Dosage of Human Prescription Drugs and Biological Products for the Treatment of Oncologic Diseases**, Draft Guidance, January 18, 2023

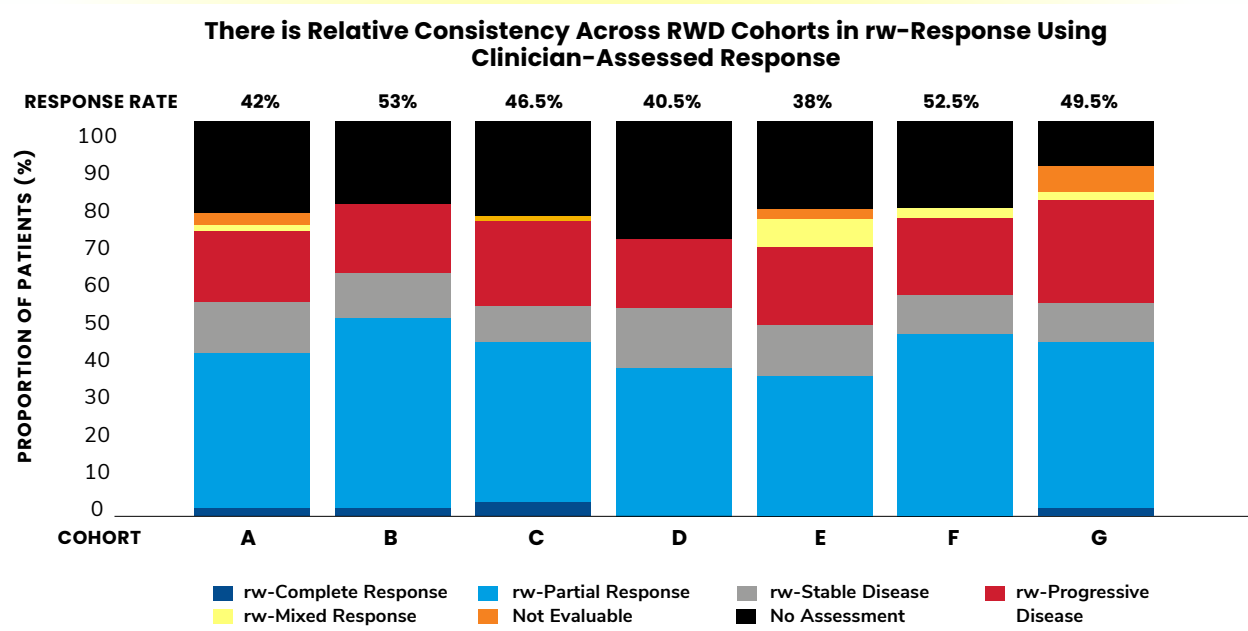
Real-World Evidence: Leveraging RWD for Insights on Real-World Response

Clinical trials provide the foundational evidence to support the safety and efficacy of new therapies. However, the patient population participating in a clinical trial is relatively small and may not reflect the broader patient population eligible to receive the therapy once it is approved. Therefore, it is crucial to leverage data generated through real-world care settings to further understand a therapy's safety and effectiveness in a larger and more representative patient population. Recent legislation and FDA guidance documents support the use of real-world data (RWD) to generate real-world evidence (RWE) and inform drug development and patient care. However, barriers to using RWD exist, as there is significant variability in the way RWD are reported within and across data sources. This inconsistency presents challenges in effectively using these data.



Zong J, Jiao X, Pan L, et al. Using real-world evidence (RWE) in regulatory decision making: A study of 6 oncology approvals with RWE included in the product label. *JCO*. 2023;41(16_suppl):6611-6611. doi:10.1200/JCO.2023.41.16_suppl.6611

Since 2017, *Friends* has facilitated collaborations (See Project Spotlight) to develop strategies and methodologies for aligning RWD. In 2023, *Friends* completed the real-world (rw)-Response Pilot, the latest research partnership in the RWE Portfolio, which proposes a framework for measuring response to treatment in RWD and assesses the consistency of the measure across RWD sources. This effort found relative consistency across data sources in an aligned patient population using clinician-stated response, demonstrating the potential for RWD sources to be used to evaluate drug effectiveness. In addition to a poster presentation at the ASCO Annual Meeting, *Friends* hosted a public meeting in 2023 to share the pilot results, and to provide considerations for leveraging rw-endpoints in oncology drug development, incorporating lessons learned from previous *Friends'* pilot projects. This work will continue to support the advancement of using RWD to generate robust evidence to support oncology drug development and patient care.



GUIDANCE DOCUMENTS

- **Real-World Data: Considerations for the Design and Conduct of Externally Controlled Trials for Drug and Biological Products**, Draft Guidance, February 2, 2023
- **Considerations for the Use of Real-World Data and Real-World Evidence To Support Regulatory Decision-Making for Drug and Biological Products**, Final Guidance, August 30, 2023
- **Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices**, Draft Guidance, December 19, 2023
- **Data Standards for Drug and Biological Product Submissions Containing Real-World Data**, Final Guidance, December 22, 2023
- **Real-World Data: Assessing Registries To Support Regulatory Decision-Making for Drug and Biological Products**, Final Guidance, December 22, 2023

Advancing RWE: Leveraging Data from Routine Care as a Research Tool

GOAL

Every day many patients are treated for cancer, and each patient's experience generates data on a treatment's effectiveness and safety to provide valuable insights and knowledge to advance cancer research and improve patient care. These data are also known as real-world data (RWD), and Friends of Cancer Research (*Friends*) leads critical work to advance methodology for using RWD as a research tool more effectively. *Friends'* unique collaborations lead to important insights into the accuracy and reliability of RWD, the ability to transform RWD into real-world evidence (RWE) related to the usage and potential benefits or risks of a treatment, and new policies to advance the use of this important data source to enhance cancer research and care for patients.

BACKGROUND

On average, fewer than 5% of patients with cancer receive treatment through a clinical trial. By leveraging RWD, the information gap between data generated from clinical trials and from routine care can be bridged. RWD are captured from sources such as insurance claims, electronic health records, and patient registries. There is growing recognition that these data sources, when analyzed appropriately, can generate RWE in broader patient populations to inform treatment effectiveness, safety, and patient outcomes.

Friends' portfolio of RWE projects informs and establishes methodology for using RWD to evaluate how treatments work in patients with cancer. Aligning best practices and frameworks for aggregating and analyzing RWD will ensure RWE is high quality and reliable for supporting oncology drug development, regulatory decision-making, and real-world use of products.

APPROACH

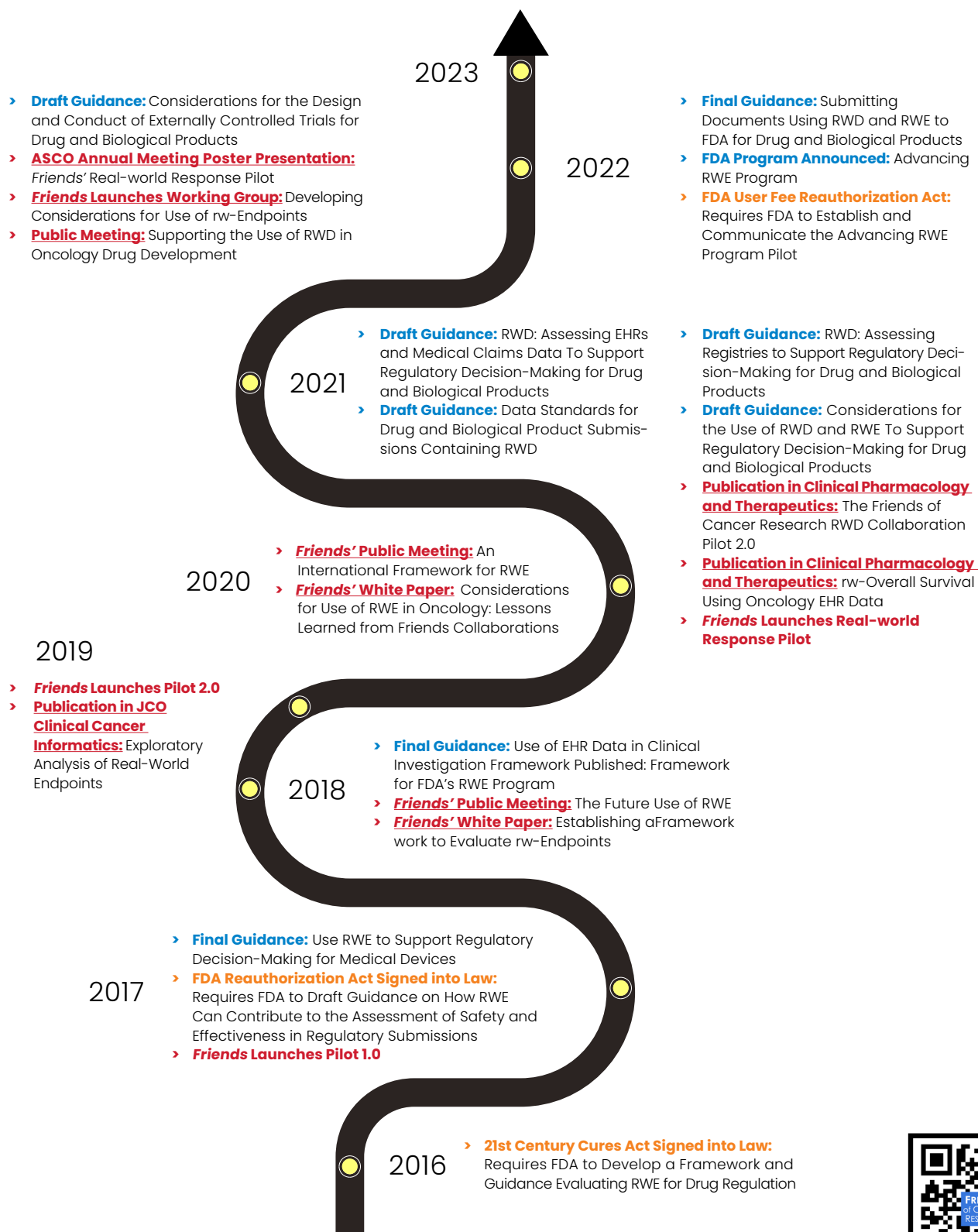
Since 2017, *Friends* facilitated collaboration among numerous RWD partners, pharmaceutical companies, government officials, and academic researchers to advance the understanding and applications of RWD and RWE in oncology:

- 1) RWE Pilot 1.0:** *Friends* developed a framework for operationalizing and validating real-world endpoints in advanced non-small cell lung cancer (aNSCLC). The pilot used a harmonized protocol and set of aligned definitions to identify similar patient populations across data sources and extract real-world endpoints.
- 2) RWE Pilot 2.0:** RWD partners showed that RWD can generate similar results to clinical trials when measuring treatment effect in patients with aNSCLC across RWD sources. The group developed a list of considerations for the design, conduct, and interpretation of RWD studies from different data sources.
- 3) rw-Response Pilot:** Project partners established a framework for measuring rw-response to treatment and assessed the consistency of the measure across RWD sources in patients with a NSCLC. Results showed that rw-response was relatively consistent across data sources in an aligned patient population using clinician-stated response.

NEXT STEPS

These multistakeholder partnerships support alignment on best practices, provide a venue to develop and test methodologies for aligning on and analyzing RWD, and help identify opportunities to proactively strategize on future use of RWD/E in oncology.

RWE Portfolio Development and Milestones

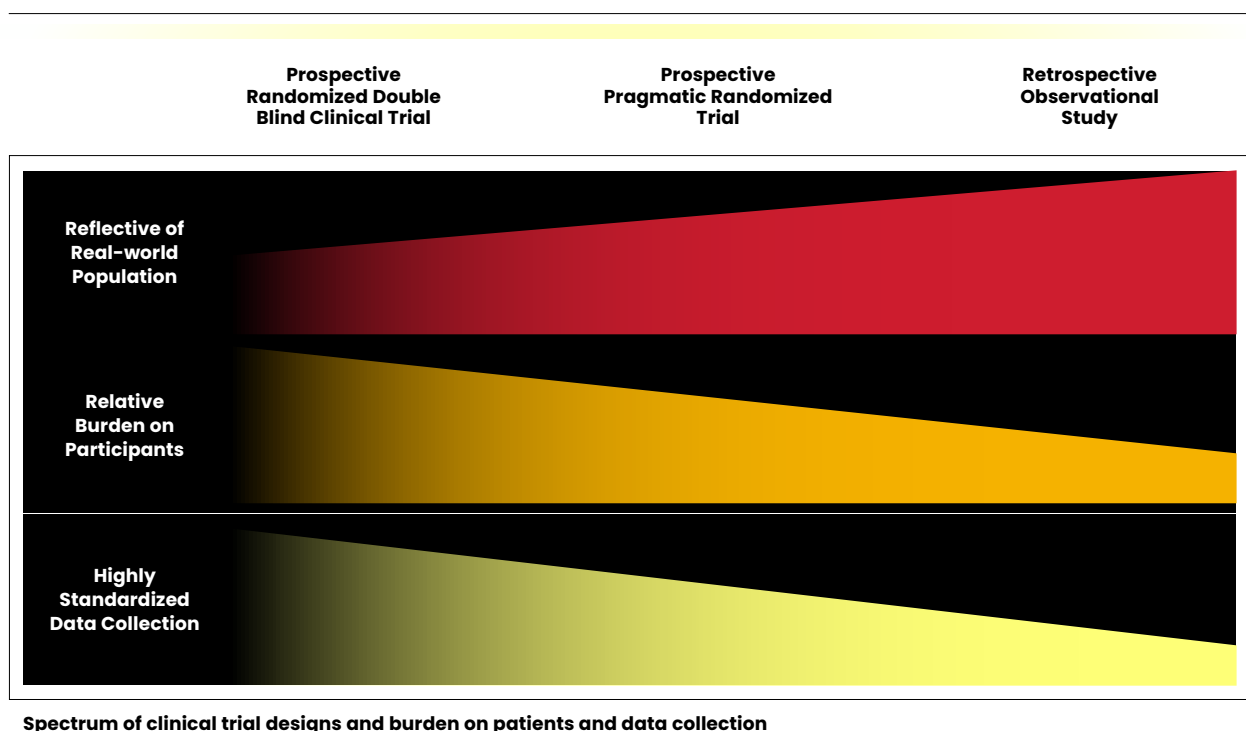


CATEGORIES ■ FRIENDS' RWE PORTFOLIO ■ FDA ■ LEGISLATION



Innovative Drug Development: Evaluating Lessons Learned to Optimize Development

Only about 5% of patients with cancer participate in clinical trials, which is driven by structural, clinical, attitudinal, and socioeconomic factors. It takes a multifactorial approach to improve patient enrollment and retention, including improving clinical trial designs by critically examining how to be more judicious about how data are collected and considered.



Clinical trialists should be thoughtful in their approach to data collection and extrapolation. Data from academic-led studies may be leveraged for regulatory decisions, but collaboration is necessary between those conducting the study and the drug sponsor to ensure data are collected in a manner that can support regulatory decision-making. *Friends' 2023 Annual Meeting* discussions identified opportunities to improve data preparation, including initiating conversations with FDA early and often. Early conversations with FDA are also important when considering data extrapolation or the use of data from other related clinical development programs to inform the development of other similar products. Cell therapy development is a key area where developers consider opportunities for data extrapolation to enable more efficient data collection and expedite development. In 2023, *Friends* hosted a public meeting about approaches to using data from other cell therapy products to inform the development of the next generation of cell therapies. As novel therapies are developed, it is critical to ensure regulatory paradigms keep pace with technological advances.

Additionally, clinical trial designs that reduce burdens on sites and patients are important. The Lung-MAP project is an umbrella trial in patients with non-small cell lung cancer that provides a case study for establishing and engaging in public-private partnerships. Findings from a Lung-MAP study spurred the Pragmatica-Lung Clinical Trial, a trial with pragmatic elements to reduce patient and site burden of data collection. At the *Friends' Annual Meeting* in 2023, stakeholders discussed opportunities for incorporating pragmatic elements in clinical trials more broadly to encourage broader patient participation. Additionally, findings from a study completed in collaboration with ASCO and published in the *JCO Oncology Practice* assessed the impact of the COVID-19 pandemic on clinical trial data collection, supporting the use of pragmatic and decentralized elements. Findings identified that flexibility in drug delivery and monitoring of therapy were not only feasible during the COVID-19 pandemic, but did not appear to impact data integrity, supporting their thoughtful incorporation into clinical trials.

Clinical trialists should be **thoughtful in their approach to data collection and extrapolation.**

As novel therapies are developed, it is critical to ensure regulatory **paradigms keep pace with technological advances.**

5% of patients with cancer

participate in clinical trials which is driven by **structural, clinical, attitudinal, and socioeconomic factors.**

GUIDANCE DOCUMENTS

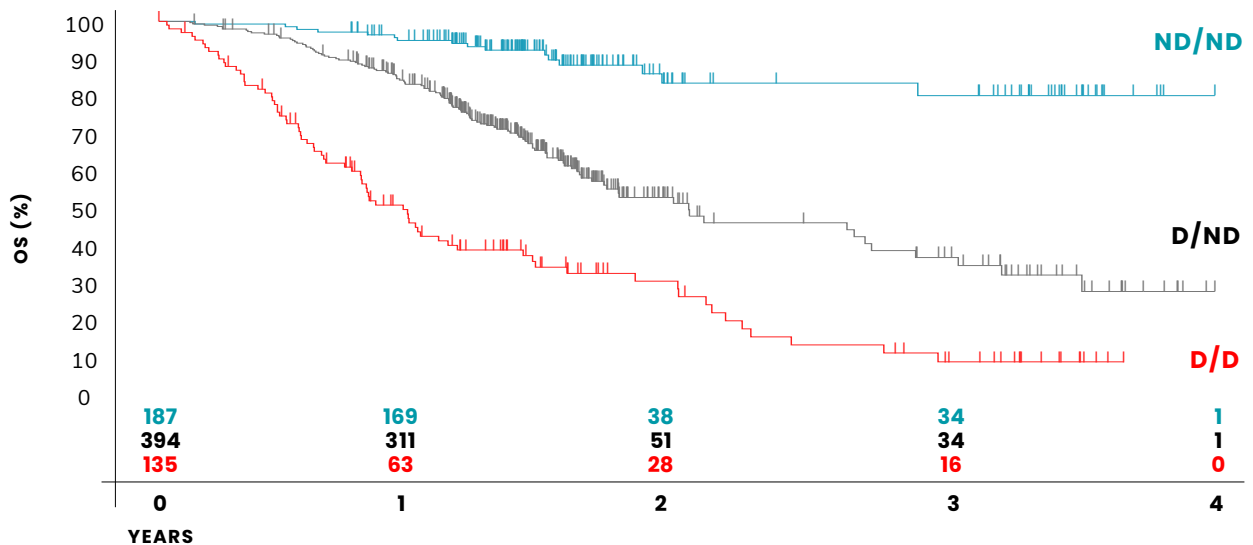
- **Clinical Trial Considerations to Support Accelerated Approval of Oncology Therapeutics**, Draft Guidance, March 27, 2023
- **Manufacturing Changes and Comparability for Human Cellular and Gene Therapy Products**, Draft Guidance, July 13, 2023
- **Postmarketing Approaches to Obtain Data on Under-Represented Populations in Clinical Trials**, Draft Guidance, August 10, 2023
- **Decentralized Clinical Trials for Drugs, Biological Products, and Devices**, Draft Guidance, May 1, 2023
- **Master Protocols for Drug and Biological Product Development**, Draft Guidance, December 22, 2023

Complex Biomarkers: Aligning Best Practices to Support Future Utilization

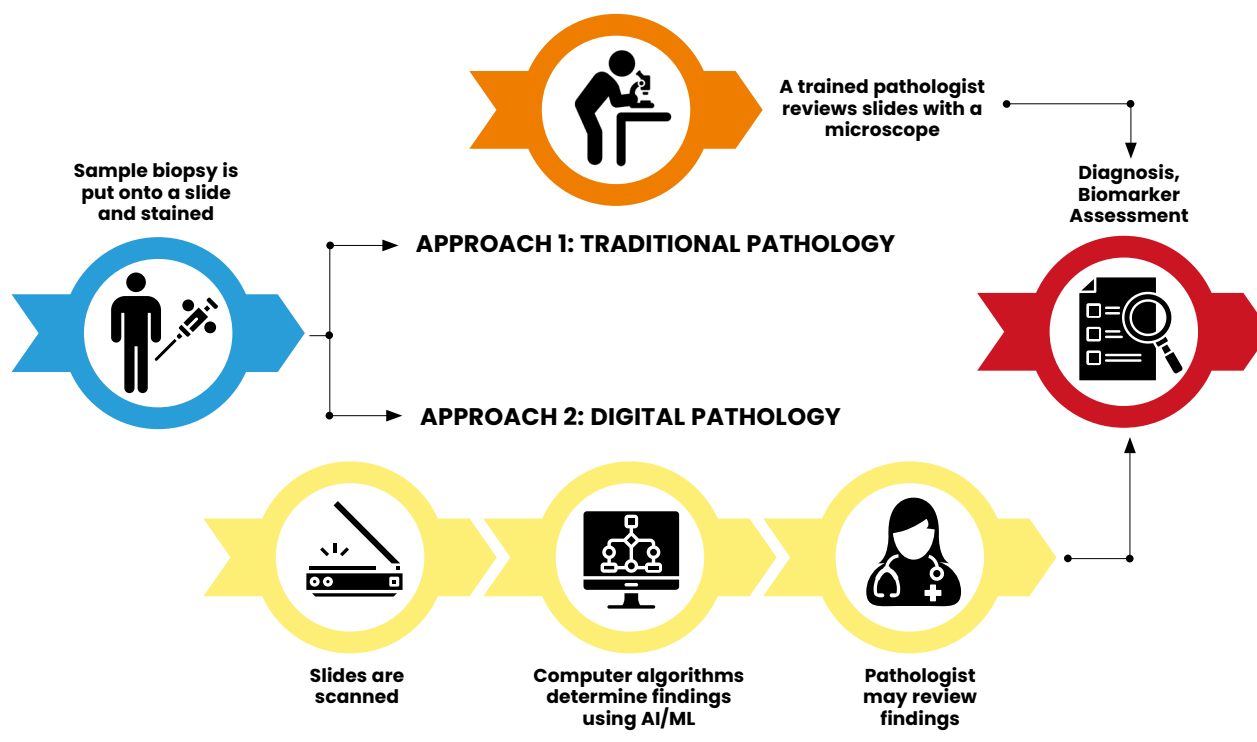
In oncology, physicians and patients use biomarker assessments to make treatment decisions, track how disease progresses, and assess patient prognosis. Diagnostic tests measure the abundance or presence of biomarkers, often in the tumor or the blood. As our understanding of the biology of cancer continues to improve and advance, so do the assays that measure various biomarkers.

Given the role diagnostic tests play in patient care, it is imperative that tests measure a biomarker accurately and precisely. However, given the current landscape of diagnostic test regulation and oversight, there can be uncertainty in the comparability of tests within the same intended uses. This may lead to inconsistencies in outputs and a lack of clarity on test and treatment decision-making. To support alignment, *Friends* established collaborative research partnerships with diagnostic developers, regulators, and academics to share datasets to assess variability in biomarker assessment across tests and determine opportunities for overcoming the differences. The homologous recombination deficiency (HRD) Harmonization Project compared outputs from 20 assays measuring HRD using ovarian cancer samples. In 2023, *Friends* completed this work with a presentation of initial findings at the AACR Special Conference in Cancer Research: Ovarian Cancer. In 2023, *Friends* also convened a working group to develop a landscape assessment of another tool that assesses complex biomarkers, digital and computational pathology platforms. These discussions set the stage for a new research partnership to launch in 2024.

Overall Survival by ctDNA Categories for Patients with aNSCLC Treated with TKI



Category	Baseline ctDNA Level	On Treatment ctDNA Level
ND/ND	Not Detected	Not Detected
D/ND	Detected	Not Detected
D/D	Detected	Detected




Another challenge is that the clinical utility of a biomarker may not always be clear. To understand whether circulating tumor DNA (ctDNA) levels are associated with long-term outcomes like overall survival, *Friends'* ctMoniTR Project combines patient-level data from multiple clinical trials. Combining and evaluating data from multiple trials creates a larger sample size, enables analyses not conducted in individual trials, and can improve precision in the estimates of effect of ctDNA change. In 2023, aggregating data from eight independently conducted clinical trials, an analysis of patients with advanced non-small cell lung cancer treated with a tyrosine kinase inhibitor was presented as a poster at the ASCO Annual Meeting. Additionally, these data and information describing ctDNA levels across cancer types and stages were presented during a public meeting hosted by *Friends* in July 2023.

GUIDANCE DOCUMENTS & POLICY PROPOSALS

- **Oncology Drug Products Used with Certain IVD Tests: Pilot Program, Final Guidance & Pilot Program**, June 21, 2023
- **Medical Devices - Laboratory Developed Tests**, Proposed Rule, September 29, 2023
- **Marketing Submission Recommendations for a Predetermined Change Control Plan for Artificial Intelligence/Machine Learning (AI/ML)-Enabled Device Software Functions**, Draft Guidance, April 3, 2023





Patient-Focused Drug Development: Advancing Patient-Centered Technologies and Trial Designs

An Analysis of Dosing-Related Postmarketing Requirements for Novel Oncology Drugs Approved by the U.S. Food and Drug Administration, 2012–2022

Grace Collins, Brittany McKelvey, Hillary S. Andrews, Jeff D. Allen, and Mark D. Stewart

ABSTRACT

The FDA's Oncology Center of Excellence's (OCE) launch of Project Optimus signals increased focus on dose optimization approaches in oncology drug development, particularly toward optimization in the premarket setting. Although sponsors continue to adapt premarket study designs and approaches to align with FDA's expectations for dose optimization, including consideration of the optimal dosage(s), there are still instances where questions remain at the time of approval about whether the approved doses or schedules are optimal. In these cases, FDA can exercise regulatory flexibility by issuing postmarketing requirements (PMR) and avoid

delaying patient access to promising therapies. This landscape analysis demonstrates that over the past decade (2012–2022), FDA frequently used PMRs to answer additional questions about dosing for novel oncology approvals. We found more than half of drugs (78/132, 59.1%) had a dosing PMR and observed a recent increase in PMRs intended to evaluate whether a lower dose could be more optimal. These results suggest there are opportunities to adapt premarket dose optimization strategies and leverage innovative development tools to ensure timely identification of the optimal dose.

Introduction

The FDA Oncology Center of Excellence's (OCE) launch of Project Optimus signals a shift in expectations for dose optimization approaches in oncology, particularly towards optimization in the premarket setting (1). Although sponsors continue to adapt premarket study designs and approaches to align with FDA's expectations for dose optimization, including consideration of the optimal dosage(s), there are still instances where questions remain at the time of approval about whether the approved doses or schedules are optimal. In these circumstances, FDA can use its authority to require sponsors to conduct additional dose optimization by issuing postmarketing requirements (PMR). A sponsor may also agree to a postmarketing commitment (PMC) to conduct additional dose optimization, but these are "studies or clinical trials the sponsor has agreed to conduct but are not required by statute or regulation" (2). PMRs are important tools, which allow the FDA to exercise regulatory flexibility and enable timely approval of potentially lifesaving drugs and biologics (collectively referred to herein as drugs) while additional studies are ongoing. This is particularly true in oncology, a disease area in which drugs are often approved on expedited timelines that speed access to innovative treatments for patients with life-threatening cancers who have exhausted all other treatment options.

Given the increased emphasis on the importance of adequate characterization of doses and schedules, we conducted a landscape analysis of dosing PMRs issued to novel oncology drugs approved

over the last decade (2012–2022). Previous research has broadly evaluated clinical pharmacology- and immunogenicity-related PMR/Cs and considered how factors such as the use of expedited programs [e.g., accelerated approval (AA)], special designations (e.g., orphan drug designation), and pivotal trial designs influence decisions to assign a PMR or PMC (3–6). These studies briefly acknowledged certain dosing PMR/Cs within the scope of their analyses but did not evaluate trends or characteristics of dosing PMR/Cs for novel oncology drugs. Our analysis provides a comprehensive review of dosing PMRs for oncology drugs to identify the types of dosing information the FDA requires sponsors to collect and how long it takes to complete these activities in the postmarketing setting. We focused our analysis on PMRs because FDA has authority to issue them and ensure they are completed (2). In addition, PMRs better reflect the types of dosing activities and information FDA views as critical to fulfilling statutory requirements that ensure safe and effective use. We also evaluated trends in dosing PMRs over time to assess the impact of the FDA's re-evaluation of the dose optimization and selection paradigm and associated policy related to dose optimization in oncology.

Materials and Methods

We identified a list of novel drugs approved to treat cancer by the FDA between January 1, 2012, and December 31, 2022. Novel drugs include original applications for drugs that have never been approved before. We focused on this group of drugs because they have no predicates or same in-class drugs, and therefore, no prior knowledge to rely on. Using the publicly available Drugs@FDA database and FDA's web page for products licensed by the Center for Biologics Evaluation and Research's (CBER) Office of Therapeutic Products (OTP), we compiled a list of PMRs included in the original approval letters for these drugs (7). Additional information collected from approval letters included PMR descriptions, statutes under which they were issued, and final report due dates.

We then identified PMRs intended to inform dosing by searching PMR descriptions for the keywords "dose," "dosage," and

Friends of Cancer Research, Washington, DC.

Corresponding Author: Mark D. Stewart, Science Policy, Friends of Cancer Research, Washington, DC 20036. E-mail: mstewart@focr.org

Clin Cancer Res 2023;XX:XX–XX

doi: 10.1158/1078-0432.CCR-23-2268

This open access article is distributed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) license.

©2023 The Authors; Published by the American Association for Cancer Research

Table 1. Characteristics of novel oncology drugs approved by the FDA (2012–2022).

	All drugs				Drugs with a dosing PMR			
	All years	2012–2015	2016–2019	2020–2022	All years	2012–2015	2016–2019	2020–2022
Total	132	42 (31.8)	44 (33.3)	46 (34.8)	78 (59.1)	26 (61.9)	23 (52.3)	29 (63)
Application type								
NDA	82 (62.1)	29 (35.4)	29 (35.4)	24 (29.3)	66 (80.5)	25 (86.2)	22 (75.9)	19 (79.2)
BLA	50 (37.9)	13 (26)	15 (30)	22 (44)	12 (24)	1 (7.7)	1 (6.7)	10 (45.5)
Approval type								
Accelerated approval	71 (53.8)	25 (35.2)	26 (36.6)	20 (28.2)	35 (49.3)	13 (52)	12 (46.2)	10 (50)
Regular approval	61 (46.2)	17 (27.9)	18 (29.5)	26 (42.6)	43 (70.5)	13 (76.5)	11 (61.1)	19 (73.1)
Drug class								
Molecular target inhibitors	68 (51.5)	23 (33.8)	27 (39.7)	18 (26.5)	56 (82.4)	20 (87)	21 (77.8)	15 (83.3)
Monoclonal antibody/ADCs	38 (28.8)	11 (28.9)	11 (28.9)	16 (42.1)	11 (28.9)	1 (9.1)	1 (9.1)	9 (56.3)
Chemotherapies ^a	8 (6.1)	5 (62.5)	—	3 (37.5)	7 (87.5)	4 (80)	—	3 (100)
Cell and gene therapies	8 (6.1)	1 (12.5)	2 (25)	5 (62.5)	—	—	—	—
Endocrine therapies/hormone antagonists and related agents	4 (3.0)	1 (25)	2 (50)	1 (25)	—	—	—	—
Radiopharmaceuticals	3 (2.3)	1 (33.3)	1 (33.3)	1 (33.3)	3 (100)	1 (100)	1 (100)	1 (100)
Other ^b	3 (2.3)	—	1 (33.3)	2 (66.7)	1 (33.3)	—	—	1 (50)
Disease setting								
Advanced	124 (93.9)	41 (33.1)	39 (31.5)	44 (35.5)	75 (60.5)	26 (63.4)	21 (53.8)	28 (63.6)
Both	4 (3)	1 (25)	2 (50)	1 (25)	1 (25)	—	—	1 (100)
Early stage	4 (3)	—	3 (75)	1 (25)	2 (50)	—	2 (66.7)	—

^aChemotherapies include three alkylating agents, two antimetabolites, one protein biosynthesis inhibitor, and two angiogenesis inhibitors.

^bOther includes two antineoplastic enzymes and one hypoxia-inducible factor (HIF) inhibitor. BLA, biologics license application; NDA, new drug application; ADC, antibody-drug conjugate.

“dosing.” Dosing PMR descriptions were reviewed and activities were categorized as: (i) “Extrinsic Factors,” which include evaluations of how extrinsic factors affect dosing such as drug interaction, drug-drug interaction, and food effect trials; (ii) “Intrinsic Factors,” which include evaluations of how intrinsic factors affect dosing such as dosing in patients with renal and hepatic impairment, pediatric populations, patients with a certain genetic marker not specified in the label, and evaluations of dosing based on body surface area or body weight; (iii) “Dose Variation” PMRs, including evaluations of dosing in a new combination, alternative regimens, levels, schedules, or infusion timelines, studies that informed dose modification and monitoring recommendations, and studies that otherwise compare doses or inform whether the approved dose(s) are optimal; and (iv) “Miscellaneous activities,” which include development of new formulation strengths, assessments of the QT interval (QT/QTc studies), long-term safety studies that do not explicitly inform dose modifications and monitoring, animal toxicology studies, and immunogenicity studies.

To understand factors influencing the types of PMR issued, we used FDA’s public databases to collect information on the approval pathway (AA vs. traditional approval), application type [new drug application (NDA) vs. biologic license application (BLA)], indicated cancer type, and disease setting (advanced vs. early stage) for each drug. We identified drug classes using the National Library of Medicine’s (NLM) RxClass database and the Kyoto Encyclopedia of Genes and Genomes (KEGG) Drug database.

Results

Between January 1, 2012, and December 31, 2022, the FDA approved 132 novel oncology drugs and we identified 376 PMRs for 112 of these novel drugs. Of the 376 PMRs, 43.9% (165/376) collected additional dosing information for 78 of the approved drugs (59.1%, 78/132).

Characteristics of drugs with a dosing PMR

Between 2012 and 2022, NDAs were more likely than BLAs to have a dosing PMR (80.5%, 66/82, vs. 24%, 12/50). The percentage of BLAs with a dosing PMR increased over time from 7.7% (1/13) of BLAs approved 2012 to 2015 to 45.5% (10/22) of BLAs approved 2020 to 2022. In contrast, the percentage of NDAs with a dosing PMR decreased slightly over time from 86.2% (25/29) of NDAs 2012 to 2015 to 79.2% (19/24) of NDAs 2020 to 2022 (Table 1).

Across most drug classes, the percentage of approvals with a dosing PMR increased or remained consistent over the last 10 years (Table 1). Although 82.4% (56/68) of molecular target inhibitors had a PMR to inform dosing, there was a slight decrease over time in the percent of drugs in this class with a dosing PMR (87%, 20/23 approved 2012–2015 vs. 83.3%, 15/18 approved 2020–2022). The drug classes with the most approvals assigned a dosing PMR were radiopharmaceuticals (100%, 3/3), chemotherapies (87.5%, 7/8), and molecular target inhibitors (82.4%, 56/68). Drugs classified as other (33.3%, 1/3) and mAbs/antibody–drug conjugates (ADC; 28.9%, 11/38) had the fewest drugs with a dosing PMR. Over time, the percentage of mAbs/ADCs with a dosing PMR increased from 9.1% (1/11) of drugs approved 2012 to 2015 to 56.3% (9/16) of drugs approved 2020 to 2022. Cell and gene therapies and endocrine therapies/hormone antagonists and related agents both had 0 drugs with a PMR to collect additional dosing information in the postmarketing setting (Table 1).

Characteristics of dosing PMRs

Most dosing PMRs (75.6%, 125/165) evaluated the impact of intrinsic factors such as renal/hepatic impairment, body weight, genetic markers, or extrinsic factors such as food effect and drug interactions (Table 2). In the past 3 years, there was an increase in the percentage of dosing related PMRs evaluating extrinsic factors (31.3%, 15/48 were issued during 2012–2015 compared with 52.1%, 25/48 issued during 2020–2022). PMRs focused on intrinsic factors had a median of 2.1 years to be completed (years from the approval date to

Table 2. PMRs by dosing category and type of information provided to inform dosing over time (2012–2022).

Dosing category	Type of information	Years of approval			
		All years	2012–2015	2016–2019	2020–2022
Intrinsic factors	Hepatic impairment	45 (58.4)	15 (33.3)	14 (31.1)	16 (35.6)
	Renal impairment	21 (27.3)	9 (42.9)	5 (23.8)	7 (33.3)
	Age (pediatric)	7 (9.1)	—	—	7 (100)
	Genetic subgroup	2 (2.6)	1 (50)	—	1 (50)
	Renal and hepatic impairment	1 (1.3)	—	1 (100)	—
	Low body weight	1 (1.3)	—	—	1 (100)
	Subtotal	77 (46.7)	25 (32.5)	20 (26)	32 (41.6)
Extrinsic factors	Drug interaction	42 (87.5)	11 (26.2)	7 (16.7)	24 (57.1)
	Drug–drug interaction	5 (10.4)	4 (80)	—	1 (20)
	Food effect	1 (2.1)	—	1 (100)	—
	Subtotal	48 (29.1)	15 (31.3)	8 (16.7)	25 (52.1)
Dosing variation	Evaluate safety and efficacy of lower dose(s)	9 (31)	2 (22.2)	2 (22.2)	5 (55.6)
	Evaluate alternative dose(s)/dosage(s)	9 (31)	6 (66.7)	1 (11.1)	2 (22.2)
	Inform dose modifications/monitoring	8 (27.6)	4 (50)	4 (50)	—
	Inform long-term use/chronic administration	2 (6.9)	1 (50)	1 (50)	—
	Determine if additional dosing trial needed	1 (3.5)	1 (100)	—	—
	Subtotal	29 (17.6)	14 (48.3)	8 (27.6)	7 (24.1)
Miscellaneous	Long-term follow-up	5 (45.5)	2 (40)	2 (40)	1 (20)
	QT/QTc assessment	4 (36.4)	3 (75)	—	1 (25)
	Animal toxicology study	2 (18.2)	1 (50)	—	1 (50)
	Subtotal	11 (6.7)	6 (36.4)	2 (18.2)	3 (27.3)
Total		165	60 (36.4)	38 (23)	67 (40.6)

the final report due date indicated in the approval letter) and those focused on extrinsic factors had a median of 1.9 years to be completed (Fig. 1).

Dose variation PMRs (17.6%, 29/165) evaluated lower doses (31%, 9/29) or alternative doses/dosages (31%, 9/29), informed dose modifications and monitoring (27.6%, 8/29), dosing for long-term/chronic use (6.9%, 2/29), and helped collect data to determine whether an additional trial would be needed to inform dose optimization (3.5%, 1/29). We also found that several drugs (8/132, 6%) had a PMR to evaluate the safety or efficacy of a lower dose. In the past 3 years, FDA requested five PMRs to evaluate lower doses for 4 of the 46 (8.7%) drugs approved. In contrast, there were only 4 PMRs to evaluate lower doses for 4 of the 86 (4.7%) drugs approved in the prior 8-year period (Fig. 2). Dose variation PMRs took a median of 4.5 years to be completed, with PMRs to inform dose modifications and monitoring and investigate lower dosing taking the greatest amount of time at a median of 6.2 years and 5.0 years, respectively (Fig. 1).

The remaining 11 miscellaneous dosing PMRs consisted of long-term follow-up studies to characterize safety ($n = 5$), QT/QTc assessments ($n = 4$), and 2 animal toxicology studies (Table 2). These took a median of 2.6 years to be completed (Fig. 1).

Discussion

For many oncology drugs, FDA uses PMRs as a tool to further inform safe and effective use of an approved drug, including the optimal dose(s). Traditionally, early-phase oncology clinical trials aimed to identify the maximum tolerated dose (MTD), a dose optimization strategy designed for cytotoxic chemotherapies with which increasing the dose is associated with increasing efficacy. Over the past decade, scientific advancements have led to more approvals of targeted therapies for which efficacy may plateau before reaching the MTD. As has previously been discussed, for

these therapies a lower dose can provide the same efficacy with improved safety and tolerability profiles for patients (8).

As our analysis showed, most novel oncology approvals over the past decade have been targeted inhibitors for which FDA continues to emphasize that identification of the MTD is no longer adequate justification for having optimized the dose (9). We also found that more than half of all oncology drugs approved in the last decade had a PMR to further inform dosing (Table 1). In addition, we observed an increase in the proportion of approvals for mAbs/ADCs over time and found the percentage of these drugs with a dosing PMR increased six-fold during the 2020 to 2022 period compared with the preceding approval periods (Table 1). A prior analysis of small molecules and ADCs for oncologic indications approved 2019 to 2021 showed use of the MTD paradigm persists in the premarket setting (10). This coincided with an increase in PMRs intended to evaluate lower or alternative dosing regimens during the past 3 years (2020–2022), compared with the preceding 8 years combined (2012–2019; Fig. 2).

Dosing PMRs designed to evaluate a lower dose had a median of 5 years to be completed after approval and evaluations of alternative doses/dosages had a median of 4.2 years (Fig. 1). During this time, there is a risk of patients being exposed to suboptimal doses. Trial design and analytical methods to support timely identification of the optimal dose other than the MTD approach, is paramount given the length of time it takes to evaluate lower and alternative dose(s). Recent Oncologic Drugs Advisory Committee (ODAC) meetings focusing on a certain class of targeted therapies, Phosphoinositide 3-kinase (PI3K) inhibitors, provide another example of challenges arising when pre-market dosing strategies fail to adequately optimize the dose and postmarketing trials designed to further inform dosing raise additional questions about safety and efficacy leading to withdrawal from the market (11).

Increased focus on dosages aligns with the OCE's recent efforts to reform approaches to dose optimization in oncology. In 2021, OCE launched Project Optimus, “an initiative to reform the dose

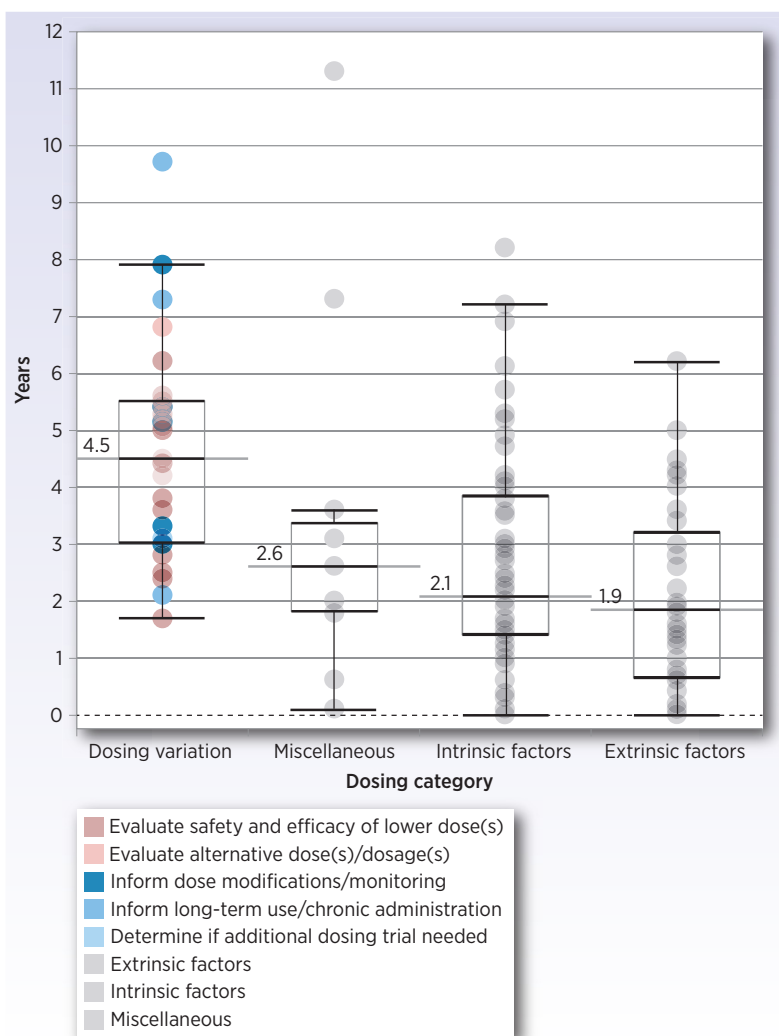


Figure 1. Median time from approval to final report due date for dosing PMRs. Dosing variation PMRs have the longest median time to be completed (4.5 years from date of approval to final report due date).

optimization and dose selection paradigm in oncology drug development” (1, 8). As more of these targeted therapies are introduced, there will be a need to develop tailored dose optimization strategies that account for the nuances that exist between drugs and drug classes. As such, opportunities to adapt dosing strategies and identify appropriate flexibilities that enable timely identification of the optimal dose will be important.

Moving Forward

PMRs are important tools that enable FDA to exercise regulatory flexibility and facilitate timely access to promising therapies, particularly for patients living with cancer. For oncology drugs approved over the past decade, FDA has frequently used PMRs to gain additional information about the optimal dose. The push for dose optimization of oncology drugs in the premarket setting is not a new concept; however, this analysis provides timely insights on the types of dosing activities FDA has requested in the postmarketing setting over the last decade which could identify areas where additional dosing information could be collected in the premarket setting. In addition, the analysis dem-

onstrated certain dosing activities take longer to complete in the postmarketing setting than others. While PMRs remain an important tool for exercising regulatory flexibility, they may be more appropriate for dosing questions that can be efficiently answered. The dosing evaluations that take longer, such as the exploration of a range of lower doses, could be prioritized earlier in development to avoid exposing patients to potentially suboptimal doses. Leveraging scientific advances and innovative trial designs can help enhance dose optimization strategies and enable more efficient dosing studies in the premarket setting. For instance, the use of novel biomarkers, such as circulating tumor DNA (ctDNA), also holds promise by providing less invasive and real-time insights into tumor dynamics and treatment responses associated with different dosages. The 3+3 trial design is frequently used in early phase dose escalation studies for oncology drugs; however, other, more flexible trial designs could enable more dynamic adjustments to dosing regimens based on accumulating trial data and allow for quicker identification of the most effective doses (12). As we continue to advance our approaches for optimizing dosage selection in oncology drug development, we should do so with the goal of bringing safer and more tolerable drugs to patients.

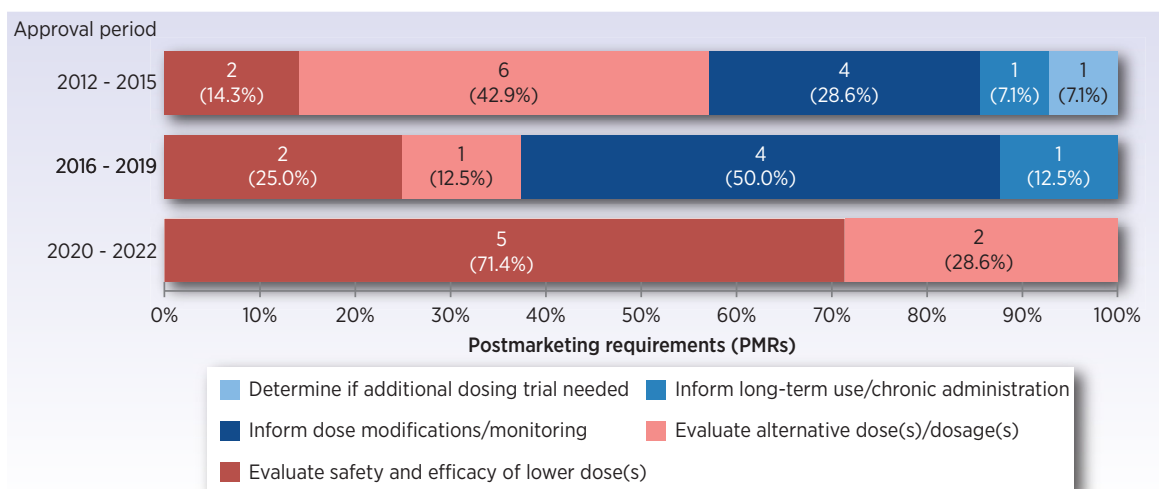


Figure 2. Dosing variation PMRs by type of information over time. FDA issued more PMRs directing sponsors to evaluate a dose lower than the one approved in the last 3 years ($n = 5$, 2020–2022) compared with the preceding 8-year period ($n = 4$, 2012–2019).

Authors' Disclosures

No disclosures were reported.

Received August 2, 2023; revised September 20, 2023; accepted October 19, 2023; published first December 12, 2023.

References

- US Food and Drug Administration (FDA). Project Optimus [Internet]. 2023 [cited 2023 Apr 27]. Available from: <https://www.fda.gov/about-fda/oncology-center-excellence/project-optimus>.
- US Food and Drug Administration. Postmarketing Requirements and Commitments: Introduction [Internet]. US FDA Website. 2016 [cited 2023 Jun 14]. Available from: [https://www.fda.gov/drugs/guidance-compliance-regulatory-information/postmarket-requirements-and-commitments#:~:text=Postmarketing%20commitments%20\(PMCs\)%20are%20studies,by%20a%20statute%20or%20regulation](https://www.fda.gov/drugs/guidance-compliance-regulatory-information/postmarket-requirements-and-commitments#:~:text=Postmarketing%20commitments%20(PMCs)%20are%20studies,by%20a%20statute%20or%20regulation).
- Ridge S, Guinn D, Pfuma Fletcher E, Zineh I, Madabushi R, Ramamoorthy A. US FDA postmarketing requirements and commitments: a systematic assessment of clinical pharmacology studies and their impact on US FDA prescribing information. *J Clin Pharmacol* 2023;63:695–703.
- Guinn D, Madabushi R, Wang YM, Zineh I, Maxfield K. Elucidating the impact of immunogenicity assessment postapproval: a targeted analysis of immunogenicity postmarketing requirements and commitments. *Clin Pharmacol Ther* 2021;109:697–704.
- Hsieh J, Sahre M, Yang X, Madabushi R, Ramamoorthy A. Clinical pharmacology information in regulatory submissions and labeling: a comparative analysis of orphan and non-orphan drugs approved by the FDA. *Clin Transl Sci* 2022;15:2583–96.
- Hyogo A, Kaneko M, Narukawa M. Factors that influence FDA decisions for postmarketing requirements and commitments during review of oncology products. *J Oncol Pract* 2018;14:e34–41.
- US Food and Drug Administration. Approved cellular and gene therapy products [Internet]. 2023 [cited 2023 Jun 14]. Available from: Approved Cellular and Gene Therapy Products.
- Shah M, Rahman A, Theoret MR, Pazdur R. The drug-dosing conundrum in oncology: when less is more. *N Engl J Med* 2021;385:1445–7.
- US Food and Drug Administration. Draft guidance for industry: optimizing the dosage of human prescription drugs and biological products for the treatment of oncologic diseases [Internet]. 2023 [cited 2023 Jun 28]. Available from: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/optimizing-dosage-human-prescription-drugs-and-biological-products-treatment-oncologic-diseases>.
- Fourie Zirkelbach J, Shah M, Vallejo J, Cheng J, Ayyoub A, Liu J, et al. Improving dose-optimization processes used in oncology drug development to minimize toxicity and maximize benefit to patients. *J Clin Oncol* 2022;40:3489–500.
- US Food and Drug Administration. FDA oncologic drugs advisory committee meeting briefing document: phosphatidylinositol 3-kinase (PI3K) inhibitors in hematologic malignancies [Internet]; 2022. Available from: <https://www.fda.gov/media/157762/download>.
- Kurzrock R, Lin C-C, Wu T-C, Hobbs BP, Pestana RC, Hong DS. Moving beyond 3+3: the future of clinical trial design. American Society of Clinical Oncology Educational Book; 2021.

Interpreting Data from Dose-Finding Studies in Early Phase Oncology Trials to Determine the Optimal Dose

Introduction

A critical aspect of drug development is identifying the appropriate dose* that leads to maximal efficacy balanced with safety and tolerability. Oncology clinical trials historically focused on a maximum tolerated dose (MTD) because early systemic therapies such as cytotoxic chemotherapies often have steep dose-response curves that suggest a higher dose equates to higher efficacy.¹ Newer therapeutic classes like molecularly targeted therapies and immunotherapies may have wide separation of dose-response curves between safety and efficacy leading to efficacious doses that are lower than the MTD, and thus resulting in better tolerability while maintaining efficacy. In addition, some agents may have an efficacy curve that is bell-shaped, with higher doses delivering less efficacy than intermediate doses. In recent years, through Project Optimus and recent draft guidance, the U.S. Food and Drug Administration (FDA)'s Oncology Center of Excellence (OCE) has emphasized the need for premarket dose optimization in clinical trials to ensure patients receive drugs that are effective, safe, *and* tolerable.^{2,3} The goal of Project Optimus is "to educate, innovate, and collaborate with companies, academia, professional societies, international regulatory authorities, and patients to move forward with a dose-finding and dose optimization paradigm across oncology that emphasizes selection of a dose or doses that maximizes not only the efficacy of a drug but the safety and tolerability as well."²

Oncology drug trial sponsors are generally moving towards early phase clinical trial designs that balance efficacy, safety, and tolerability to identify an optimized dose. However, a key uncertainty is how to establish the appropriate totality of evidence from these different endpoints and how to interpret the data to select optimal dose(s), which is a dose that can maximize the benefit/risk profile or provide the desired therapeutic effect while minimizing toxicity,³ that align with the goals of Project Optimus. Specifically, a clear understanding of how to assess and generate evidence for tolerability and how it fits into the totality of evidence is needed. Several potential trial designs and statistical analyses that support improved approaches to collecting early phase trial data have been identified.^{4,5} However, the desire for additional data collection adds complexity to study design and data interpretation. As such, it is also critical to be forward thinking and consider how emerging technologies can assist with data collection and analysis, including how to integrate new data with what is included in existing collection approaches.

* The term dose is used throughout this document to refer both to dose, the amount of the drug, and dosage, the amount of the drug and its schedule.

Thank You to Our Contributors

Giovanni Abbadessa, Sanofi
Hillary (Stires) Andrews, Friends of Cancer Research
Ethan Basch, University of North Carolina at Chapel Hill
Vishal Bhatnagar, U.S. FDA
Julie Bullock, Certara
Alicyn K. Campbell, AstraZeneca
Amylou C. Dueck, Mayo Clinic
Judith Medeiros Fitzgerald, Breast Cancer Advocate & Founder of Sisters4Prevention
Houston N Gilbert, Human Immunology Biosciences (HI-Bio)
Lee Jones, Research Advocate
Ramon K. Kemp, GSK
Bellinda L. King-Kallimani, LUNgevity Foundation
Chunze Li, Genentech
Sheetal Patel, Genentech
Nalin Payakachat, Eli Lilly and Company
Devin Peipert, Northwestern University Feinberg School of Medicine
Eric H. Rubin, Merck & Co., Inc, Rahway, NJ, USA
Enrique Sanz Garcia, Princess Margaret Cancer Centre, University of Toronto
Mirat Shah, U.S. FDA
Gita Thanarajasingam, Mayo Clinic
Peter C. Trask, Genentech
Lynne I. Wagner, University of North Carolina at Chapel Hill
Stefan Zajic, GSK
Li Zhu, Bristol Myers Squibb

This document reflects discussions that occurred among stakeholder groups on various topics. This document should not be construed to represent FDA's views or policies.

To support a comprehensive approach to data integration and interpretation for oncology drug dose optimization, Friends of Cancer Research (*Friends*) convened stakeholders to outline the types of data that are collected during dose-finding trials, consider how to prioritize data collection, and propose ways to interpret these data in the identification and selection of the optimal dose(s) for registrational trials. Given the current drug development environment where only 9% of Phase 1 experimental agents make it to registration,⁶ there is risk in any decision-making. It is critical to make a concerted effort to identify the best possible dose that maximizes efficacy while reducing toxicity and asks the minimum possible number of patients needed to contribute to such an effort.

Data that Establish the Totality of Evidence

Data collected from dose-finding trials are encompassed within five main categories, each with a different purpose: pharmacokinetics (PK), pharmacodynamics (PD)/ target engagement (TE), efficacy, safety, and tolerability. For each of these categories, the purpose of including the data category, the type of data currently collected, challenges with the current data collection approaches, and opportunities for improving data collection are described below. When determining the data collection approach within these five categories, trialists should consider not only the methodological approach or assay used to collect these data, but also the appropriate assessment frequency of data collection.

Pharmacokinetics (PK)

Pharmacokinetics (PK) establishes how the body interacts with the drug and evaluates the absorption, distribution, metabolism, and elimination of the therapeutic. PK is often analyzed via serial plasma/serum concentrations collected within hours or days after the administration of a drug. Collecting information on food intake (e.g., through using a food diary) and concomitant medications can aid PK interpretation. Currently, drug distribution to specific tissues of interest (i.e., the tumor) is not commonly assessed, but novel techniques are emerging to assess distribution to target sites.

Given the importance of exposure-response analyses for dose decision-making, trialists should plan to include PK sampling in all patients during dose-finding trials. The extent of the PK sampling can vary from intensive (i.e., 8-10+ samples/patient) to sparse (i.e., 2-4 samples/patient) or a combination of both. Population PK modeling is a tool that should be leveraged to derive modeled parameters from both intensive and sparse PK data. When designing studies, it is critical to consider the time toxicity of cancer treatments for patients, which includes the time spent coordinating care and frequency of visiting the healthcare facility.⁷ Incorporating flexibility into protocol language for the safety committee to make decisions about stopping or re-starting full PK sample collection based on emerging data can save time for the trialist rather than submitting and waiting for protocol amendments to be approved.

The main challenges in measuring PK in dose-finding studies are the operational and logistical considerations of sample collection due to the frequency/intensity, questions about which cycles to collect data, and the number of patients contributing PK samples. To help with the operation and recruitment burdens of PK sampling, at-home sampling and dried blood spot sampling⁸ have emerged and could ultimately result in increased data collection and more accurate PK profiling because of the ability to collect PK samples more frequently at the timepoints that are important

for PK characterization. The use of these newer approaches requires additional validation steps to include as part of the totality of evidence.

Pharmacodynamics and Target Engagement (PD/TE)

Pharmacodynamics and target engagement (PD/TE) aim to assess how the drug interacts with the body and the tumor. Most PD/TE studies measure TE through tumor biopsies or peripheral sampling such as blood or cerebral spinal fluid. Depending on the location of the disease, performing multiple biopsies may be impractical or impossible. Protocols for early phase solid tumor trials that require multiple tumor biopsies might cause some patients to not enroll, ultimately precluding them from accessing potential life-prolonging therapy. To overcome this, imaging methods to assess receptor occupancy are increasing in use and can provide insights into tumor dynamics.

The clinical relevance of many PD biomarkers in the context of antitumor effects is often unknown in the first-in-human study and it is unclear how much receptor occupancy is necessary to elicit a drug response. There may be differences in timing to evaluate PD/TE according to the mechanism of action, which may be challenging for first-in-class drugs due to the lack of prior knowledge. Characterizing the dose to PD to activity relationship in relevant preclinical models in both the tumor and the periphery improves the ability to leverage PD biomarkers for decision-making.

When available, circulating PD biomarkers may be used, some of which are indicators of activity linking the impact of the drug on the tumor while others are purely mechanistic. The priority should be for early efficacy markers that help establish PK-response relationships. Some biomarkers that are indicators of activity are specific for certain cancer indications (e.g., protein derived tumor markers such as prostate specific antigen (PSA) in prostate cancer, M protein in multiple myeloma). Novel techniques like measuring the kinetics of circulating tumor DNA (ctDNA) may support an understanding of PD. For mechanistic biomarkers, there may be opportunities to monitor quantitative and qualitative changes in immune cell populations (e.g., T cells) in plasma specifically for therapies that target the immune system (i.e., immunotherapies). Peripheral biomarkers (e.g., T-cell activation and cytokines), when relevant like in the case of T-cell engagers, can help characterize the pharmacologically active dose range. However, analyses of circulating immune cells may not reflect tumor dynamics. Preclinical and clinical studies that aim to address whether peripheral blood reflects tumor PD (especially leveraging novel single cell technologies), will further improve the utility of peripheral blood-based assessment.

Overall, low specificity and high variability of circulating biomarkers and assays can make interpretability in clinical trials challenging. Characterizing PD biomarkers in clinically relevant samples to validate the assay (e.g., signal to noise, variability in longitudinal samples), should be leveraged to prioritize PD biomarkers and assays, prior to first-in-human studies. There are gaps regarding clinical relevance of thresholds and timing for measuring PD/TE, as circulating PD modulation may not correlate with anti-tumor effect. Standardization and alignment of many PD biomarkers (e.g., ctDNA) is ongoing and identifying the right biomarker to inform the dose selection is critical.

Efficacy

Efficacy provides information about whether the therapy treats the patient's disease. In solid tumors, assessment using Response Evaluation Criteria in Solid Tumors (RECIST) criteria is a common approach based on analyzing tumor measurements from radiographic imaging at different timepoints, while in hematologic malignancies, disease-specific imaging and/or blood test-based criteria have been defined.^{9,10} Tumor burden as measured by imaging for solid tumors and/or blood test for hematologic malignancies can support the development of tumor growth kinetics models. There may be opportunities to compare the tumor growth kinetics before and after experimental therapy, and between doses or treatment options.¹¹ An emerging technology is the use of radiomics, which can provide further granularity into solid tumor dynamics. For hematological malignancies, minimal Residual Disease (MRD) is an emerging approach to measure the depth of response.¹²

One potential challenge to efficacy assessments is that the efficacy endpoints used for dose selection may not be the same as those used for marketing decisions. Overall survival (OS) is important for evaluating overall efficacy in clinical trials, however, using OS in dose-finding studies is not practical as the endpoint takes a long time to generate. Additionally, time-to-event endpoints are not reliable in single-arm cohorts due to confounding by baseline prognostic factors. Therefore, identifying relevant early efficacy endpoints is crucial for dose decisions. Prospective assessment of early efficacy endpoints (i.e., objective response rate (ORR), model-based tumor growth inhibition/ctDNA dynamic metrics, MRD) and an understanding of how they could relate to long-term clinical benefit might be valuable to support the selection of the appropriate earlier endpoints for dose decisions.¹¹

Another challenge with measuring efficacy is that many emerging drug targets may be tissue agnostic and companies often consider multiple tumor types in their clinical development strategy; therefore, the earliest stages of trials may include multiple cancer types. When developing trial designs and analytic approaches, consider the level of homogeneity in the patient population, including whether it is by a biomarker or a histological type (or both). When considering dose-finding in multiple cancer types, one option is to focus dose-finding on one cancer type or a cluster of cancer types (e.g., cancer types driven by the same mutation, those with similar sensitivity to a certain class of agents) in a trial. Alternatively, patients can be stratified by tumor type and analyses can be performed on all patients and by tumor type if tumor type drives efficacy. A newer approach to analyzing the efficacy of multiple cancer types in early phase trials is using a pruning and pooling approach, where potentially inactive tumor indications are removed, and the efficacy data across the remaining doses is pooled for the analysis to enable the dose decision.¹³

Safety

A common approach to measuring safety is to use investigator reporting via the Common Terminology Criteria for Adverse Events (CTCAE), which includes a severity scale for each adverse event (AE). Typically, dose-finding trials focus on rates of serious Grade 3-4 events to determine safety. Together with laboratory results that also measure AEs, CTCAE graded AEs support an understanding of dose limiting toxicities (DLTs), or side effects that are serious enough to prevent an increase in dose. DLTs are generally defined as the presence of any Grade 3 or

higher nonhematological or Grade 4 or higher hematological toxicity at least possibly related to treatment within the DLT assessment window (i.e., the first few weeks of treatment).¹⁴ In early phase clinical trials, there are sometimes difficulties with associating AEs to a drug rather than underlying disease because patients are often sicker, and there is no control arm. Paying close attention to Adverse Events of Special Interest (AESI) may help in focusing on AEs specific to the treatment and not the disease alone.

A key challenge to safety measurements is timing, which can be complicated by AEs that emerge later or are compounded as time goes on (i.e., those that are chronic, cumulative, or delayed). Early safety signals may not fully represent the safety events that happen outside of the DLT period, which is increasingly more common in newer classes of cancer therapeutics such as immuno-oncology drugs, targeted agents, and antibody-drug conjugates. Additionally, low-grade toxicities like Grade 1 and 2 AEs that occur frequently and/or compound over time impact patients more substantially when they receive therapy for months or years. Therefore, the assessment of AEs needs to consider these these later and compounding effects.

In the future, there may be opportunities to use biometrics measured by wearable devices, mobile applications, biosensors, and biomarkers for real-time monitoring signs of AEs to enable earlier intervention once biasing “noise” (i.e., excessive data collected) is sorted out. Real-time monitoring of certain health parameters (e.g., vital signs, physiological events) may support a clearer understanding of safety signals. If used successfully in clinical trials, these interventions would be expected to be used in clinical practice as well.

Tolerability

The tolerability of a medical product is the degree to which symptomatic and non-symptomatic AEs associated with the product’s administration affect the ability or desire of the patient to adhere to the dose or intensity of therapy.¹⁵ Because the goals of Project Optimus focus on tolerability and approaches to measuring tolerability are emerging, there is an increased emphasis on this topic included below. Currently, tolerability assessments in dose-finding studies are primarily measured by the number of dose reductions, interruptions, and discontinuations as well as physician-reported AEs as a proxy for the patient’s ability or desire to adhere. Sometimes, dose modifications may be driven by physician or patient preferences, or logistical reasons unrelated to tolerability (e.g., due to the patient’s schedule, including modifications for travel). Documentation of the reason for dose modifications or discontinuation, including a differentiation of dose changes due to tolerability versus other reasons, may support a more precise assessment of the relationship between dose intensity and tolerability.

It is increasingly recognized that any assessment of tolerability in a clinical trial without systematically collecting data about the patient’s experience is incomplete.¹⁵ In 2022, *Friends* developed a white paper highlighting key considerations for collecting patient-reported outcomes (PROs) in dose-finding studies.¹⁶ PROs capture the patient perspective, are considered the gold standard when measuring patient experiences, and include key elements of tolerability such as symptomatic AEs, and bother with side effects of treatment.¹⁷ Certain side effects measured by PROs can provide insights into larger problems as they precede long-term consequences of a drug, including nausea or anorexia that causes profound weight loss or neuropathy that becomes

irreversible. A challenge to using PROs in dose-finding studies is that this is a novel approach, and as a result there are not standard methods for how to use and interpret PROs to assist in making decisions about dose. Despite this, there are a variety of proposals for collecting PROs in early phase trials.^{3,16,18}

A few outstanding considerations about incorporating PROs include:

- **Many AEs occur outside of office visits.** Ideally, PROs would assess the patients' experience on an event-driven basis (i.e., symptomatic AE onset or worsening) in addition to a calendar-driven basis at a regular cadence through an electronic PRO (ePRO) platform, which would allow for push notifications, time stamping, and assessing key domains when most relevant to the patients' experience (i.e., maximum experience of symptomatic AEs) independent of scheduled clinical encounters. Using ePRO collection requires effort to initially set up including implementation time, cost, considering patient factors (e.g., technology literacy, age, frailty), and practice factors (e.g., infrastructure and staffing of clinical team to review and respond to alerts). There is precedent and feasibility for using ePROs during later-stage trials and outside of trial settings, such as observational research for PRO evaluation and clinical assessment of PROs, which can be leveraged to inform approaches for ePRO collection in early phase trials.¹⁹ Paper PRO collection could be employed when remote collection is not practical for patients who lack access to or are not comfortable with the use of technology. Awareness of the challenges, including confirming when the paper PRO collection was completed and by whom, should be addressed. When considering collection approaches, PRO instruments like the PRO-CTCAE are generally equivalent regardless of the mode in which they are administered, meaning that PRO-CTCAE surveys completed directly by the patient may be interchangeably administered by electronic system, paper, or automated telephone system, based on the preferences and circumstances of a given patient or study design.²⁰ The potential rigor lost by accepting multiple modes for self-administered PRO collection and the balance with what is gained in terms of more complete data and approaches that suit all types of patients should be considered.
- **The optimal timing of when PROs should be analyzed, including how this information may impact interpretation of tolerability.** One option is to analyze PROs at the end of the trial, which means that clinical trial staff would not have access to patient-level PROs as they arise. However, this approach can prevent PRO data from being used to inform clinician assessment. An emerging approach of interest is to share PRO data with site investigators during trial conduct to inform management of patients' symptoms. By sharing PRO data in real-time, clinicians can use patient responses to inform their own CTCAE reporting, which also ameliorates potential concerns about reconciliation of tolerability data. This approach has been shown to be feasible and improve alignment of CTCAE reporting with the patient experience.²¹ As an example, **Figure 1** represents a form used in the NCI cooperative group randomized clinical trial, N1048. Patients reported the PRO-CTCAE electronically and this information auto-populated an AE form for clinical investigators to review and complete at the point of care during trial conduct.²² A similar approach was used in early-phase trials at Memorial Sloan Kettering Cancer Center, in which patients report PROs in the waiting room prior to visits, and then the PROs populate a software interface through which investigators enter their own CTCAE scores (**Figure 2**). A benefit of this approach is that the patient's perspective on their treatment is used at the point of care to inform trial conduct. Patients

have noted concerns that their PRO data might be used as a rationale to remove them from trials, however, findings from prior cooperative group trials where PRO data was shared with investigators noted no increase in trial discontinuation even among patients with severe toxicities based on PRO data. Patient education is critical for each PRO approach at the outset of the trial, so patients understand how this information is and is not being used within the study.

Solicited Adverse Event

(Page 1 of 4)

Alliance for Clinical Trials in Oncology

Protocol Number: N1048
 Patient ID: demo-nt
 Institution (Inst. Number): [REDACTED]

Instructions: This form should be completed by the CRA per the Test Schedule (Section 4.0) using patient’s medical records, starting from the first day since the prior reporting period if post-baseline. When completing this form, the patient’s self-reported adverse event ratings (shown in the table below for the reporting period) should be used as a reference.

Reporting Period End Date (date on which the clinician evaluated patient’s adverse events) (dd-MMM-yyyy) 29-MAY-2015

Adverse Event Text Name (CTCAE v 4.0)	MedDRA Adverse Event Code (v. 12.0)	Patient Self-Reporting Adverse Event Ratings			Check the circle next to event(s) not evaluated at this visit	Adverse Event Grade (highest grade in this reporting period) INCLUDE GRADE 0's					Has an adverse event expedited report been submitted?*	
		Severity	Frequency	Interference with Activities of Daily Life		0	1	2	3	4		5 (death)
Dysphagia	10013920	None	-	-	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	5 (death)	
Diarrhea	10012727	-	Rarely	-	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	5 (death)	
Nausea	10028813	Mild	Rarely	-	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>		
Vomiting	10047700	-	Never	-	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	5 (death)	
Dyspnea	10013963	None	-	-	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	5 (death)	
Peripheral sensory neuropathy	10034620	None	-	-	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	5 (death)	
Constipation	10010774	Mild			<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	5 (death)	
Mucositis oral	10028130	None			<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	5 (death)	

Figure 1. Example of a form used to populate PROs to provide information to clinicians during clinical trials by the Alliance for Clinical Trials in Oncology. In the example, the patient portion of the form has been populated, and clinician reporting of CTCAE needs to be added to the form. This is an approach to generate patient-informed clinician-reported AEs in real time during a clinical trial.

Patient: P081091, Demo Survey date: 6/15/2011 10:55 AM

Adverse symptom	Patient self report	Date	Agree?	Clinician reassign	Attribution
ALOPECIA	GRADE 0	6/15/2011 10:54 AM	Agree	GRADE 0	N/A
ANOREXIA	GRADE 1	6/15/2011 10:53 AM	Disagree	GRADE 2	Unrelated
COUGH	GRADE 1	6/15/2011 10:53 AM	Agree	GRADE 1	N/A
DYSPNEA	GRADE 1	6/15/2011 10:51 AM	Disagree	GRADE 2	Unlikely
EPIPHORA	GRADE 0	6/15/2011 10:55 AM	Agree	GRADE 0	N/A
EPISTAXIS	GRADE 0	6/15/2011 10:55 AM	Agree	GRADE 0	N/A
FATIGUE	GRADE 0	6/15/2011 10:51 AM	Disagree	GRADE 1	Possibly
KPS	100%	6/15/2011 10:55 AM	Agree	GRADE 0 GRADE 1 GRADE 2 GRADE 3 GRADE 4 GRADE 1	N/A
MUCOSITIS/STOMATITIS	GRADE 1	6/15/2011 10:54 AM	Agree	GRADE 1	N/A
MYALGIA	GRADE 1	6/15/2011 10:51 AM	Agree	GRADE 1	N/A
NAUSEA	GRADE 0	6/15/2011 10:54 AM	Agree	GRADE 0	N/A
PAIN	GRADE 0	6/15/2011 10:51 AM	Agree	GRADE 0	N/A
SENSORY NEUROPATHY	GRADE 1	6/15/2011 10:50 AM	Agree	GRADE 1	N/A
VOICE CHANGES/HOARSENESS	GRADE 1	6/15/2011 10:54 AM	Agree	GRADE 1	N/A

Lock Submit

Figure 2. A second example of a form used to populate PROs to provide information to clinicians during clinical trials from Memorial Sloan Kettering Cancer Center.

- Strategies for selecting which PROs to include for dose-finding trials are emerging.^{23–28} One approach is to start with an established group of core items from the PRO-CTCAE, then include additional PRO-CTCAE items for expected toxicities based on drug class or prior publications, a single-item global side effect impact item to capture the cumulative experience of toxicities, and a free text box for unsolicited AEs.²⁶ FDA guidance provides directions on which core domains are important to measure in cancer trials, although the guidance does not specify which measures are best suited for use in dose optimization trials.³ Additional suggested approaches are included in the *Friends’* white paper from 2022, such as the use of a free-text item to capture newly emerging toxicities.¹⁶ The use of a free-text item can inform selection of patient-reported symptoms for later drug development when the toxicity profile of a drug may be otherwise unknown.

Overall, how PRO data are considered in the totality of evidence and how they can contribute to decisions about dosing is not yet fully established and would benefit from additional standards or guidelines. PROs can complement investigator-derived safety information to determine the benefit-risk of different doses, particularly for treatment-related toxicities that are poorly captured

by investigator assessment (e.g., low-grade diarrhea, blurry vision that is transient but recurs daily, etc.). While standards and guidelines are being developed, including PROs in dose-finding studies to optimize dose is encouraged to capture a comprehensive assessment of tolerability.

Interpreting the Data that Establish Totality of Evidence to Determine the Optimal Dose

Dose decisions from dose-finding studies do not occur at a single timepoint, as the data that establish the totality of evidence are different at each decision point and should be interpreted as such. An idealized dose-finding clinical trial(s) includes two phases, the Dose Escalation phase and the Dose Expansion phase, which are often part of or completely encompass Phase 1 and Phase 2 clinical trials. Dose-finding trials have three decision points for dose selection (Figure 3): 1) during Dose Escalation, to determine whether more patients should be included in that dose level, whether the level should be increased or decreased, and whether to evaluate intermediate doses, 2) at the end of Dose Escalation, to identify the dose(s) and schedule(s) for Dose Expansion, and 3) at the end of Dose Expansion, to identify the dose(s) for subsequent clinical investigations or a registrational trial.

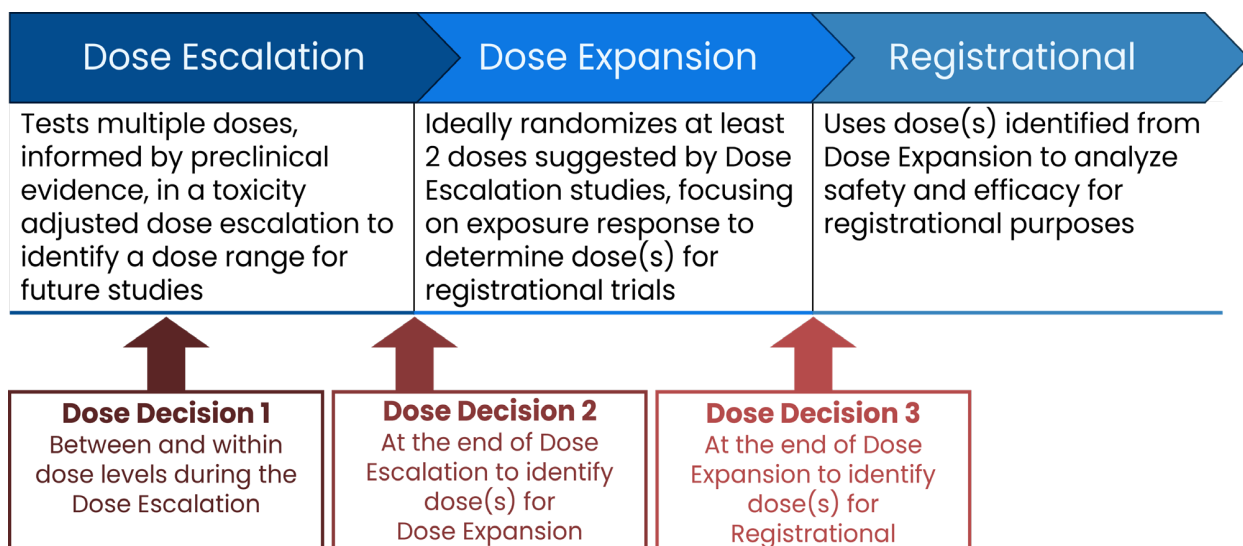


Figure 3. Dose-finding trials and the decisions about dose that occur throughout.

The analysis at each decision point should be a benefit-risk assessment using the totality of data available at that decision point, as not all categories of data will be available at all decision points and in the context of some of the data, there will not be enough of it at certain timepoints to make meaningful conclusions. Although data driven, the decisions are not necessarily statistically powered for each data element.

The approach to interpretation outlined below considers a monotherapy as a first indication in a population with metastatic disease.

Dose Decision 1 – During Dose Escalation

The Dose Escalation phase tests multiple doses and schedules, adjusting based on toxicity, to identify a dose range for future studies. The starting dose, and considerations for how many doses should be included at this point, are often driven by preclinical data. A dose range is an important output of Dose Escalation, which may include the MTD, as a necessary step to characterize the drug. The MTD is defined using DLTs and other severe toxicities that may happen outside of the DLT period. A common approach to defining MTD involves selecting a target DLT rate such as 25% and using a model-based or model-assisted approach for dose escalation.^{29,30} Backfill cohorts of a select subset of doses are sometimes included in parallel during Dose Escalation to assess the safety, tolerability, and activity. Not all trial designs and Dose Escalation trial populations are well-suited for backfilling, therefore it is important to know if backfilling escalation cohorts will provide meaningful data and how these data would be interpreted and used to make decisions about dosing. Backfilling may use a significant amount of patient resources and limitations to control the number of patients enrolled into backfill should be established.

Dose-Escalation trials are often first-in-human and conducted in a heterogeneous patient population with respect to tumor type, prior treatments and patient co-morbidities which may confound detailed data interpretation at this stage but should yield useful trends to define a dose range for further evaluation. The patients enrolled typically have exhausted standard of care options. While understanding the lower limit of the dose range is critical, it is also important to not start too low. Preclinical data and/or clinical data from other treatments in the same class can support a starting dose. There is an increasing trend to not expose patients to inactive doses and rather use an accelerated titration, especially for those drugs which are not first in class.

When deciding about increasing doses within Dose Escalation, the focus is largely on safety (i.e., DLT criteria and severe AEs) and to some extent PK and PD data. However, PK and PD data availability and analysis typically lag safety and therefore are not often included in early Dose Escalation decision criteria. As PK and PD data emerge, even if they lag 1-2 cohorts behind, these data can be considered for decision-making during later parts of Dose Escalation and into Dose Expansion. Additional dose escalation may not always be warranted if exposure remains unchanged with dose due to saturation of absorption (i.e., solubility) or if a target threshold PK level is reached. Emerging safety signals and tolerability from earlier dose cohorts that occur after the DLT period should also be considered in the Dose Escalation decision, especially if they limit not only the dose, but how long a patient is likely to remain on treatment.

TE (i.e., receptor occupancy; RO) may play a role in the Dose Escalation decision provided these results are available within a reasonable turnaround time. Target-mediated drug disposition could provide indirect evidence of TE/RO for easily accessible targets. When relevant, PD/TE biomarkers may be used to define the range of active doses to backfill with safe and potentially active doses at the end of Dose Escalation.

While collecting and analyzing PRO data may not be the focus of Dose Escalation, these data can support an understanding of profiles of symptomatic AEs by dose^{31,32} and support an understanding of tolerability in subsequent trials. Including PROs in Dose Escalation can help sites establish processes for PRO implementation to carry into later phase trials allowing the patient perspective to inform all phases of drug development. Failing to include PROs in this phase may miss the opportunity to consider the patient perspective in an appropriate and purposeful manner.

Dose Decision 2 – Selecting dose(s) for Dose Expansion

Dose Expansion ideally takes two or more doses and/or schedules from Dose Escalation and assesses them in a larger and potentially more homogeneous population with a focus on dose and exposure response analyses for safety and efficacy to determine the dose(s) for the registrational trials. Draft FDA guidance recommends randomized, parallel dose-response designs, where randomization helps to avoid selection bias.³

From a totality of evidence perspective, decisions about which dose(s) to bring to Dose Expansion should incorporate safety and PK, but also consider PD/TE, tolerability, and activity and should be supported by exposure-response analyses when feasible. Transitioning from Dose Escalation to Dose Expansion allows for an analysis of safety data collected during the entirety of Dose Escalation to identify emerging safety signals that may not have been evident during the DLT period. For PK, it is important to assess linearity to ensure that doses chosen for Dose Expansion do not have significantly overlapping exposures. Activity tracked by tumor dynamics or changes in ctDNA can give initial glimpses of efficacy. For tolerability, in addition to reviewing dose modifications and dose intensity, an assessment of patient-reported tolerability can be included with a focus on symptomatic adverse events, and side effect bother, assessed with validated PROs. For example, a single global side effect impact item can assess side effect bother (e.g., Functional Assessment of Cancer Therapy General item GP5 “I am bothered by side effects of treatment” or EORTC IL46 “troubled by your side effects” item). Co-administration of selected PRO-CTCAE items associated with symptomatic AEs, or other tools’ symptom scales can help inform which side effects are contributing to tolerability-related concerns or confirm a signal seen in safety data.

From a decision-making perspective, there are limited examples of how including PROs influences decisions about dose to date. Currently, PROs are unlikely to change which doses are pursued in the Dose Escalation phase, however, they can aid in providing confidence in an AE profile to support the doses selected for further evaluation or signal the need for approaches that mitigate certain AEs. Additionally, PROs can help detect unanticipated toxicities or influence approaches to defining safety and tolerability in subsequent dose-finding studies. Future research should consider the best approaches for interpreting data about PROs.

When deciding about dose(s) to bring to Dose Expansion, the interplay between activity/efficacy and TE should be considered. Dose optimization without some level of observed efficacy, or at least of PD activity, may lead to choosing ineffective doses and may prevent optimization in the proper indication(s). Selection of a dose well above tumor RO saturation may not be warranted as it is unlikely to provide additional antitumor activity and may lead to increased toxicity. Caution should be made when RO is calculated but not measured unless the assumptions are validated

clinically. Because of these limitations (i.e., uncertainties about timing/ relationship to efficacy), PD/TE data should be used in conjunction with other data to identify doses for evaluation in Dose Expansion.

Regarding safety and tolerability, when evaluating exposure-response relationships, it may be helpful to consider exposure-response relationships for multiple safety and tolerability measures to support the dose(s) for further evaluation. Interpretation of exposure-response relationships should involve experts in quantitative pharmacology. When determining which doses to evaluate further, even if there are doses predicted to have efficacy and not associated with serious toxicity, but tolerability is poor, it would be helpful to include this dose and a lower dose or alternate regimen which could improve tolerability in the Dose Expansion study.

Dose Decision 3 – Selecting dose(s) for Registrational Trials

At the end of Dose Expansion, the totality of evidence is greater enabling more robust quantitative approaches to dose selection. The population in Dose Expansion is generally more focused on the final target indication, allowing for more accurate decision-making about dose regarding efficacy, safety, and tolerability. Additionally, findings from Dose Expansion will set the stage for the measurement of more targeted safety and tolerability endpoints. In some cases, when results of a randomized Dose Expansion are inconclusive, further randomized dose selection may be incorporated into a registrational trial.

When determining which dose(s) to evaluate in Registrational Trials, analyses will continue to incorporate PK and PD/TE and include longer term data for efficacy and tolerability. The use of population PK, exposure-response modeling, and longitudinal PK/PD model (e.g., PK-tumor dynamic or lab values if there is a lab AE) to characterize trends in exposure and activity, efficacy, safety, and tolerability is expected to support a dose for registration. It is important to consider the overall benefit-risk of the various doses, and clinical judgment will likely be required to evaluate potential tradeoffs between efficacy and safety.

Conclusions and Next Steps

Recent FDA draft guidance³ and a recently posted FDA toolkit³³ provides considerations for dose optimization³ and ongoing studies focused on dose-finding will provide supplementary information about additional settings. Increasingly in oncology, therapies are administered in combination. In September 2023, FDA co-hosted a workshop with ASCO focused on dose-finding in combination therapies.²⁴ Pediatric drug dosing is another area that will benefit from additional focus, and FDA hosted an Oncologic Drugs Advisory Committee meeting focused on dosing in drugs indicated for pediatric populations.³⁴ Further, it will be helpful to define the criteria necessary for extrapolating doses from one therapy, or therapeutic class, to another. Whether the same dose or a new dose would be necessary will depend on the available data and appropriate justification by the sponsor. In each of these situations, discussions outlined in this white paper should be considered as principles regarding what is included in the totality of evidence will remain. Future studies to support approaches to data extrapolation, which information to include in dose-finding trials, and how to interpret the data to select the dose will ensure patients receive the optimal dose that provides efficacy balanced with safety and tolerability.

References

1. Martin JH, Dimmitt S. The Rationale of Dose–Response Curves in Selecting Cancer Drug Dosing. *Br J Clin Pharmacol*. 2019;85(10):2198. doi:10.1111/BCP.13979
2. FDA’s Project Optimus. <https://www.fda.gov/about-fda/oncology-center-excellence/project-optimus>
3. FDA. Optimizing the Dosage of Human Prescription Drugs and Biological Products for the Treatment of Oncologic Diseases: Draft Guidance for Industry. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/optimizing-dosage-human-prescription-drugs-and-biological-products-treatment-oncologic-diseases>
4. Araujo D, Greystoke A, Bates S, et al. Oncology Phase I Trial Design and Conduct: Time for a Change - MDICT Guidelines 2022. *Annals of Oncology*. 2023;34(1):48–60. doi:10.1016/j.annonc.2022.09.158
5. Kurzrock R, Lin CC, Wu TC, et al. Moving Beyond 3+3: The Future of Clinical Trial Design. *Am Soc Clin Oncol Educ Book*. 2021;41(41):e133–e144. doi:10.1200/EDBK_319783
6. Dowden H, Munro J. Trends in Clinical Success Rates and Therapeutic Focus. *Nat Rev Drug Discov*. 2019;18(7):495–496. doi:10.1038/D41573-019-00074-Z
7. Gupta A, Eisenhauer EA, Booth CM. The Time Toxicity of Cancer Treatment. *Journal of Clinical Oncology*. 2022;40:1611–1615. doi:10.1200/JCO.21.02810
8. Evans C, Arnold M, Bryan P, et al. Implementing Dried Blood Spot Sampling for Clinical Pharmacokinetic Determinations: Considerations from the IQ Consortium Microsampling Working Group. *AAPS J*. 2015;17(2):292. doi:10.1208/S12248-014-9695-3
9. Cheson BD, Fisher RI, Barrington SF, et al. Recommendations for Initial Evaluation, Staging, and Response Assessment of Hodgkin and Non-Hodgkin Lymphoma: The Lugano Classification. *Journal of Clinical Oncology*. 2014;32(27):3059–3067. doi:10.1200/JCO.2013.54.8800
10. Kumar S, Paiva B, Anderson KC, et al. International Myeloma Working Group Consensus Criteria for Response and Minimal Residual Disease Assessment in Multiple Myeloma. *Lancet Oncol*. 2016;17(8):e328–e346. doi:10.1016/S1470-2045(16)30206-6
11. Bruno R, Bottino D, De Alwis DP, et al. Progress and Opportunities to Advance Clinical Cancer Therapeutics Using Tumor Dynamic Models. *Clinical Cancer Research*. 2020;26(8):1787–1795. doi:10.1158/1078-0432.CCR-19-0287
12. The Promises of MRD Testing in Hematologic Malignancies. ACCCBuzz Blog. <https://www.accc-cancer.org/acccbuzz/blog-post-template/accc-buzz/2022/05/05/the-promise-of-mrd-testing-in-hematologic-malignancies>
13. Zhou H, Liu F, Wu C, Rubin EH, Giranda VL, Chen C. Optimal Two-Stage Designs for Exploratory Basket Trials. *Contemp Clin Trials*. 2019;85:105807. doi:10.1016/J.CCT.2019.06.021
14. Lee SM, Wages NA, Goodman KA, Lockhart AC. Designing Dose-Finding Phase I Clinical Trials: Top 10 Questions That Should Be Discussed With Your Statistician. *JCO Precis Oncol*. 2021;5(5):317–324. doi:10.1200/PO.20.00379
15. Friends of Cancer Research White Paper: Broadening the Definition of Tolerability in Cancer Clinical Trials to Better Measure the Patient Experience; 2018. https://friendsofcancerresearch.org/wp-content/uploads/Comparative-Tolerability-Whitepaper_FINAL.pdf

16. Friends of Cancer Research White Paper: Supporting a Patient-Centric Approach to Dose Optimization in Oncology: The Essential Role of Patient-Reported Outcomes (PROs); 2022. https://friendsofcancerresearch.org/wp-content/uploads/Supporting_Patient-Centric_Approach_Dose_Optimization_Oncology-PROs.pdf
17. FDA. Core Patient-Reported Outcomes in Cancer Clinical Trials: Draft Guidance for Industry. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/core-patient-reported-outcomes-cancer-clinical-trials>
18. Bergerot CD, Pal SK, Tripathi A. Patient-Reported Outcomes in Early Phase Clinical Trials: An Opportunity to Actively Promote Patient-Centered Care. *Oncologist*. 2022;27:714-715. [doi:10.1093/oncolo/oyac122](https://doi.org/10.1093/oncolo/oyac122)
19. Coons SJ, Eremenco S, Lundy JJ, et al. Capturing Patient-Reported Outcome (PRO) Data Electronically: The Past, Present, and Promise of ePRO Measurement in Clinical Trials. *Patient*. 2015;8(4):301-309. [doi:10.1007/s40271-014-0090-z](https://doi.org/10.1007/s40271-014-0090-z)/METRICS
20. Bennett AV, Dueck AC, Mitchell SA, et al. Mode Equivalence and Acceptability of Tablet Computer-, Interactive Voice Response System-, and Paper-Based Administration of the U.S. National Cancer Institute's Patient-Reported Outcomes Version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). *Health Qual Life Outcomes*. 2016;14(1):1-12. [doi:10.1186/s12955-016-0426-6](https://doi.org/10.1186/s12955-016-0426-6)/TABLES/5
21. Basch E, Wood WA, Schrag D, et al. Feasibility and Clinical Impact of Sharing Patient-Reported Symptom Toxicities and Performance Status with Clinical Investigators During a Phase 2 Cancer Treatment Trial. *Clin Trials*. 2015;13(3):331-337. [doi:10.1177/1740774515615540](https://doi.org/10.1177/1740774515615540)
22. Dueck AC, Mitchell SA, Rogak LJ, et al. A Cluster-Randomized Study of Clinician-Patient Shared vs. Standard Reporting of Symptomatic Adverse Events Using PRO-CTCAE Nested in a Multicenter Trial of Multimodal Therapy for Rectal Cancer (Alliance N1048 PROSPECT). *Quality of Life Research*. 2015;24(1 Suppl.) [doi:10.1007/s11136-015-1078-4](https://doi.org/10.1007/s11136-015-1078-4)
23. Cancer Treatment Tolerability Consortium. <https://prevention.cancer.gov/major-programs/cancer-treatment-tolerability-consortium>
24. Second Annual FDA ASCO WORKSHOP on Getting the Dosage Right September 6-7, 2023. <https://www.fda.gov/news-events/fda-meetings-conferences-and-workshops/second-annual-fda-asco-workshop-getting-dosage-right-september-6-7-2023-09062023>
25. Emens LA, Bruno R, Rubin EH, et al. Report on the Third FDA-AACR Oncology Dose-Finding Workshop. *Cancer Immunol Res*. 2017;5(12):1058-1061. [doi:10.1158/2326-6066.CIR-17-0590](https://doi.org/10.1158/2326-6066.CIR-17-0590)
26. Trask PC, Dueck AC, Piau E, Campbell A. Patient-Reported Outcomes Version of the Common Terminology Criteria for Adverse Events: Methods for Item Selection in Industry-Sponsored Oncology Clinical Trials. *Clin Trials*. 2018;15(6):616-623. [doi:10.1177/1740774518799985](https://doi.org/10.1177/1740774518799985)
27. Janse van Rensburg HJ, Liu Z, Watson GA, et al. A Tailored Phase I-Specific Patient-Reported Outcome (PRO) Survey to Capture the Patient Experience of Symptomatic Adverse Events. *Br J Cancer*. 2023;129(4):612. [doi:10.1038/s41416-023-02307-w](https://doi.org/10.1038/s41416-023-02307-w)
28. O'Donnell PH, Arkenau HT, Sridhar SS, et al. Patient-Reported Outcomes and Inflammatory Biomarkers in Patients with Locally Advanced/Metastatic Urothelial Carcinoma Treated with Durvalumab in Phase 1/2 Dose-Escalation Study 1108. *Cancer*. 2020;126(2):432-443. [doi:10.1002/CNCR.32532](https://doi.org/10.1002/CNCR.32532)
29. Ji Y, Liu P, Li Y, Bekele BN. A Modified Toxicity Probability Interval Method for Dose-Finding Trials. *Clin Trials*. 2010; 7(6):653-63. [doi:10.1177/1740774510382799](https://doi.org/10.1177/1740774510382799)

30. Yuan Y, Hess KR, Hilsenbeck SG et al. Bayesian Optimal Interval Design: A Simple and Well-Performing Design for Phase I Oncology Trials. *Clin Cancer Res* 2016;22:4291–4301. [doi:10.1158/1078-0432.CCR-16-0592](https://doi.org/10.1158/1078-0432.CCR-16-0592)
31. Regnault A, Loubert A, Gorsh B, et al. A Toolbox of Different Approaches to Analyze and Present PRO-CTCAE Data in Oncology Studies. *JNCI Journal of the National Cancer Institute*. 2023;115(5):586. [doi:10.1093/JNCI/DJAD018](https://doi.org/10.1093/JNCI/DJAD018)
32. Xia H, Fourie JZ, Booth B, et al. Exploration of PRO-CTCAE Data Used for Exposure-Response Relationships in an Oncology Clinical Trial. *Blood*. 2017;130(Supplement 1):5670–5670. [doi:10.1182/blood.V130.Suppl_1.5670.5670](https://doi.org/10.1182/blood.V130.Suppl_1.5670.5670)
33. FDA. Oncology Dosing Tool Kit. <https://www.fda.gov/about-fda/oncology-center-excellence/oncology-dosing-tool-kit>
34. Pediatric Oncology Subcommittee of the Oncologic Drugs Advisory Committee Meeting Announcement. <https://www.fda.gov/advisory-committees/advisory-committee-calendar/june-16-2023-pediatric-oncology-subcommittee-oncologic-drugs-advisory-committee-meeting-announcement>

A Study Design to Harmonize Patient-Reported Outcomes Across Data Sets

Brittany A. McKelvey, PhD¹; Alexandra Berk, MA²; Lynda Chin, MD^{3,4}; Stacie Hudgens, MSc⁵; Ian Kudel, PhD⁶; Ronan C. O'Hagan, PhD³; Amila Patel, PharmD⁷; Julie Scott, MSN⁸; Hillary Stires, PhD¹; Sam Wang, MD³; Debra Wujcik, PhD⁸; Mark Stewart, PhD¹; and Jeff Allen, PhD¹

PURPOSE Using patient-reported outcomes (PROs) provides important insights from the patient's perspective and can be valuable to monitor and manage treatment-related adverse events during cancer treatment. Additionally, the digital administration of PROs (electronic PROs [ePROs]) provides real-time updates to clinical care teams on treatment-related symptoms in-between clinic visits. However, given the variability in the methodology and timing of the data collection, using and harmonizing these data across different systems remains challenging. Identifying data elements to capture and operating procedures for harmonization across ePRO tools will expedite efforts to generate relevant and robust data on use of ePRO data in clinical care.

METHODS Friends of Cancer Research assembled a consortium of project partners from key health care sectors to align on a framework for ePRO data capture across ePRO tools and assessment of the impact of ePRO data capture on patient outcomes.

RESULTS We identified challenges and opportunities to align ePRO data capture across ePRO tools and aligned on key data elements for assessing the impact of ePRO data capture on patient care and outcomes. Ultimately, we proposed a study protocol to leverage ePRO data for symptom and adverse event management to measure real-world effectiveness of ePRO tool implementation on patient care and outcomes.

CONCLUSION This work provides considerations for harmonizing ePRO data sets and a common framework to align across multiple ePRO tools to assess the value of ePROs for improving patient outcomes. Future efforts to interpret evidence and evaluate the impact of ePRO tools on patient outcomes will be aided by improved alignment across studies.

JCO Clin Cancer Inform 7:e2200161. © 2023 by American Society of Clinical Oncology

Creative Commons Attribution Non-Commercial No Derivatives 4.0 License 

INTRODUCTION

Patient-reported outcomes (PROs) are a mechanism for understanding, monitoring, and managing a patient's treatment-related adverse events (AEs). PROs include symptoms, treatment side effects, function, and health-related quality of life self-reported by the patient. Capturing these data can be valuable to evaluate the tolerability, safety, and effectiveness of oncology products in both the premarket and postmarket settings. In clinical trials, PROs can inform safety and toxicity during drug development and regulatory decision making.¹ There are also significant opportunities to collect PROs in the real-world (rw) setting, including for symptom assessment and enhancing communication between the patient and provider.² Although individual academic institutions have implemented PRO collection in routine cancer care,³ the use of PROs in oncology clinical care is not ubiquitous.

PRO data may provide information on the risks and benefits of an oncology product, inform future research

studies or additional clinical trials, and support labeling decisions. All patient side effects may not be adequately addressed or reported during short encounters in the clinic,⁴ suggesting a need to not only collect PROs during clinic visits, but to systematically capture PRO data throughout treatment. To do this, digital administration of PROs (electronic PROs [ePROs]) provides patients and clinical care teams with an opportunity to address treatment-related symptoms via computer or smartphone applications in real time. This approach can lead to earlier detection and improved management of AEs, resulting in lower use of acute care services, longer time on therapy, and improved outcomes, including overall survival.^{5,6}

To use ePRO data for rw data and rw evidence, the data must be fit for purpose, reliable, and relevant. However, current PRO collection through available ePRO tools is heterogeneous because each tool has its own methodology and data elements. There are many different cancer-related PRO instruments, including the Functional

CONTEXT

Key Objective

To develop an aligned framework for assessing the use of and data generated from electronic patient-reported outcomes (ePROs) collected by various ePRO tools.

Knowledge Generated

We have developed a common study protocol to harmonize key data elements across ePRO tools to maximize data generated and assess the value of ePROs for improving patient outcomes.

Relevance

Strategies and methodologies for harmonizing disparate real-world data collected by different ePRO tools are needed to produce robust data on the use of ePROs and patient care. Our work supports opportunities for aligning multiple data sources to generate real-world evidence.

Assessment of Cancer Therapy-General, the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30, and the PRO version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE).⁷ Furthermore, data collected in the rw setting is less controlled and less standardized than in clinical trials, which may lead to discordance in the timing of assessments across patients in a single data source and across data sources. Data missingness, both at the individual symptom level and entire assessment level, may complicate the interpretation of the data and be caused by methodologic, logistic, administrative, and patient issues.⁸ These issues may make it difficult to interpret and harmonize data across ePRO tools to provide more robust evidence.

To analyze the impact of ePRO data collection on patient outcomes in the rw setting, it is critical to identify data elements to capture and operating procedures for harmonization across ePRO tools to expedite efforts to generate relevant data, while enabling continued innovation of individual technologies that support integration of core PROs in clinical care. Friends of Cancer Research assembled a multistakeholder group of experts including the US Food and Drug Administration, ePRO technology vendors, academic clinicians and researchers, patient advocacy group representatives, and health data aggregators (Data Supplement) to accomplish several goals:

- Identify opportunities and challenges for aligning ePRO data across ePRO tools.
- Align on key data elements for assessing whether ePRO data capture improves patient care and outcomes.
- Develop a common study design that could be used to leverage ePRO data to measure rw effectiveness of ePRO tool implementation on patient care and outcomes.

CHALLENGES TO ePRO DATA SET ALIGNMENT

The reason for including ePROs in clinical care, such as personal tracking between clinic visits or completion as part of a clinic visit, influences which PRO data are captured and timing for collection. This significant heterogeneity in

ePRO data capture can affect the use and alignment of ePRO data to understand the patient experience while on treatment in the rw setting.

Approach to ePRO Data Collection

Each ePRO tool uses a different approach for deploying the PRO instrument. Some may include sending questionnaires to the patient to track their symptoms at a regular frequency, while others may also allow for ad hoc reporting as symptoms arise outside of the regularly scheduled questionnaires. Platforms that allow for ad hoc reporting may capture symptoms before they become a significant issue for the patient, rather than the patient waiting for a regularly scheduled assessment.

Frequency of ePRO Assessment

The variability in the frequency of symptom reporting will affect the interpretation of PRO data. This challenge may arise from within the same ePRO tool, as well as when harmonizing across tools. The frequency of sending questionnaires may be dictated by the clinician or rationale for collecting PROs. Some sites administer short daily or weekly check-ins for frequent symptom and tolerability monitoring, while other sites may schedule assessments once per treatment cycle to reduce patient burden.

Types of PRO Data

The types of PRO instruments integrated into ePRO tools may vary, and there may be feasibility or practical considerations affecting which PRO data are captured through questionnaires in the rw setting. Although comprehensive item libraries capture most possible AEs, such a list of items may be too burdensome for patients and include items the clinical care team cannot mediate. Therefore, depending on the context, the types of PRO data included will vary. Furthermore, platforms may have different instruments used to capture symptoms, including CTCAE categorical severity grades,⁹ European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 scale,¹⁰ or PRO-CTCAE grades.¹¹ These systems have different

response options, recall periods, and item wording that will require consideration.

Variability in Clinical Management

Capturing patient-reported treatment-related AEs may be affected by the variability in clinical management of AEs, as well as the variability in ePRO alerting and mitigation strategies in response to the reported symptoms and their severity. In the rw setting, ePRO symptom reporting will result in a mitigation response, which may be patient-facing, clinic-facing, or both, in response to reported symptoms. A patient-facing strategy may be an alert sent directly to the patient, on the basis of the patient's reported symptoms, to either suggest how to mitigate a low-grade AE or to seek care for serious AEs. A clinic-facing strategy may be to alert the care team of certain patient-reported symptoms to prompt mitigation. ePRO tools have different alerting algorithms and disparate methods of integration into an alert system for care teams, which may affect downstream clinical management. Differences in the management of AEs (eg, steroid/nonsteroidal immunosuppression, dose reductions, therapy hold, and discontinuation) may affect patient outcomes in the rw.

Patient and Clinician Engagement

Although patients generally adhere to study protocols including PRO reporting requirements in clinical trials,⁶ adherence in the rw setting is variable. A patient may interact with the platform at varying times during treatment and complete symptom reporting at changing frequencies, including choosing not to complete some of the questionnaires sent to them. There is a concern that both intermittent missing ePRO data (eg, misses reporting for 2 weeks, but then starts reporting again) and patients lost to follow-up (eg, stops reporting completely) may influence interpretation of PROs.¹² The absence of a symptom report may not mean the patient did not experience any symptoms.

There is variability in staff follow-up and reminding patients to complete PRO assessments as well as how engaged the clinician is with discussing and acting on reported symptoms with their patient. If a lack of engagement is perceived by the patient, it may affect their willingness to continue to report their symptoms or use the ePRO tool. However, the level of engagement is not easily quantified.

A FRAMEWORK TO HARMONIZE AND EVALUATE THE USE OF ePRO DATA TO IMPROVE PATIENT OUTCOMES

To evaluate whether using ePROs improves patient outcomes, we established a study protocol in which ePRO tools are used to capture symptom data in patients with non-small-cell lung cancer treated with a standard-of-care immunotherapy immune-oncology (IO) therapy, as there are unique immune-related AEs associated with IO therapy.^{13,14} The study protocol is a prospective, observational study with the interventional arm including patients who received IO therapy and had access to an ePRO remote monitoring system and the control arm including patients receiving IO therapy who

did not have access to an ePRO system to communicate treatment-related symptoms. The study protocol outlines an approach for using an external control, but randomly assigning patients to a concurrent control arm is also an option. If using an external control arm, it is important to recognize potential limitations in the ability to harmonize variables between the two arms and account for this in the interpretation of outcomes.

Although this use case is specific to IO therapies and associated symptomatic AEs and immune-related AEs, the common study design provides a generalizable framework that may be applicable to other drug classes and settings. The ePRO vendors reconciled differences in the collection and aggregation of data across tools so that a more homogenized data set could be realized. Additionally, a statistical analysis plan (detailed in the Data Supplement) was also generated to detail the key data variables, end points, and analyses.

Key PRO Data Components to Align Across Platforms

We identified key data components to characterize adherence and use across ePRO tools (Table 1; Data Supplement).

Defining an active user and platform engagement. Platform engagement is defined as any time a patient interacts with the ePRO tool, in any capacity, and should be recorded and analyzed for variability across patients and platforms. The time from treatment initiation to platform engagement, defines an active user. An active user is defined as a patient who completes at least one ePRO symptom reporting activity, either through completion of a questionnaire or ad hoc symptom reporting, within 90 days of treatment initiation, and subsequently completes another symptom reporting activity at a later time. This time frame was chosen because a patient may be overwhelmed at the beginning of treatment and require a reminder or reinforcement to become active. Furthermore, the total length of time the patient engaged with the platform (first engagement to last noted engagement) should also be recorded and analyzed. The totality of engagement is important, as a patient who reports once within 90 days of treatment initiation and then never again should not be considered an active user. Further alignment is needed to create an operational definition for the time frame and engagement for an active user and platform engagement over time.

ePRO symptom reporting and completion rate. To align symptom reporting across ePRO tools, a common terminology for treatment-related AEs should be used. We aligned on use of a CTCAE-like Severity Grade instrument to capture symptomatic AEs, because of the ease of analysis across platforms and because vendors that administer PRO-CTCAE items can convert those scores to CTCAE grades.

The approach for defining ePRO questionnaire completion depended on the platform. Some may only allow submission of the questionnaire if every question is

TABLE 1. Outcome Measures to Evaluate the Impact of ePRO Symptom Reporting

Outcome Measure	Key Data Elements
Duration of treatment	Treatment initiation date
	Treatment discontinuation date
	Treatment last continuing date
	Date of death
	Start date (if applicable) of regimen after initial study treatment
	Last confirmed activity date
ePRO compliance and utilization	ePRO questionnaire completion rate
	ePRO questionnaire send rate
	ePRO ad hoc symptom reporting frequency
Reporting of AEs and symptom burden	Symptom date reported
	CTCAE severity grade
AE intervention and management	Date of active AE intervention
	Type of active intervention
	IO therapy management
	New relevant treatment regimen for AE management (steroid, nonsteroidal immunosuppression)
Health care utilization	Type of health care setting
	Frequency of visits

NOTE. Definitions for the terms included in Table 1 can be found in the Data Supplement.

Abbreviations: AE, adverse event; CTCAE, Common Terminology Criteria for Adverse Events; ePRO, electronic patient-reported outcome; IO, immune-oncology.

answered, while others allow submission of a partially filled questionnaire. ePRO questionnaire completion rate should be reported as the number of entire surveys completed per patient out of the total number sent to the patient. When tools allow incomplete questionnaire submission, the partial completion rate can also be reported as the number of surveys with at least one item completed in the survey, out of the total surveys sent. Since the ePRO tools represented in our group all require patients to complete the entire questionnaire to submit, we did not align on an approach to analyze partially completed questionnaires, but such an approach should be considered.

The frequency at which patients receive scheduled questionnaires may vary between technology platforms and should be reported, as this will inform the completion rate. Computing this information as a rate adjusts for the variability in frequency of questionnaires sent and the time the patient was an active user. The frequency in which a

patient reports a symptom into the ePRO tool outside of the scheduled questionnaires (ie, ad hoc) should also be reported.

Key Patient Outcome Data

Aligning on an approach to capture patient outcome data is important to ascertain the impact of PRO reporting on patient care and outcomes. To understand if ePRO symptom reporting can improve patient outcomes, rw end points, such as rw-time to treatment discontinuation, rw-time to next treatment, and rw-overall survival, can assess the effectiveness of the intervention (Data Supplement). Additionally, clinical variables such as anticancer treatment and supportive care use are critical to contextualize the findings. Finally, the setting (eg, emergency room) and frequency of health care visits should be analyzed to evaluate whether the collection of ePRO data and monitoring has an impact on health care utilization (Data Supplement). Table 1 highlights possible outcomes to assess the impact of ePRO data collections, as well as key data elements to support measurement of the possible outcomes.

In conclusion, in this study, we propose a model framework for use across multiple ePRO tools in the rw setting to evaluate whether using ePROs improves patient outcomes. The protocol design allows for different approaches for ePRO data capture, with the presentation and analysis of the data in a common framework. The framework can be applied to understand how standardized AE mitigation strategies can elevate the quality of rw evidence for determining rw end points. We highlight the methodologic and alignment considerations to enable robust evaluation and use of ePRO data capture. Although ePRO tools may operate differently, we aligned on the key data components and a common framework for harmonized analyses. The findings from the study protocol may allow for more consistent results, improve patient care, and develop the evidence base to support the use of ePRO technology. As demonstrated by the recent release of the Center for Medicare and Medicaid Innovation’s Enhancing Oncology Model, there is a growing push to gradually implement PROs in clinical care for a patient-centric approach.¹⁵ Therefore, creating consensus on approaches to analyzing emerging data in the literature will provide more interpretable data to support patient care and generate evidence to impact drug development by informing clinical trials or labeling changes for approved therapeutics.

AFFILIATIONS

- ¹Friends of Cancer Research, Washington, DC
²Invitae, San Francisco, CA
³Apricity Health, Houston, TX
⁴Dell Medical School, Austin, TX
⁵Clinical Outcomes Solutions, Tuscon, AZ
⁶Varian, a Siemens Healthineers Company, Palo Alto, CA
⁷Navigating Cancer, Seattle, WA
⁸Carevive Systems Inc, Nashville, TN

CORRESPONDING AUTHOR

Brittany A. McKelvey, PhD, 1800 M St NW, Ste 1050 South, Washington, DC 20036; e-mail: bmckelvey@focr.org.

SUPPORT

Supported by Friends of Cancer Research.

AUTHOR CONTRIBUTIONS

Conception and design: Brittany A. McKelvey, Alexandra Berk, Lynda Chin, Stacie Hudgens, Ronan C. O'Hagan, Amila Patel, Debra Wujcik, Mark Stewart, Jeff Allen

Collection and assembly of data: Brittany A. McKelvey, Stacie Hudgens, Amila Patel, Julie Scott, Sam Wang, Debra Wujcik, Mark Stewart

Data analysis and interpretation: Brittany A. McKelvey, Ian Kudel, Amila Patel, Hillary Stires, Sam Wang, Debra Wujcik, Mark Stewart, Jeff Allen

Manuscript writing: All authors

Final approval of manuscript: All authors

Accountable for all aspects of the work: All authors

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

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

Open Payments is a public database containing information reported by companies about payments made to US-licensed physicians ([Open Payments](http://OpenPayments)).

Alexandra Berk

Employment: InVitae

Stock and Other Ownership Interests: InVitae

Lynda Chin

Employment: Apricity Health

Leadership: Apricity Health

Stock and Other Ownership Interests: Sporos Bio (I), Tvardi Therapeutics (I), Apricity Health

Consulting or Advisory Role: Tvardi Therapeutics (I), Asyilia Therapeutics Inc (I), Nirogy Therapeutics (I)

Stacie Hudgens

Consulting or Advisory Role: Clinical Outcomes Solutions (Inst)

Research Funding: Clinical Outcomes Solutions (Inst)

Ian Kudel

Employment: Varian Medical Systems

Stock and Other Ownership Interests: Varian Medical Systems

Research Funding: Varian Medical Systems

Travel, Accommodations, Expenses: Varian Medical Systems

Ronan C. O'Hagan

Employment: Bectas Therapeutics Inc, Apricity Health, Xilio Therapeutics Inc, Nimbus Therapeutics Inc (I)

Leadership: Xilio Therapeutics Inc, Apricity Health, Bectas Therapeutics Inc

Stock and Other Ownership Interests: Xilio Therapeutics

Consulting or Advisory Role: Duke Street Bio, OncoMyx Therapeutics

Amila Patel

Employment: Navigating Cancer, Flatiron Health, Roche/Genentech (I)

Stock and Other Ownership Interests: Roche/Genentech (I), Navigating Cancer

Travel, Accommodations, Expenses: Navigating Cancer

Julie Scott

Employment: Carevive

Hillary Stires

Consulting or Advisory Role: Avalere Health

Sam Wang

Employment: Apricity Health

Leadership: Apricity Health

Stock and Other Ownership Interests: Apricity Health

Debra Wujcik

Employment: Carevive Systems

No other potential conflicts of interest were reported.

REFERENCES

1. Basch E: Patient-reported outcomes: An essential component of oncology drug development and regulatory review. *Lancet Oncol* 19:595-597, 2018
2. Basch E: Patient-reported outcomes—Harnessing patients' voices to improve clinical care. *N Engl J Med* 376:105-108, 2017
3. Basch E, Barbera L, Kerrigan CL, et al: Implementation of patient-reported outcomes in routine medical care. *Am Soc Clin Oncol Ed Book* 38:122-134, 2018
4. Laugsand EA, Sprangers MAG, Bjordal K, et al: Health care providers underestimate symptom intensities of cancer patients: A multicenter European study. *Health Qual Life Outcomes* 8:104, 2010
5. Basch E, Deal AM, Kris MG, et al: Symptom monitoring with patient-reported outcomes during routine cancer treatment: A randomized controlled trial. *J Clin Oncol* 34:557-565, 2016
6. Atherton PJ, Burger KN, Pederson LD, et al: Patient-reported outcomes questionnaire compliance in Cancer Cooperative Group Trials (Alliance N0992). *Clin Trials* 13:612-620, 2016
7. Voon PJ, Cella D, Hansen AR: Health-related quality-of-life assessment of patients with solid tumors on immuno-oncology therapies. *Cancer* 127:1360-1368, 2021
8. Cruz Rivera S, Aiyegbusi OL, Ives J, et al: Ethical considerations for the inclusion of patient-reported outcomes in clinical research: The PRO ethics guidelines. *JAMA* 327:1910-1919, 2022
9. Common Terminology Criteria for Adverse Events (CTCAE) | Protocol Development | CTEP. https://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm
10. EORTC QLQ-C30—EORTC—Quality of Life. <https://qol.eortc.org/eortc-qlq-c30/>
11. Overview of the PRO-CTCAE. <https://healthcaredelivery.cancer.gov/pro-ctcae/overview.html>
12. Bell ML, Fairclough DL: Practical and statistical issues in missing data for longitudinal patient-reported outcomes. *Stat Methods Med Res* 23:440-459, 2014
13. Schneider BJ, Naidoo J, Santomasso BD, et al: Management of immune-related adverse events in patients treated with immune checkpoint inhibitor therapy: ASCO guideline update. *J Clin Oncol* 39:4073-4126, 2021
14. Gu L, Khadaroo PA, Su H, et al.: The safety and tolerability of combined immune checkpoint inhibitors (anti-PD-1/PD-L1 plus anti-CTLA-4): A systematic review and meta-analysis. *BMC Cancer* 19:559, 2019
15. Centers for Medicare & Medicaid Services: Enhancing Oncology Model. 2022. <https://innovation.cms.gov/innovation-models/enhancing-oncology-model>





Real-World Evidence: Leveraging RWD for Insights on Real-World Response

Real-world response endpoints in patients with mNSCLC treated with chemotherapy across real-world datasets.

Brittany Avin McKelvey, Elizabeth Garrett-Mayer, Andrew J. Belli, Thomas D. Brown, Jessica Dow, Janet L. Espirito, Paul Kluetz, Xinran Ma, Andrea McCracken, Pallavi Shruti Mishra-Kalyani, Yanina Natanzon, Danielle Potter, Donna Rivera, Hillary Stires, Mark Stewart, Jeff Allen; Friends of Cancer Research, Washington, DC; American Society of Clinical Oncology, Alexandria, VA; COTA Inc, Boston, MA; Syapse, San Francisco, CA; Tempus Labs, Inc., Chicago, IL; Ontada, Irving, TX; Oncology Center of Excellence, U.S. Food and Drug Administration, Silver Spring, MD; Flatiron Health, New York, NY; Guardian Research Network, Spartanburg, SC; US Food and Drug Administration, Ellicott City, MD; ConcertAI, Cambridge, MA; IQVIA, Horsham, PA; Oncology Center of Excellence, Office of the Commissioner, U.S. Food and Drug Administration, Silver Spring, MD

Background: Response Evaluation Criteria in Solid Tumors (RECIST) based response rate (RR) is used for efficacy evaluation in clinical trials and relies on imaging data collected at specified timepoints for uniform assessment. In routine clinical practice, the method and timing of response assessment can vary, and imaging data from electronic health records (EHR) and other real world (rw) sources may not be available, making RECIST-based assessment of rw-response rate (rwRR) using rw data (RWD) challenging. Friends of Cancer Research formed a multi-stakeholder partnership to assess available data attributes to measure response across RWD sources to inform development of a consistent method for measurement. **Methods:** The study included seven EHR data partners who identified and analyzed a cohort of 1,380 patients (pts) with metastatic non-small cell lung cancer (mNSCLC) treated with first-line platinum doublet chemotherapy, following a common protocol and statistical analysis plan. The availability and frequency of data components to assess response including raw images, radiology imaging reports, and clinician response assessments from provider notes were assessed. Response endpoints measured included rwRR, rw-duration of response (rwDOR), and the association of rwR with rw-overall survival (rwOS), rw-time to treatment discontinuation (rwTTD), and rw-time to next treatment (rwTTNT). **Results:** The availability of data components varied across RWD sources (Table). Images were not widely accessible, thus response was analyzed using clinician response assessments (median proportion of pts evaluable, 77.5%). Of these assessments, the majority relied on imaging interpretation. The median rwRR was 46% with a median rwDOR of 119 days. The table provides median rwTTD, rwTTNT, and rwOS across data sources. **Conclusions:** The rwRR among pts with mNSCLC calculated using the clinician assessment was relatively consistent across all RWD sources, with consistent trends in time to event endpoints. While variability in the availability of data components to assess response was observed, the demonstrated feasibility of response endpoints based on clinician assessment suggests further exploration may inform drug effectiveness evaluation with RWD. Research Sponsor: Friends of Cancer Research (Non-profit).

Group*	Pts Evaluable for rwR (Pts Ev) by Images	Pts Ev by Radiology Reports	Pts Ev by Clinician Assessment	rwRR	Median rwDOR, days (95% CI)	Median rwTTD, days: Responders /Non-Responders (R/NR)	Median rwTTNT, days: R/ NR	Median rwOS, days: R/ NR
A	3.5%	73%	79.5%	42%	115 (86, 199)	142/69	200/100	375/245
B	0.5%	55%	80.5%	53%	133 (108, 182)	128/84	209/98	464/314
C	40.5%	77%	77.5%	46%	146 (102, 210)	147/63	234/93	832/213
D	0%	0%	74%	40%	100 (74,-)	105/70	140/115	614/414
E	79.5%	79.5%	76%	38%	119 (98, 231)	132/48	235/93	474/184
F	0%	66.5%	69%	52%	182 (147, 287)	99/43	219/109	436/353
G	0%	85.5%	88.3%	49%	105 (7, 672)	112/21	198/61	392/86

*n=200 pts except G: n=180.

Scan to view poster



Considerations for Leveraging Real-World Endpoints in Oncology Drug Development

Use of real-world data (RWD) to generate real-world evidence (RWE) can support oncology drug development and regulatory decision-making. There is growing recognition that RWD, when analyzed appropriately, can generate RWE in broader patient populations than are able to be treated in clinical trials to inform medical product effectiveness, safety, and patient outcomes. Unlike traditional clinical trial settings where data are collected per protocol at pre-specified timepoints and reported uniformly for participants, there is significant heterogeneity in RWD within and across data sources. Inconsistent definitions and data missingness present challenges to using real-world (rw) endpoints for measuring treatment effectiveness. Strategies and methodologies for mitigating these challenges and alignment across stakeholders are needed to fully realize the potential of RWD. Friends of Cancer Research (*Friends*) initiated multiple research partnerships^{1,2,3,4} to develop and establish aligned methodologies for measuring rw-endpoints across RWD sources. Based on lessons learned from these research partnerships, a multi-stakeholder working group considered opportunities for using rw-endpoints and developed this resource to optimize use of rw-endpoints in oncology drug development (see table below).

There are multiple intended uses of RWD to support oncology development and may include generating RWE for signal detection to inform clinical development strategies, inform clinical trial design and patient access strategies, or directly be included as part of a regulatory submission. The intended use will impact the applicability of RWD and potential data quality considerations. For example, there should be justification for using RWD as part of a regulatory submission as well as evidence that the selected real-world dataset is fit-for-purpose. Further, caution should be taken when comparing rw-endpoints to clinical trial endpoints, given the inherent limitations of differing populations and measurements. Therefore, this work focuses on alignment across RWD sources, rather than comparison to clinical trial endpoints, through standardized methodologies for assessing rw-endpoints.

The table provides initial considerations for selecting rw-endpoints to measure treatment effectiveness. While rw-endpoints may be leveraged in many ways to support oncology drug development (e.g., rw-overall survival establishing natural history of a specific disease) that may be seen as more a benchmark, the definitions and minimum data elements listed are intended for comparative studies attributing an outcome to a specific treatment (e.g., causal inference). The definitions and data elements provided were jointly developed and implemented across collaborators participating in *Friends'* pilots evaluating rw-endpoints, which focused on patients with metastatic non-small cell lung cancer (mNSCLC) receiving systemic treatments (platinum doublet chemotherapy and/or immunotherapies). While the definitions and data elements listed herein are likely relevant to other solid tumor malignancies, additional data or validation may be needed to support use of these rw-endpoints in other tumor types and indications with disease specific requirements or endpoints. Furthermore, the strengths and limitations noted are informed by the mNSCLC rw-endpoint pilots conducted and may not be generalizable to other disease states.

1. [Establishing a Framework to Evaluate Real-World Endpoints, 2018 Friends of Cancer Research White Paper](#)
2. [The Friends of Cancer Research Real-World Data Collaboration Pilot 2.0: Methodological Recommendations from Oncology Case Studies, Rivera 2022, Clinical Pharmacology & Therapeutics](#)
3. [Real-world Overall Survival Using Oncology Electronic Health Record Data: Friends of Cancer Research Pilot, Lasiter 2022, Clinical Pharmacology & Therapeutics](#)
4. [rw-Response Endpoints in Patients with mNSCLC Treated with Chemotherapy Across rw-Datasets, 2023 ASCO Poster](#)

Working Group Collaborators

Thank you to our working group collaborators for informing the development of this table and considerations for using real-world endpoints in oncology drug development.

Amanda Bruno, Syneos Health, Formerly Bayer Pharmaceuticals

Gil Carrigan, Amgen Inc.

Victoria Chia, Amgen Inc.

Colleen Costello, Sanofi

Janet Espirito, Ontada

Laura Fernandes, COTA, Inc.

Elizabeth Garrett-Mayer, American Society Clinical Oncology

Shrividya Iyer, Eisai

Monika Izano, Syapse

Maria Karasarides, Bristol Myers Squibb Company

Sudeep Karve, Abbvie

Beata Korytowsky, Mirati

Mark Lanasa, BeiGene

Yanina Natanzon, ConcertAI

Irene Nunes, Flatiron Health

Vivek Pawar, EMD Serono

Danielle Potter, IQVIA

Lynn Sanders, Takeda Pharmaceutical Company

Regina Schwind, Tempus Labs, Inc.

Alessandria Strübing, Daiichi-Sankyo

We thank colleagues at the U.S. FDA for their input and collaboration. This document reflects discussions that occurred among stakeholder groups on various topics. This document should not be construed to represent FDA's views or policies.

rw-Endpoint and Definition	Minimum Data Elements Needed	Strengths & Limitations	Statistical Analysis Considerations
<p>rw-Overall Survival (rwOS) Length of time from the index treatment date to the date of death; for patients without a date of death, patients will be censored at the date of last structured recorded clinical activity, or end of follow-up period, whichever occurs earliest.</p>	<ul style="list-style-type: none"> Date of index treatment initiation Date of death or end of follow-up 	<p>Strengths</p> <ul style="list-style-type: none"> Objectively defined. <p>Limitations</p> <ul style="list-style-type: none"> Missingness of mortality information, which may not be random and could lead to biased estimates. Insufficient follow up time can also lead to a high proportion of censored patients which may overestimate survival. Survival attributed to index treatment may be impacted by subsequent activities or therapies. These subsequent activities or therapies may be unavailable in EMR (e.g., start dates for oral medications may be difficult to obtain) due to incompleteness of data capture. Real-world mortality information may not include cause of death to understand disease specific survival. 	<ul style="list-style-type: none"> Capture median rwOS as well as landmark rwOS (e.g., 1 year and 5 year). Additional data elements noting subsequent activity (subsequent therapies, etc.) or intercurrent events may be used to provide context to the rwOS endpoint. Reduce immortal time bias (i.e., stratifying rwOS curves on factors that are determined after date of index treatment initiation).
<p>rw-Progression-Free Survival (rwPFS) Length of time from the index treatment date to the date of progression event or date of death. Patients without a progression event or date of death will be censored at the date of last structured recorded clinical activity reporting disease status, or end of follow-up period, whichever occurs earliest.</p>	<ul style="list-style-type: none"> Date of index treatment initiation Date of progression event through assessment of tumor response by clinician-based assessments Date of death or end of follow-up Date of index treatment discontinuation* Date of next treatment initiation* <p>*Optional, to attribute progression event to index treatment</p>	<p>Strengths</p> <ul style="list-style-type: none"> Less follow-up time is needed than rwOS, which may limit data missingness concerns. Captures more direct effect of treatment activity on disease. <p>Limitations</p> <ul style="list-style-type: none"> Subject to interval censoring bias, i.e., assessments may occur at different time intervals and using different methodologies or modalities. Length of intervals may also be related to response to therapy. Assessments may occur outside of available data source and lead to data missingness. Capture of rwPFS based on clinician-based assessments is subjective (variable) and not based on RECIST criteria or have confirmation of progression. 	<ul style="list-style-type: none"> Capture median rwPFS as well as landmark (e.g., 6 months and 1 year). Account for interval censoring in data analysis. Present breakdown of the type of PFS event: n(%) patients with progression event against n(%) patients with treatment discontinuation or n(%) patients with next line treatment start to understand the progression information captured. Consider sensitivity analyses that assess rwPFS based on the type of progression data (e.g., tumor measurements from imaging, symptomatic progression). Consider sensitivity analyses censoring patients who had no sign of progression but switched therapies.

rw-Endpoint and Definition	Minimum Data Elements Needed	Strengths & Limitations	Statistical Analysis Considerations
<p>rw-Response (rWR)</p> <ul style="list-style-type: none"> • Occurrence of a rWCR or rWPR after index treatment initiation during the study period among all patients. • This is often assessed as a rWR rate (rWRR), which is the proportion of patients with a rw-best overall response (rWBOR) of rWCR or rWPR. • $rWCR > rWPR > rWSD$ rWPD. 	<ul style="list-style-type: none"> • Assessment of tumor response by clinician-based assessments and date of assessment • Date of index treatment initiation • Date of index treatment discontinuation 	<p>Strengths</p> <ul style="list-style-type: none"> • Direct measurement of drug antitumor activity. • Less follow-up time is needed than rWOS and possibly rWPFS (depends on rWRR of drug). <p>Limitations</p> <ul style="list-style-type: none"> • For clinician-based assessments, subjective measure due to lack of standardized assessment framework in routine practice, including timing and frequency of assessment. • Absence of confirmatory scans on response in clinical practice. • Subject to observer or information bias. • Clinician-based assessments from one assessment to the next may use the last assessment as the new comparator, rather than the pre-treatment baseline. • Assessments may not be appropriately adjusted based on if the patient received any surgical resection or radiotherapy during index treatment. 	<ul style="list-style-type: none"> • Consider sensitivity analyses of patients with both images or image reports (to conduct a RECIST-like assessment) and clinician assessment to evaluate concordance of response. • Consider analysis of interval timing and frequency of clinician-based assessments to inform findings.
<p>rw-Duration of Response (rWDOR)</p> <p>The length of time from the date of the first documented assessment of rWCR or rWPR after the index date to the date of the first subsequent documented assessment of rWPD, rWMR or death, whichever comes first. For patients without rWPD, rWMR, or death, the patient will be censored at their last known response assessment of rWCR, rWPR, or rWSD, or the date of treatment discontinuation, whichever comes first.</p>	<ul style="list-style-type: none"> • Date of index treatment initiation • Date of first assessment of rWCR or rWPR • Date of first subsequent assessment of rWPD, rWMR, or death • Date of last assessment of rWCR or rWPR • Date of index treatment discontinuation • Date of next treatment initiation (Optional, if missing index treatment discontinuation) 	<p>Strengths</p> <ul style="list-style-type: none"> • Provides understanding of response durability. <p>Limitations</p> <ul style="list-style-type: none"> • Subject to interval censoring bias, i.e., assessments may occur at different time intervals and using different methodologies or modalities. • For clinician-based assessments, subjective measure due to lack of confirmatory scans and varying methodologies. • Requires various data points that may not be captured adequately for assessment. 	<ul style="list-style-type: none"> • Capture durable 6-month rWRR: The proportion of patients with at least one assessment of rWCR or rWPR who have not had an assessment of rWPD or rWMR or discontinuation of therapy within 6 months after the first documented assessment of rWCR or rWPR. • Capture median rWDOR and landmark (e.g., 3, 6, 9 months).

rw-Endpoint and Definition	Minimum Data Elements Needed	Strengths & Limitations	Statistical Analysis Considerations
<p>rw-Time to Treatment Discontinuation (rwTTD) Length of time from the index treatment date to the date the patient discontinues treatment. The treatment discontinuation date is defined as the last administration or non-cancelled order of the therapy. Discontinuation is defined as having a subsequent systemic therapy after the index therapy, having a gap of more than 120 days with no systemic therapy following the last administration prior to treatment completion, or having a date of death while on the therapy. Patients without a discontinuation will be censored at their last known usage of study treatment.</p>	<ul style="list-style-type: none"> • Date of index treatment initiation • Date of index treatment discontinuation • Date of death or end of follow-up 	<p style="text-align: center;">Oral Drugs</p> <p>Strengths</p> <ul style="list-style-type: none"> • Integrates total amount of time a patient is treated on therapy, regardless of reason for discontinuation, whether due to effectiveness or tolerability. • Not subject to bias associated with variable tumor burden assessments (e.g., time intervals, assessment methodologies). • May be associated proxy for PFS. <p>Limitations</p> <ul style="list-style-type: none"> • Treatment cycles may vary, both on a treatment and a per patient basis, impacting the ability to define discontinuation. • Date of treatment discontinuation is not always (and often not) available, and patient adherence may be unknown. Therefore, might need to rely on assumptions based on start dates or other algorithms. • Discontinuation may be due to tolerability or causes other than treatment ineffectiveness. 	<ul style="list-style-type: none"> • In combination therapies, discontinuation should be considered when both therapies are discontinued (one therapy in the combination can be discontinued and still be considered on index treatment), however, timing of discontinuation of the one therapy in the combination should be noted. • Consider sensitivity analyses on the 120-day period, as this criterion may differ based on clinical opinion in the disease of interest or data source.
	<ul style="list-style-type: none"> • Date of index treatment initiation • Date of last administration • Date of death or end of follow-up 	<p style="text-align: center;">Infused Drugs</p> <p>Strengths</p> <ul style="list-style-type: none"> • Integrates total amount of time a patient is treated on therapy, regardless of reason for discontinuation, whether due to effectiveness or tolerability. • Often based on structured data and easier to implement in an EMR system. • Not subject to bias associated with variable tumor burden assessments (e.g., time intervals, assessment methodologies). <p>Limitations</p> <ul style="list-style-type: none"> • Discontinuation may be due to tolerability or causes other than treatment ineffectiveness. 	<ul style="list-style-type: none"> • In combination therapies, discontinuation should be considered when both therapies are discontinued (one therapy in the combination can be discontinued and still be considered on index treatment), however, timing of discontinuation of the one therapy in the combination should be noted. • Consider sensitivity analyses on the 120-day period, as this criterion may differ based on clinical opinion in the disease of interest or data source.

rw-Endpoint and Definition	Minimum Data Elements Needed	Strengths & Limitations	Statistical Analysis Considerations
<p>rw-Time to Next Treatment (rwTTNT) Length of time from the index treatment date to the date the patient received their next systemic therapy (next-line therapy) or date of death. If patient has not received subsequent treatment or has not died, patients will be censored at their last known activity or end of follow-up.</p>	<ul style="list-style-type: none"> • Date of index treatment initiation • Date of index treatment discontinuation • Date of next line treatment • Date of death or end of follow-up 	<p>Strengths</p> <ul style="list-style-type: none"> • A proxy for disease control or potential benefit (e.g., duration of effect). • Integrates total amount of time a patient is treated on therapy, regardless of reason for discontinuation, whether due to effectiveness or tolerability. • Captures possible response durability of initial treatment. • Advantage over rwTTD given the data availability of timing of next therapy initiation compared to discontinuation. • Not subject to bias associated with variable tumor burden assessments (e.g., time intervals, assessment methodologies). <p>Limitations</p> <ul style="list-style-type: none"> • Endpoint as defined is specific to next line systemic therapy and is not inclusive of other interventions such as surgery or radiation which could result in bias. • Censoring is likely not independent of prognosis (violation of censoring assumption). • Missingness of data if patients receive treatment outside of the system. 	<ul style="list-style-type: none"> • Consider analyses that account for intercurrent events, if data are available.

EMR, electronic medical record; rwCR, real-world complete response; rwPR, real-world partial response; rwPD, real-world progressive disease; rwMR, real-world mixed response; rwSD, real-world stable disease.



Innovative Drug Development: Insights for Advancing Oncology Trials and Therapies

Maximizing Data from Academic-Led Studies for Regulatory Decision-Making

Introduction

Clinical trials sponsored or conducted by academic investigators or through clinical trial network groups are an important component of the oncology clinical research landscape. These trials are integral to advancing our knowledge of cancer and improving patient care. Industry-sponsored trials are most often the primary data source for regulatory submissions to support new indications or other label updates (e.g., dose adjustments, safety updates). However, additional sources of data exist on the safety and effectiveness of a drug that may support regulatory submissions and approvals. One additional source that contributes to the scientific understanding around the benefits and risks of cancer therapies comes from trials conducted by academic investigators and clinical trial network groups, broadly referred to as “academic drug development trials” in this white paper.

Academic drug development trials offer a unique opportunity to address four critical aspects of cancer research and treatment:

1. They play an important role in generating additional data that can address key regulatory questions, including post-marketing commitments related to safety, alternative dose or administration schedules.
2. They may target rare cancers, underrepresented patient groups, or patients excluded in the pivotal trial (e.g., older adults or those with organ dysfunction), filling evidentiary gaps and expanding treatment options for additional patient populations.
3. They can provide access to clinical trials to more diverse patient populations through community networks, enhancing the representativeness of clinical findings.
4. They often focus on pressing scientific questions, such as exploring novel combination treatments, driving innovation in cancer therapy.

Collaborations between industry and academic drug development trial investigators can harness these opportunities, which can advance research, drug development, and patient care. These collaborations can take various forms, including funding or providing experimental agents for clinical trials, sharing expertise, providing access to patient populations, contributing resources to accelerate cancer research, and leveraging well-established infrastructures. For example, the

Thank You to Our Contributors

Sarang Abhyankar, Eli Lilly & Co

Sandra Casak, U.S. FDA

Scot Ebbinghaus, Merck & Co., Inc.

Doug Fecteau, Johnson & Johnson Innovative Medicine

Anmarie Galli, GSK

Viktoriya Ilaria, Eli Lilly & Co

Percy Ivy, NCI

Abigail Johnston, Patient Advocate

Tarik Khaznadar, F. Hoffmann-La Roche, Basel, Switzerland

Kristina Laumann, Mayo Clinic

Seth Miller, GSK

Flora Mulkey, U.S. FDA

Nancy Nair, Johnson & Johnson Innovative Medicine

Christy Osgood, U.S. FDA

Russ Palmer, EMD Serono

Mark Stewart, Friends of Cancer Research

Kathleen Winson, Genentech

Sunita Zalani, Merck & Co., Inc.

This document reflects discussions that occurred among stakeholder groups on various topics. This document should not be construed to represent FDA's views or policies.

National Cancer Institute's National Clinical Trial Network (NCI NCTN) is one such infrastructure comprised of five US network groups (formerly known as cooperative groups), encompassing collaborative networks of researchers, clinicians, and institutions that conduct large-scale, multi-center clinical trials. The NCI NCTN serves as a valuable resource for coordinating and supporting cancer clinical trials by engaging in independent research initiatives and trials. By industry working directly with these network groups or even individual academic investigators with patient consultation, there can be greater alignment on shared research goals in specific therapeutic areas or patient populations to ultimately contribute to improving patient care and the development of new cancer treatments. Indeed, the U.S. Food and Drug Administration (FDA) Oncology Center of Excellence (OCE) has highlighted the NCI network as a potential opportunity to reach patients with clinical trials to obtain a patient population more representative of the U.S. population to support regulatory submissions, either alone or as a part of a larger multi-regional global clinical trial.¹

Yet, challenges exist in leveraging data generated from academic studies for regulatory purposes. Not all academic studies are intended for regulatory use. For those that may potentially be used to support regulatory decision-making, industry partners and those conducting the trial should align on study designs and optimize data collection practices. For academic drug development trials where the industry partner indicates an interest in the potential use of the data to support registration or labeling updates, considerations should be given to enable proactive planning of the data collected and align with expectations of regulatory submissions. Submitting data to the FDA as well as other health-regulatory authorities for regulatory decision-making requires data to be comprehensive and formatted in well-defined and internationally recognized standardized ways.² This can be difficult to achieve if statistical designs, study conduct, data collection methodologies, and other processes do not meet the expectations of the FDA and other health-regulatory authorities.

Due to increased interest in leveraging data from academic drug development trials for regulatory submissions, Friends of Cancer Research (*Friends*) brought together key stakeholders from industry, academia, advocacy, and government to characterize challenges encountered in this space and propose ways to enhance the use of data from these studies. This white paper aims to address factors impacting the use of data from academic drug development trials, with a focus on streamlining processes to expedite results, ultimately advancing oncology drug development and care for patients.

Factors That Can Impact Use of Data from Academic Drug Development Trials

An industry partner's decision to engage in a collaboration with an academic group or investigator for a registrational trial can be influenced by various factors such as the study's prioritization within the overall clinical development plan, the study's design and complexity, intellectual property rights, as well as timeline-related considerations; however, the primary focus of this white paper is addressing issues that arise when industry partners pursue collaborations to use data produced from academic drug development trials for regulatory purposes. Several methodological, operational, and communication-related challenges have been identified as barriers affecting the use of data from academic drug development trials for regulatory decision-making, impacting both industry partners and health-regulatory authorities.

Lack of Early Engagement with FDA

One prominent issue is the lack of early engagement with the FDA by those conducting or supporting academic drug development trials, which can impact the ability to support registration. This is often an issue where trials are not clearly identified as potentially label-enabling at study inception due to ambiguity by the industry partner and can lead to skipping pre-study engagement with FDA resulting in downstream issues that may not be able to be mitigated mid- or post-study. Lack of early engagement can result in study designs that do not meet regulatory expectations or missing data points that impact the content and/or quality of data packages necessary to meet regulatory requirements and support regulatory decision-making.

Varying Data Capture and Monitoring Requirements

The way in which data are collected in an academic drug development trial as compared to an industry trial can vary. Industry sponsored trials follow the Clinical Data Interchange Standards Consortium (CDISC), a registrational-compliant format for data collection, programming, and analysis, which enables more streamlined regulatory submissions and more efficient FDA review of patient-level data. Academic drug development trials often employ varying data collection methods that do not always align with the intent of producing the required format for regulatory review including Study Data Tabulation Model (SDTM), which provide the raw data for FDA's review, and Analysis Data Model (ADaM) datasets, which facilitates the Agency's ability to replicate study results. Standardized medical terminology for safety data and approaches to safety data collection and reporting may also differ between industry-sponsored trials and academic institutions. Furthermore, data monitoring and quality aspects can vary. As such, these variations can result in a time and resource-intensive process of cleaning, reviewing, and programming of data generated from academic drug development trials to achieve regulatory-compliant SDTM source and ADaM datasets.

Limited Data Access During Trial Conduct

Trials conducted either within the NCI NCTN or with individual US network groups often have policies that only allow for limited interim data sharing between those conducting the trial and their industry partner. Data sharing policies are in place to maintain the integrity of the study but as a consequence restrict real-time access to data by the industry sponsor. Academic investigators, US network groups, and NCI have policies for data locks to protect the integrity of data, but these may be misaligned with steps to proactively identify potential issues and needs for data review. Resource-intensive tasks to clean and map data can delay primary analyses and interpretation and compilation of a submission-ready data package, thereby resulting in a time lag for getting important therapies to patients.

Together, several factors can lead to delays (up to 12+ months) in submitting data to the FDA for regulatory review (Figure 1). The timeline presented is one hypothetical scenario of event timing for preparing data from an academic drug development trial for regulatory use. The time from the study's last patient last visit to data sharing to submission to the FDA can be quick or delayed depending on a number of factors including 1) whether the study had registration-intent from the beginning or whether that determination was made after the data readout; 2) the level of planning, communication, and collaboration between those conducting the study, the industry partner, and the FDA; 3) the resource availability of the group conducting the trial; 4) the data

cleaning and reviewing process; and 5) whether additional data collection is required.

Addressing these challenges should involve both short-term measures and broader system-wide initiatives aimed at standardizing data management processes and enhancing collaboration between academic researchers, industry partners, and regulatory agencies. The strategies noted below are aimed at reducing the timeframe between the completion of an academic drug development trial and a regulatory submission to the FDA, ultimately speeding access of these therapies for patients.

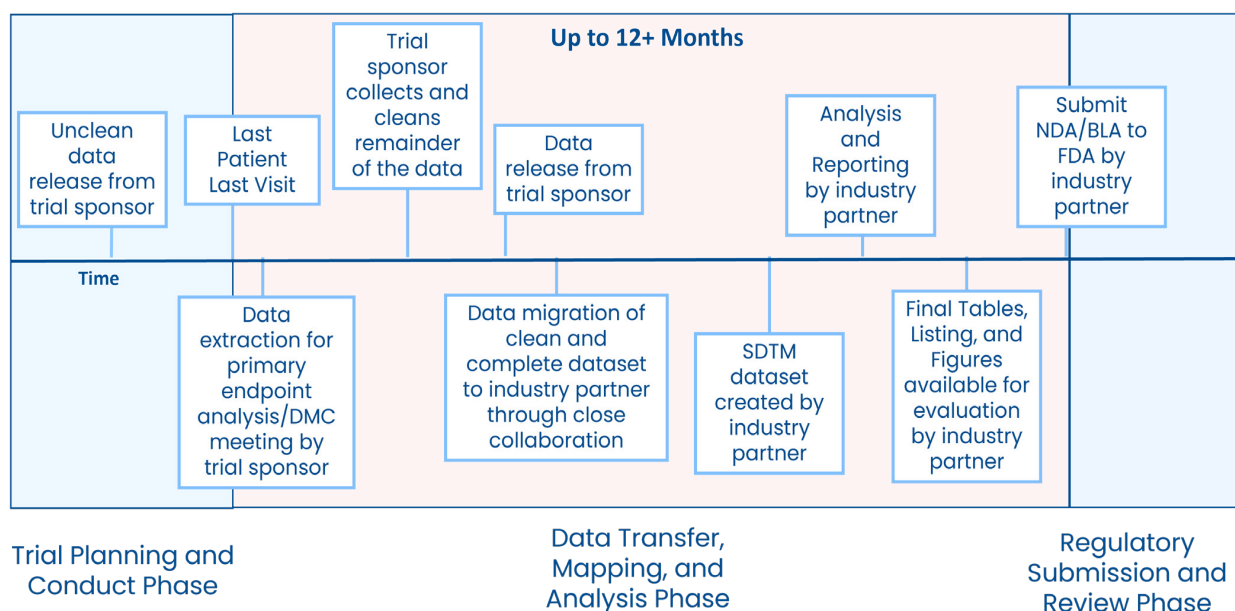


Figure 1. Example of a timeline of typical events associated with preparing data from an academic drug development trial for regulatory use.

Strategies to Reduce Time Between Completion of Academic Drug Development Trials and FDA Submission

Initiate Early Commitment and Communicate Registrational Intent

Industry partners should establish internal processes for identifying the inclusion of academic drug development trials within their development plans. This process could lead to or support potential registrational intent thereby facilitating proactive planning versus waiting until the study reads out to determine registrational potential and enabling industry partners to be more systematic and intentional with those conducting the academic drug development trial, including prioritizing initial planning conversations. Early commitment can help trigger discussions around requirements and expectations of all parties engaged, including availability of required systems, data collection requirements, data sharing needs and platforms, and regulatory engagement strategies. This proactive approach will increase the likelihood that the data from the trial will

meet regulatory standards and allow the industry partner to better plan for timely data transfers to support critical regulatory submission activities.

Initiate Discussions Early with FDA

When an academic drug development trial is identified as having the potential for registrational intent after review by the industry partner, joint meetings should be requested between the appropriate FDA review division and those conducting the trial (e.g., the academic investigator/network group, industry partner, NCI) with patient consultation to discuss the scientific rationale and strategy of the proposed trial design, clinical endpoints, and statistical analysis plan. This will also provide an opportunity for discussion and feedback on operational elements (e.g., data collection and cleaning) that may not align completely with standard industry practices and are not routinely discussed with FDA. Subsequently, those conducting the trial can align study designs, methodologies, and data collection strategies with regulatory requirements. Discussions can focus on key phases of the study:

- Study concept development: Conduct joint meetings to discuss the protocol, study design, endpoints, safety reporting considerations, and statistical analysis plan for potential registration studies; Discuss case report forms, collection of data (e.g., blinded independent central review, adverse event terminology)
- Study ongoing: Conduct joint meetings to discuss data cleaning, data transfer and mapping, database lock planning
- Study conclusion: Conduct joint meeting to discuss the results and dossier preparation, final data transfer, dataset, and tables, listings, and figure generation

These interactions can be achieved through Prescription Drug User Fee Act (PDUFA) meetings, and content and format meetings. Overarching feedback can also be received through workshops and collaborative forums where academic researchers, industry partners, and FDA can have informational exchanges to share learnings, expectations, and best practices for success. These interactions should ideally introduce opportunities to discuss and gain early agreement on the optimal type/amount of data needed to address the specific scientific question, leading to more compliant downstream regulatory submissions and reducing submission delays. By involving the FDA prior to study initiation, potential regulatory concerns that may impact the readiness of the data for regulatory submission can be identified and mitigation strategies discussed. In addition, there are numerous FDA/OCE guidance documents aimed at providing insight to potential applicants on topics, such as endpoint selection, typical analyses expected in specific disease areas, and other considerations when planning a trial for submission.³

In instances where a trial starts as a non-registrational trial and later intends to support regulatory decisions, it will be necessary to identify mechanisms for mid-study check-ins. Potential future FDA guidance documents specific to academic drug development trials could further clarify expectations and types of meetings that can be leveraged for these interactions.

Establish a Regulatory Track for Studies with Potential Registrational Intent

In instances where academic drug development trials are identified as having potential for registrational intent, a “regulatory track” could be established within the network group or NCI NCTN.

The regulatory track would trigger certain expectations for data lock procedures, study protocols, interactions with the FDA, and outline data requirements to meet regulatory submission needs. Additional needs for the study may also be agreed upon by those conducting the study and the industry partner. The primary objective is to ensure uniformity in data collection methodologies, encompassing crucial aspects such as demographic information, patient outcomes, disease characteristics, treatment specifics, and adverse event documentation. Minimum expectations around what and how data is collected and the types of questions that need to be addressed should be outlined. Moreover, these guidelines should include standardized definitions, particularly for adverse event categorization and the criteria for defining treatment response and endpoints. Notably, the level of safety data collection and the use of verbatim terms in academic trials often differ from what is required for FDA regulatory submissions. Harmonizing these practices will help to ensure that data can be appropriately mapped to meet the stringent regulatory requirements which are in place to ensure the safety of these agents, thereby expediting the evaluation and approval of promising therapies.

Evaluate Data Sharing Policies for Studies with Potential for Regulatory Intent

Earlier evaluation of data quality and formatting could enable more proactive efforts to clean and map data, but current policies can limit access to data during trial conduct. Appropriate data transfers between trial collaborators while the trial is ongoing can enable an iterative data review process that accelerates the identification of potential issues to enable programming for SDTM-compliant mapping to occur and would increase the overall data quality and scientific rigor of the trial. This could be accomplished by establishing secure blinded data-sharing policies that allow for the exchange of relevant data throughout the trial's lifecycle while also maintaining appropriate trial oversight, patient and trial confidentiality, and data and statistical integrity. Alternatively, the use of third-party organizations that can engage with those conducting the academic drug development trials for access to blinded data for the purposes of data cleaning and/or SDTM mapping could be explored if current policies or concerns around data integrity are encountered when sharing directly with an industry partner.

Establish a Streamlined Process for Submitting Data to the FDA

Traditionally, complete datasets from academic drug development trials are submitted to the FDA, which can be time-consuming and resource-intensive to evaluate if the data are not captured with the intent of conforming to regulatory standards. In these instances, one approach for more efficient data submissions could involve providing an abbreviated or summary data package to the FDA earlier for review with full datasets to follow. With FDA agreement, initial submissions could prioritize the mapping and transfer of key subsets of data, including initial submissions with primary/secondary endpoints and select safety data. This approach aims to improve efficiency, reduce redundant efforts, and accelerate the review process while maintaining data integrity and regulatory compliance. The Real-Time Oncology Review (RTOR) program at the FDA provides some general principles and practices that can be adapted to help structure submission of data from academic drug development trials. Specifically, this framework could include several necessary components:

1. Pre-submission activities to discuss the data that will be included in the application,
2. Submission of initial abbreviated data that includes the clinical study report and datasets,

3. Review of the initial data including the study design, efficacy data, and safety data, and
4. Submission of the final data that includes any additional data that was not included in the initial submission.

Conclusions

Leveraging academic drug development trials presents significant opportunities for enhancing evidence generation and bringing innovative therapies to patients faster. While efforts to implement standardized practices across all academic drug development trials are important, near-term opportunities center around improving collaboration and coordination of academic drug development trials intended for regulatory decision-making to reduce the delays from study readout to FDA submission, which can slow access to potentially practice-changing trial results. Given resource limitations for many academic drug development trials and significant efforts to streamline data collection and workflows for site staff, it is important to recognize that it may not be feasible for those conducting academic drug development trials to program every study to meet regulatory requirements (e.g., SDTM/CDISC format) due to limited resources and differing objectives. As such, industry partners should consider long-term partnerships with academic investigators or US network groups that allow for more sustained support for these efforts and help develop the infrastructure for these types of studies.

Addressing challenges through near-term and longer-term solutions will enable a more efficient and impactful process for leveraging academic drug development trials for regulatory use. In the future, it is important to establish early collaboration with the FDA to synchronize data collection and analysis approaches, evaluate data sharing guidelines, and specify preferred data formats for academic drug development trials.

References

1. Derrick Gingery. US FDA's Pazdur Suggests NCI Should Help Sponsors Enroll Clinical Trials For Innovative Drugs. Pink Sheet. November 18, 2022. <https://pink.citeline.com/PS147341/US-FDAs-Pazdur-Suggests-NCI-Should-Help-Sponsors-Enroll-Clinical-Trials-For-Innovative-Drugs>
2. Data Standards Catalog. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/data-standards-catalog>
3. Oncology Center of Excellence Guidance Documents. <https://www.fda.gov/about-fda/oncology-center-excellence/oncology-center-excellence-guidance-documents>

Accelerating The Development of Engineered Cellular Therapies: A Framework for Extrapolating Data Across Related Products

Introduction

Engineered cellular therapies^a have emerged as a new treatment pillar and are poised to change the therapy landscape for patients with serious or life-threatening malignancies. To date, the U.S. Food and Drug Administration (FDA) has approved six autologous cell-based immunotherapies, each showing remarkable activity in certain hematologic malignancies. However, considerable scientific and operational obstacles must be overcome to enable broader application of this therapeutic approach in additional cancers, including solid tumors, and advance emerging approaches such as allogeneic and in vivo targeted cell engineering. Novel scientific approaches that build on current products and enhance product safety and efficacy, overcome biological limitations, and reduce manufacturing costs and time are necessary to develop the next generation of engineered cellular therapies.

During engineered cellular therapy development, sponsors investigating an autologous chimeric antigen receptor (CAR) T-cell product may also test different versions of the primary product (e.g., an altered CAR protein domain to enhance CAR T-cell activity, additional functional enhancements, a CAR-T cell derived from an alternative starting material, a more purified cell subtype) in parallel or in tandem. As such, leveraging data from related product versions combined with prior platform knowledge may support a more streamlined and effective development strategy across product versions and for future product versions. Accordingly, adaptations of clinical development models and regulatory frameworks are needed to support more flexible development strategies and allow for product improvements based on empirical learnings. The approach should consider the totality of evidence collected from preclinical research, clinical trials, and characterization of the manufactured product as well as any available published literature or post-marketing surveillance from related products to inform the safety and biological activity of iterative product versions. Ultimately, this strategy can optimize the development of these therapies and bring them to patients in a rapid, safe, and efficient manner.

^aThis document primarily focuses on genetically engineered cell-based gene therapies. The term engineered cell therapies includes a variety of immune therapies, such as T-cell receptor (TCR) or chimeric antigen receptor (CAR) based tumor infiltrating lymphocytes (TILs) and other T-cell based therapies.

Thank You to Our Contributors

Marc Better, Pharmefex Consulting

Vicki Coutinho, Independent (previously affiliated with Takeda Pharmaceutical Company)

Miriam Fuchs, Novartis

Shalini Gidwani, Allogene Therapeutics

Carsten Goessl, GSK

Jane Healy, Merck & Co., Inc

Julie Jadowsky, University of Pennsylvania

Jonathan Jazayeri, Kite, A Gilead Company

Michael Kalos, Next Pillar Consulting

Wen Liu, Lyell Immunopharma

Laura Pearce, Canadian Cancer Trials Group (previously affiliated with GSK)

Shari Pilon-Thomas, Moffitt Cancer Center

Naik Snehal, Spark Therapeutics

Monica Veldman, Genentech, Inc.

Judy Vong, A2 Bio

Susan Weinbach, Bristol Myers Squibb Company

Jennifer Yohrling, Janssen R&D

We thank Dr. Ingrid Markovic and colleagues at the Centers for Biologic Evaluation and Research at the FDA for their thoughtful review. This paper reflects discussions that occurred among stakeholder groups on various topics. This paper should not be construed to represent FDA's views or policies.

The FDA continues to refine guidance to increase efficiencies and facilitate development of engineered cellular therapies and has released several guidance documents focused on informing development and streamlining regulatory processes for novel cellular and gene therapies.^{1,2,3} Specifically, FDA outlines an innovative approach to investigate different versions of a cellular or gene therapy in a single umbrella trial during early clinical evaluation, rather than the traditional approach of initiating individual trials for each product version. FDA provides several examples of changes that result in different versions (see Appendix), which would require separate investigational new drug applications (INDs). Within these different versions, one version would be the primary version with the “Primary IND” containing the clinical protocol, the chemistry, manufacturing, and controls (CMC), and pharmacology/toxicology information. Each of the “Secondary INDs” would cross-reference the clinical information in the Primary IND and contain additional CMC and pharmacology/toxicology information specific to each of the secondary versions. The recent passage of the Food and Drug Omnibus Reform Act of 2022 also includes a provision for FDA to create a designation program for platform technologies that have the potential to be used with more than one drug and may be eligible for certain expedited development or review actions.⁴ Within this program, sponsors may “reference or rely upon data and information” from a previous drug/biologics licensing application incorporating the same platform technology.

As our understanding of engineered cellular therapies continues to improve and FDA’s expectations for the types of data necessary to support product changes are clarified, opportunities for leveraging data from product versions across the stages of development will likely increase. Extending the concept of cross-referencing information from one product to a related product version could enable informed trial designs and refined data collection to improve operational and developmental efficiencies as well as streamline regulatory data packages. Because there is not a “one size fits all” approach for extrapolating data across product versions, a risk-based approach can help evaluate when, to what extent, and how data from one product can support development of another version. This white paper provides a conceptual, risk-based approach to leverage the totality of evidence—available manufacturing, product quality, analytical characterization, and non-clinical and clinical knowledge—to support development of multiple product versions, minimize redundant data collection, and optimize development of next generation engineered cellular therapies.

Leveraging Data Across Product Versions to Support Clinical Development

Data extrapolation to advance new versions of investigational products has occurred for several decades across therapeutic classes due to an understanding of the biology, mechanism of action, and manufacturing processes (Appendix Supplemental Table 1). Lessons learned from leveraging the totality of evidence in other therapeutic classes to support inferences for new product versions or indications provide a basis for data extrapolation in engineered cellular therapies.

The extent to which data can be meaningfully extrapolated from a primary product to related engineered cellular therapy products depends on the type of modification (including prior knowledge of its impact on related constructs) and phase of development of the primary and secondary products, as well as how “similar” the two versions are to each other. Notably, a case-by-case assessment should be done to determine if it may be considered the “same” therapeutic.⁵ The appropriateness of data extrapolation between two product versions may vary throughout the product lifecycle (e.g., first-in-human studies, early phase, late phase, and post-market) and across product versions.

YESCARTA® and TECARTUS® provide an example of extrapolation in engineered cellular therapy products. The secondary product, TECARTUS®, shares the same anti-CD19 CAR construct, the vector used in the manufacturing, the final drug product composition, and cryopreservation method as YESCARTA®, the primary product. However, TECARTUS® has a modified manufacturing process, which includes a white blood cell enrichment process. Nonclinical, clinical, and certain CMC data were extrapolated from YESCARTA® to support development and approval of TECARTUS® (Table 1). The concept of leveraging prior data and the totality of evidence seen in this example can be extended to other engineered cellular therapy products in development.

Table 1: Use of Data Extrapolation between YESCARTS® and TECARTUS® CAR T-cell Therapies Targeting CD19

Publicly available FDA review documents include examples where data extrapolation has been used in the development and approval of CAR T-cell therapies.^{6,7,8}

Data Type Extrapolated	Data Extrapolation Noted in FDA Review Documents
Non-Clinical Data	<ul style="list-style-type: none"> Due to several identical features between YESCARTA® (axicabtagene ciloleucel) and TECARTUS® (brexucabtagene autoleucel)—the same anti-CD19 CAR construct, the vector used in the manufacturing, the final drug product composition and cryopreservation method—further safety pharmacology, pharmacokinetic, toxicology, tumorigenicity, and genotoxicity studies were not required for TECARTUS®.
Clinical Data	<ul style="list-style-type: none"> The starting dose in the clinical study (ZUMA-2) to assess the safety and efficacy of TECARTUS® in subjects with relapsed/refractory (r/r) mantle cell lymphoma (MCL) was selected based on the prior explored dose of YESCARTA® in subjects with r/r MCL in the same clinical study. Therefore, the typical dose escalation cohorts, inter-patient intervals and stopping rules were minimized. Due to several identical features existing across the two product versions, including the anti-CD19 CAR expressed, the vector used in manufacturing, and the similar safety profiles of cytokine release syndrome (CRS) and neurological toxicities, the FDA supported a combined risk evaluation and mitigation strategies (REMS) program for YESCARTA® and TECARTUS®.
CMC Data	<ul style="list-style-type: none"> Due to several similarities in the manufacture (vector construct, vector manufacturing process, product manufacturing process, controls, formulation, container closure system validation, storage, equipment, and same manufacturing sites) of the two product versions, several relevant sections of CMC data were not generated for TECARTUS®, but rather FDA required the information be resubmitted in the TECARTUS® biologics license application (BLA). Certain facility inspections were waived due to YESCARTA® and TECARTUS® sharing the same licensed manufacturing site, which could leverage overlaps in the planned cGMP/surveillance inspections. For TECARTUS®, drug product batch analysis, stability and stability stress studies were conducted to confirm analytical methods, as well as container closure integrity testing was performed.

Developing a Risk-Based Approach to Support Data Extrapolation Between Product Versions

Extrapolating data across engineered cellular therapy product versions necessitates a fundamental understanding of the primary product and its functional and biophysical properties (Table 2), which in turn requires sufficient non-clinical, CMC, and clinical data, and adequate scientific justification for extrapolation. A framework for evaluating risk in pharmaceutical development is well established in the International Council for Harmonization (ICH) Q9(R1) and Q8(R2) guidelines on Quality Risk Management and Product Development.^{9,10} Extensive knowledge of critical process parameters, product quality attributes, and well-established, robust analytical methods are essential to justify extrapolation and support development of subsequent product versions. To support this, qualified and fit-for-purpose analytical methods that characterize quality attributes are necessary for a variety of critical parameters (e.g., safety, purity, potency, and identity) to define risk categories.

Table 2: Proposed Best Practices in Process and Product Development to Support Data Extrapolation

1. Generate comprehensive product knowledge

Gather appropriate non-clinical, clinical, and CMC knowledge based on the stage of drug development.

2. Evaluate the relationship between product attributes (process parameters and critical quality attributes [CQA]) and safety and efficacy using non-clinical or clinical data sets

While the initial assessment can be performed based on non-clinical and clinical data, as the product advances through later clinical development stages more robust information on the product efficacy and safety profile will enable a more meaningful determination of how a potential change can impact CQAs or product safety and efficacy. Thus, a stepwise approach will be necessary as multiple products advance through development:

- 1) Assess the relationship between manufacturing process parameters and CQAs (e.g., identity, purity, potency, and safety).
- 2) Assess the impact of each CQA on product safety and efficacy (i.e., clinical activity).

3. Develop parameters to define risk and perform risk assessment to facilitate development of secondary products

Based on the defined relationships between any changes in quality attributes and safety and efficacy profiles between the primary and secondary product, define:

- 1) The relative risk of a change on product safety and efficacy, and
- 2) Appropriate action(s) to be taken based on the assigned risk.

4. Develop data packages based on identified risk and actions to mitigate risk to support regulatory submission of a new product version

Based on the totality of evidence from the primary and secondary products and assigned level of risk of the change(s) on safety and efficacy of the secondary product, determine the appropriate actions. Such actions could include extrapolation of data from the primary product, generation of additional or new data or development of clinical risk mitigation strategies to facilitate clinical development of the secondary product. There should be frequent and early discussions with FDA particularly when there are uncertainties regarding regulatory and clinical pathways (i.e., will the data extrapolation package be acceptable, will safety run in data or additional data be necessary to support the use of the new secondary products, etc.).

Based on the magnitude of difference in assay outputs relative to the original product version and other data governing the modification that may exist, a risk assessment can demonstrate the probability and severity of risk to patients due to a product modification. Of note, especially for autologous products with variable incoming starting material, variability between final products can be expected, especially early in development, making extrapolations potentially more challenging. Furthermore, the sensitivity of the assays utilized for in-process controls and final product release must be considered. Consequently, evaluating the totality of the manufacturing, characterization, and release data as well as clinical data are critical when extrapolating between product versions.

The type and amount of required additional data for extrapolation will vary and depend on whether a change has a minor or major impact on product quality, efficacy, or safety. A modification that results in a low-risk impact may allow for data extrapolation across products with targeted data collection to address data gaps and support regulatory requirements, whereas a modification that results in a high-risk impact may require more extensive studies. For example, a low-risk impact that has a minor impact only on product quality may require an analytical comparability assessment, while a moderate-risk impact that impacts patient safety/efficacy may require a clinical bridging study, and a high-risk impact may require a larger clinical trial to confirm safety and efficacy in accordance with the degree of expected similarities. The patient population and magnitude of unmet need should also be considered in thinking about risk and may lead to a shift in risk tolerance for a particular development program as well.

Classifying the impact of modifications and product changes as low- or high-risk may not be easily determined at the outset of development of the new product. The extent to which prior data can be extrapolated to inform development of a new product version will depend on several factors, including the intended development plan of the new product version and risk determination for the impact of the changes in the new product on safety and efficacy. In a risk evaluation, it will be important to assess the robustness and types of existing data available from the primary product such as information from analytical and in vitro studies, non-clinical in vivo studies, clinical pharmacokinetic/dynamic (PK/PD) studies (i.e., biomarker correlates, product correlates of response), and clinical efficacy and safety studies (Table 3). The analytical methods deployed will vary based on the type of engineered cellular therapy product (e.g., autologous, allogeneic, CAR, TCR, etc.) as well as the types and extent of modifications introduced. Methods to analyze risk should be defined early in development and have an adequate level of sensitivity to identify expected differences between two product versions and support a risk-based extrapolation plan.

Table 3: Select Product Attributes, Analytical Assays, and Studies for Formulating an Extrapolation Strategy for Secondary Versions of Engineered Cellular Therapy Products

Parameter	Assessment Stage	Measure	Readout(s)	Actionable Output
Safety	Non-clinical/ Preclinical	Binder identity	<ul style="list-style-type: none"> High-content proteomic screening Tissue panel screening 	Assess off-target binding potential (e.g., weak potential for off-target binding to non-essential and essential targets) vs. primary product
		In vivo pharmacology and toxicology and histopathology	<ul style="list-style-type: none"> Tolerability In-life parameters (e.g., body weight, physical appearance, behavior, etc.) Tissue biodistribution Deaths 	Assess statistical differences vs. primary product
	CMC	Copies vector/cell	<ul style="list-style-type: none"> Vector copy number 	Assess average vector copy/cell vs. primary product
		Integration site and rearrangement analyses	<ul style="list-style-type: none"> On- and off-target integration sites Genomic rearrangement status 	Identify and quantify frequency of on- and off-target genome editing sites and quantify genomic rearrangement events vs. primary product
		Cytokine production	<ul style="list-style-type: none"> Cytokine profiling (e.g., basal, target dependent) 	Assess statistical fold-change of effector cytokines values vs. primary product
		Proliferation potential	<ul style="list-style-type: none"> Target dependent-rate, doublings Antigen-/cytokine-independent proliferation 	Assess statistical differences in proliferation rate and maximum proliferation vs. primary product
	Clinical	Immunogenicity assessment	<ul style="list-style-type: none"> Anti-product antibody assay Anti-transgene antibody assay 	Assess titers and isotypes of anti-product antibodies vs. primary product
		Clinical measures	<ul style="list-style-type: none"> Frequency and severity of adverse events Clinical laboratory measurements Product expansion kinetics 	Identify statistically significant differences vs. primary product

Potency	Preclinical	In vivo efficacy studies	<ul style="list-style-type: none"> Tumor growth 	Quantify statistical differences in dose required to achieve complete response vs. primary product
	CMC	Functional response	<ul style="list-style-type: none"> Target-specific cytokine production/ cytotoxicity 	Quantify statistical differences in target-dependent cytotoxicity and effector cytokine activity vs. primary product
		Transgene expression	<ul style="list-style-type: none"> % transgene-positive cells Mean fluorescence index of transgene on engineered cells 	Assess statistical differences in engineering efficiency and transgene expression vs. primary product
		Phenotypic/ genotypic assessment	<ul style="list-style-type: none"> Flow cytometry-based T-cell immunophenotyping 	Compare immune activation, memory, exhaustion phenotype and genetic evaluation vs. primary product
	Clinical	In vivo dose/ response evaluation	<ul style="list-style-type: none"> Expansion kinetics and persistence Minimum efficacious dose 	Assess statistical differences in rate of expansion, maximal expansion, 30-day area under the curve (AUC), 30-, 60-, and 90-day persistence vs. primary product
Identity	CMC	Transgene cassette sequence	<ul style="list-style-type: none"> Full sequencing of transgene cassettes and regulatory elements 	Identify any changes in protein sequence vs. primary product

This table provides examples for how product attributes, analytical assays, and studies can help evaluate the impact of a modification on product biology including potential safety and efficacy. Not all measures are relevant for each type of engineered cellular therapy.

Leveraging the Totality of Evidence to Support Product Development at Specific Stages of Clinical Development

As products progress through development, the amount of data available to determine risk and extrapolate across versions increases (e.g., extrapolating data from a primary product in early phase, a primary product in late phase, or an already approved product). Table 4 provides examples of how, when justified, data extrapolation can streamline evidence generation, assist in a more seamless transition from one phase of development to another (i.e., academic to industry, early- to mid-phase, and late-phase to post-market), minimize repetitive data collection, and potentially shorten clinical development timelines. A few example strategies are also noted below.

1) **Early Phase Clinical Development**

Early phase safety and efficacy data from the primary product could support an understanding of the preliminary safety and efficacy profile, the context to establish dosing and schedule, and an approach to data collection in later-phase studies for the secondary product. For example, if appropriately justified, sponsors could propose a similar starting dose for a secondary product as the recommended phase 2 dose for the primary product

and/or use the primary product profile to inform more targeted dose limiting toxicity (DLT) criteria to advance a secondary product through early phase studies more efficiently. In early and late phase trials, prior product knowledge could help prepare for expected toxicities and/or inform approaches to reduce or mitigate symptomatic adverse events.

2) **Late Phase Clinical Development**

In instances where a primary product is in late phase development or is approved, the totality of data from the primary product may allow a secondary version to move straight into a Phase 2/3 clinical trial. Additionally, data extrapolation may be appropriate to justify a reduced clinical dataset for the secondary product based on the similarities with the primary product. For instance, a Phase 3 randomized controlled trial (RCT) readout of the primary product paired with a single-arm clinical bridging study of the secondary product in the same indication to support registration of the secondary product. This could dramatically improve patient access to improved variations of products which have already demonstrated robust safety and efficacy (i.e., via Phase 3 RCT).

3) **Post-Market Phase**

Prior product knowledge and the totality of evidence could aid in identification of potential longer-term treatment effects, inform safety surveillance activities, and support clinical management in clinical practice for a secondary product. Additionally, post-market data from a related product may justify a shorter duration of patient safety follow-up for a secondary product in late-stage development or reduce the 15-year long-term follow-up period in the post-market setting to decrease costs, resources, and patient burden.

Table 4: Opportunities for Data Extrapolation from a Primary Product

Data	Opportunities
CMC	<ul style="list-style-type: none"> • Extrapolate viral vector/gene editing tools/cell engineering product information, and product/process characterization data • Extrapolate drug product presentation information including container and closure systems, fill volumes and cell concentration to support process qualification • Use stability data from primary product to support initial stability for secondary product • Implement reduced stability programs leveraging previous programs and/or matrixing beyond initial stability studies • Include only representative engineering batches in the initial IND of a secondary product and commit to provide certificate of analysis from good manufacturing practice (GMP) batch prior to initiating patient dosing • Reuse gene editing safety data (i.e., translocation information, on and off target editing data) if same edits are used with different CAR • Use a risk-based microbiology control strategy based on experience with the primary product to minimize redundant safety testing requirements • Use same analytical methods including potency assays (qualified or validated as appropriate) • Use orthogonal assays to support similar characteristics of potency with the secondary product • Extrapolate residual control strategy as applicable, and apply to new product • Leverage specifications development of primary product to enable more refined set CQAs

<p>Non-clinical/ Preclinical</p>	<ul style="list-style-type: none"> • Use same relevant animal model and, if not available, justify not conducting toxicity studies • Potential to reduce/waive in vivo studies and use in vitro studies for proof of concept by referencing data generated with the primary product • Use potency data from primary and secondary product to support in vivo study design for secondary product (i.e., dose)
<p>Clinical Safety</p>	<ul style="list-style-type: none"> • Inform starting dose using primary product data • Extrapolate safety data from primary product to optimize, reduce testing (i.e., replication competent lentivirus [RCL]/replication competent retrovirus [RCR]), and timepoints required to assess long-term safety • Extrapolate potency data to determine potential support for or differentiation of the safety profile for the secondary product as a supplement of secondary drug safety data with supportive key safety data (or conclusions) from the primary drug data • Extrapolate safety data from the primary product for the secondary product in a regulatory filing(s)
<p>Clinical Efficacy</p>	<ul style="list-style-type: none"> • Support the starting dose and number of dose levels needed to be tested in early clinical studies, where appropriate • Extrapolate certain clinical data from one indication to support development of other clinical indications with the secondary product using the primary drug product efficacy as supportive or primary evidence to support the secondary drug clinical development and regulatory filings • Pending the nature of the modification and stage of development, the clinical trial may require fewer patients treated with the new product version for clinical comparability • Consider whether shortened follow up time for the patients treated with the new product version may be appropriate • Extrapolate biomarkers/assays for measuring clinical efficacy based on similarity to primary product or to support clinical cutoff for patient selection

Mechanisms for Exploring Data Extrapolation Opportunities and Engaging with FDA

Considerable progress is being made in the development and use of engineered cellular therapies and the field is still evolving. The conceptual framework outlined in this white paper intends to accelerate investigation and development of the next generation of engineered cellular therapy products and may also act as a guide when expanding to other indications and patient populations. As the use of data extrapolation across product versions becomes more commonly explored in development programs for engineered cellular therapies, optimal approaches to analyze, interpret, and present data in a rigorous and standardized manner will be critical. As product and process knowledge increases within individual development programs and within the field, adaptive regulatory processes that adjust based on the potential risks associated with the modification or stage of development should be in place and support data extrapolation in development of iterative product versions. An assessment aid-like tool (see prototype in Appendix Supplemental Table 2) could support a more systematic approach for determining the appropriateness of data extrapolation within clinical development programs of secondary products and serve as a vehicle for transparent information exchange when meeting with the FDA.

during product development and, if applicable, warrant the use of science- and risk-based regulatory approaches allowing streamlining of CMC development activities, so that clinical benefits of earlier patient access to these products can be realized.

- **Designation Program for Platform Technologies:** This is a designation program for platform technologies that have the potential to increase efficiencies in drug development. Applications for drugs or biologics that use or incorporate platform technologies may be eligible for certain expedited development or review actions. The intent of this designation program is to bring significant efficiencies to the drug development or manufacturing process as well as to the review process for products across the platform. Many of the concepts and areas for data extrapolation outlined above may be within scope of cell therapy platforms and thus able to be successfully leveraged in subsequent platform products.

In addition to the meeting types and mechanisms noted above, the Initial Targeted Engagement for Regulatory Advice on CBER/CDER Products (INTERACT) and CBER Advanced Technology Team (CATT) may be appropriate to discuss data extrapolation plans or use of new technology/methods to enable data extrapolation.

Moving Forward

Given the uniqueness of engineered cellular therapies, opportunities for continued dialogue in the post-approval setting with the FDA, including the Office of Therapeutic Products (OTP), will be important to encourage continued innovation. Additional data and evidence generation, as well as learnings from leveraging safety data across different versions of products, should inform risk-based approaches to defining the optimal safety follow-up period as the field of engineered cellular therapies continues to grow and evolve. FDA workshops could help inform updated guidance on, for example, generating long-term follow-up data for engineered cellular therapy products and clarifying opportunities to streamline data or compress development timelines based on known or expected safety events. Additionally, workshops and other mechanisms should be explored to capture and disseminate best practices and case studies of data extrapolation in clinical development as well as learning from pilot projects like CDRP, which will help educate sponsors in exploring adequate development pathways. A question-and-answer resource could provide timely answers to questions that are commonly asked and applicable across development programs. The concepts and proposals in this white paper hold promise in streamlining data requirements, while still adequately and robustly assessing products, and ultimately shortening the timelines for bringing these transformative therapies to patients.

The field continues to progress, and numerous developers are investigating engineered cellular therapies to not only expand into new disease areas and lines of therapy, but also to improve upon available engineered cellular therapies. For innovation to reach patients in a meaningful timeframe, leveraging available data and extrapolation from related product versions is one mechanism to accelerate development. Additional approaches for accelerating investigation and development of the next generation of engineered cellular therapy products should also be explored. Specifically, in addition to extrapolation, trial design considerations, alternative study designs, real-world data sources, novel endpoints, and use of bioinformatic approaches may accelerate investigation and will require thoughtful discussion among key stakeholders, including regulators, investigators, patient advocacy groups and sponsors.

References

1. Studying Multiple Versions of a Cellular or Gene Therapy Product in an Early-Phase Clinical Trial. <https://www.fda.gov/media/152536/download>
2. Considerations for the Development of Chimeric Antigen Receptor (CAR) T Cell Products. <https://www.fda.gov/media/156896/download>
3. Human Gene Therapy Products Incorporating Human Genome Editing. <https://www.fda.gov/media/156894/download>
4. Text – H.R.2617 – 117th Congress (2021–2022): Consolidated Appropriations Act, 2023 | Congress.gov | Library of Congress
5. Interpreting Sameness of Gene Therapy Products Under the Orphan Drug Regulations. <https://www.fda.gov/media/134731/download>
6. YESCARTA Approval History. <https://www.fda.gov/media/109140/download>
7. TECARTUS Approval History. <https://www.fda.gov/media/141164/download>
8. Approved Risk Evaluation and Mitigation Strategies (REMS) for YESCARTA and TECARTUS: <https://www.accessdata.fda.gov/scripts/cder/remis/index.cfm?event=IndvRemsDetails.page&REMS=375>
9. ICH Harmonized Guideline. Quality Risk Management. https://database.ich.org/sites/default/files/ICH_Q9%28R1%29_Guideline_Step4_2022_1219.pdf
10. ICH Harmonized Tripartite Guideline. Pharmaceutical Development Q8(R2). https://database.ich.org/sites/default/files/Q8_R2_Guideline.pdf
11. Drugs@FDA: FDA-Approved Drugs. <https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm>
12. FDA Website Vaccines Licensed for Use in the United States. <https://www.fda.gov/vaccines-blood-biologics/vaccines/vaccines-licensed-use-united-states>

Appendix

Examples of Changes that Result in Different Versions of an Engineered Cellular or Gene Therapy Product

FDA provides several examples of changes that result in different versions of an engineered cellular or gene therapy product!:

- Changing a cellular product from bulk tumor-infiltrating lymphocytes (TILs) to purified CD8+ TILs.
- Changing from dendritic cells (DCs) pulsed with a recombinant tumor antigen to DCs pulsed with immunodominant peptides from the same antigen.
- Altering the differentiation state of a stem cell product to a more mature cell type along the same lineage (e.g., neural progenitor cells vs. neurons).
- Changing the cell source (e.g., allogeneic vs. autologous, or cord blood vs. bone marrow) for a mesenchymal stromal cell product.
- Changing from an embryonic stem cell bank to an induced pluripotent stem cell (iPSC) bank to produce the same cell type (e.g., retinal pigment epithelial cells).
- Replacing the CAR transgene of a CAR T cell product with a new CAR transgene.
- Modifying a CAR T cell product by adding a second transgene that expresses a costimulatory protein.
- Modifying a gene therapy vector to express the same transgene with a different codon usage, promoter, enhancer, microRNA (miRNA) target or other control element.
- Deleting one or more genes from a viral-based or bacterial-based gene therapy vector.
- Modifying the transgene sequence in a gene therapy vector, resulting in a change to the amino acid sequence of the encoded protein.
- Changing a capsid protein of a viral-based gene therapy vector.

Supplemental Table 1. Examples of Data Extrapolation in Drug Development. A review of publicly available FDA summary documents^{1,12} includes examples where data extrapolation has been appropriately used in drug development.

Therapeutic Class	Data Type Extrapolated	Select Examples	Examples from Review Documents
Small molecule drugs	Clinical pharmacology and exposure-response data	COREG®/ COREG CR®	<p>Clinical Pharmacology/Pharmacodynamics: Based on equivalencies in PK and PD data in COREG CR® compared to COREG the conclusion was drawn that the indications for which the immediate-release (IR) formulation had been approved can be inferred and claimed for the controlled-release (CR) formulation.</p>
Peptide products (synthetic)	Non-clinical and clinical for rDNA derived peptides	Liraglutide/ future liraglutide ANDAS	<p>Non-clinical Pharmacology/Toxicology: Safety margins for toxicities calculated using steady state systemic exposure in healthy adults were similar based on plasma liraglutide area under the curve (AUC) supporting the basis for inclusion of boxed warning and REMS on the risk of thyroid C-cell tumors observed in rodents.</p> <p>Clinical Pharmacology/Pharmacodynamics: Exposure results VICTOZA® in the thorough QTc trial were compared with exposures (C_{max}) following SAXENDA® in weight management trials and found to be largely overlapping supporting extrapolation of results from VICTOZA®'s QTc trial to support approval of SAXENDA® for weight management.</p>

Therapeutic Class	Data Type Extrapolated	Select Examples	Examples from Review Documents
Antibody-based biologic agents	Manufacturing/CMC and clinical data	HERCEPTIN®/ HERCEPTIN HYLECTA® RITUXAN®/ RITUXAN HYCELA®	<p>HERCEPTIN®/HERCEPTIN HYLECTA® Clinical Data: Data extrapolation possible due to the same drug substance and only a formulation change and comparable PK profiles of IV trastuzumab across the neoadjuvant-adjuvant/adjuvant treatment settings in patients with early breast cancer and metastatic breast cancer.</p> <p>Manufacturing and CMC data: Due to the same manufacturing processes and drug substances, cross referencing to the BLA was possible.</p> <p>RITUXAN®/RITUXAN HYCELA® Manufacturing/CMC: Manufacturing processes cross referenced in product quality review.</p> <p>Pharmacology/Toxicology: PK Bridging studies used as primary source to support approval/comparable benefit of RITUXAN® and RITUXAN HYCELA®.</p>
Vaccines	Manufacturing/CMC and clinical primary immunogenicity leverages parent profile as a control (PVN 13 vs PVN20)	PREVNAR 13® (PVN13)/PREVNAR 20® (PVN20)	<p>Manufacturing/CMC: P20VN and PVN13 vaccines have nearly identical manufacturing processes for the 13 common serotypes.</p> <p>Clinical/Primary Immunogenicity: Vaccine induced opsonophagocytic activity (OPA) activity was used in the licensure of PVN13 by comparing the OPA titers induced by PVN13 with the licensed 7-valent pneumococcal conjugate vaccine.</p>

Supplemental Table 2. Data Extrapolation Assessment Aid Prototype. This document could be submitted as part of an initial IND and/or subsequent IND amendments for a secondary product or as justification for subsequent amendments to a protocol based on new learnings from another product version to aid in discussion with FDA. Part A and Part B describe supportive information and data to justify and evaluate data extrapolation in the clinical development of secondary products.

Supportive Data	Key Information	Guidance for Providing Information
Part A – Background/Overview		
Overview of the Primary Product	<ul style="list-style-type: none"> What is the stage of development of the primary product in? Summary of product characteristics (e.g., type of engineered cellular therapy, mechanism of action, target, CMC overview) Summary of data related to safety and efficacy data and pharmacologic properties (e.g., safety summary, efficacy summary, dosing, dose/response relationships, any correlations or association between CQAs and clinical data, PK characteristics, clinical studies) 	Articulate key non-clinical, CMC, preclinical and clinical safety, and efficacy data set.
Overview of the Secondary Product	<ul style="list-style-type: none"> What is the stage of development of the secondary product in? Summary of shared characteristics and differences between secondary and primary product Summary of data from secondary product [if applicable] Summary of known information gaps 	Articulate similarities and differences between primary and secondary product with a focus on impact to patient safety and pharmacologic properties.
Summary of Development Plan for Primary and Secondary Product	<ul style="list-style-type: none"> Summary of development strategy for primary and secondary product (i.e., will both products be developed in parallel, or will the secondary product replace the primary product?) Timeline of development strategy 	Describe development strategy for the primary and secondary product. Outline anticipated/expected timelines for data readouts and how this will inform development decisions for the secondary product.
Part B – Extrapolation Strategy		
Data Extrapolation Details	<ul style="list-style-type: none"> What data are being extrapolated? How will the extrapolated data from the primary product be used in the development of the secondary product? 	Information collected in this section could be presented in a tabulated format: <ul style="list-style-type: none"> Data being extrapolated Sponsor assessment of associated risk Mitigation strategy
Justification for Data Extrapolation	<ul style="list-style-type: none"> What is the rationale and justification for data extrapolation (i.e., risk assessment)? 	
Risk Mitigation	<ul style="list-style-type: none"> How will known information gaps and risks be mitigated? 	

Novel Approach to Accelerate Lung Cancer Research: Lung-MAP and the Potential of Public-Private Partnerships

Roy S. Herbst¹, Charles D. Blanke², and Ellen V. Sigal³

ABSTRACT

The National Cancer Institute recently found that death rates for non-small cell lung cancer (NSCLC) have been reduced by over 6% overall in recent years. This reduction in mortality has been accompanied by an average increase in overall survival and largely credited to the therapeutic advancements for the effective treatment of NSCLC. Numerous molecular alterations have been identified in NSCLC that have enabled the development of new drugs capable of targeting these changes and efficiently kill cancerous cells. New treatments to modulate patients' immune systems have been shown to be effective in stimulating natural immune cells to have an improved anti-cancer effect. While these types of approaches to treat cancer are providing new options for patients, leadership from the Food and Drug Administration (FDA) recognized that the

expansion of targeted therapy in NSCLC presented significant promise, but evaluation of the safety and efficacy of these new drugs would be slowed if new models for conducting clinical studies were not identified. Specifically, the FDA recommended that a comprehensive approach be implemented to identify the patients that are the best candidates for these, and other new treatments based upon the molecular characteristics of their tumors, and more efficiently conduct the clinical studies necessary to evaluate the safety and efficacy of new drugs. To address this growing challenge, leading lung cancer experts and stakeholders across academia, government, industry, and patient advocacy came together to design a clinical research approach that could serve as a sustainable infrastructure for new lung cancer treatments called the Lung Cancer Master Protocol.

Introduction

Over 230,000 Americans are diagnosed with lung cancer each year and it remains the leading cause of cancer death (1). However, over the past several years, notable progress has been made. Researchers from the NCI recently found that death rates for non-small cell lung cancer (NSCLC), the subtype of lung cancer that comprises approximately three quarters of all lung cancers, have been reduced by over 6% overall in recent years (2). This reduction in mortality has been accompanied by an average increase in overall survival times and largely credited to the therapeutic advancements for the effective treatment of NSCLC. Numerous molecular alterations have been identified in NSCLC that have enabled the development of new drugs capable of targeting these changes and efficiently kill cancerous cells. In addition, new treatments to modulate patients' immune systems have been shown to be effective in stimulating natural immune cells to have an improved anticancer effect (3, 4).

While these types of approaches to treat cancer are providing new options for patients, leadership from the Food and Drug Administration (FDA) recognized that the expansion of targeted therapy in NSCLC presented significant promise, but evaluation of the safety

and efficacy of these new drugs would be slowed if new models for conducting clinical studies were not identified. Specifically, the FDA recommended that a comprehensive approach be implemented to identify the patients that are the best candidates for these, and other new treatments based upon the molecular characteristics of their tumors, and more efficiently conduct the clinical studies necessary to evaluate the safety and efficacy of new drugs. To address this growing challenge, leading lung cancer experts and stakeholders across academia, government, industry, and patient advocacy came together to design a clinical research approach that could serve as a sustainable infrastructure for new potential lung cancer treatments called the Lung Cancer Master Protocol (Lung-MAP).

Background

Lung-MAP is an innovative, groundbreaking clinical trial model designed to advance the development of targeted therapies in a manner that is more efficient than if individual clinical trials were conducted for each drug candidate independently of one another. The primary objective of Lung-MAP is to evaluate the overall survival of biomarker-selected patients treated with standard of care versus the experimental targeted therapy. This first-of-its-kind trial model provides broad-based molecular screening to each patient and matches them to various substudies testing new therapies based on their unique tumor profiles. Substudies are regularly completed, and new studies added to keep pace with the rapidly evolving molecular understanding of lung cancers. For patients that do not match an existing biomarker-driven substudy, a "non-match" immunotherapy-based substudy is continuously available. This helps ensure that there is an available substudy for any patient that enrolls in Lung-MAP as part of their care, and creates a new framework for more efficient, collaborative trials.

At the origin of Lung-MAP, leaders from the NCI and the FDA sought to develop a consensus on how to establish a disease-specific standing research network capable of conducting large trials with diverse populations. The Lung Master Protocol Trial Design Proposal

¹Yale Cancer Center and Smilow Cancer Hospital, New Haven, Connecticut.

²SWOG Cancer Research Network, Oregon Health & Science University and Knight Cancer Institute, Portland, Oregon. ³Friends of Cancer Research, Washington, District of Columbia.

Corresponding Author: Roy S. Herbst, Yale School of Medicine, Yale Cancer Center, New Haven, CT 06520. E-mail: roy.herbst@yale.edu

Clin Cancer Res 2023;XX:XX-XX

doi: 10.1158/1078-0432.CCR-23-2690

This open access article is distributed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) license.

©2023 The Authors; Published by the American Association for Cancer Research

was introduced at the 2012 Conference on Clinical Cancer Research, hosted by Friends of Cancer Research and the Brookings Institution, and was further developed in conjunction with government and industry partners over a series of workshops, forums, and working groups (5). Lung-MAP launched in 2014 as a public-private partnership between the FDA, NCI, the Foundation for the NIH, leading academic researchers and institutions, patient advocacy groups, and industry. The initial iteration of Lung-MAP was led by the Southwest Oncology Group (SWOG), one of the four adult cooperative groups that comprise the NCI's National Clinical Trial Network.

While SWOG continues to hold the Investigational New Drug Application for the protocol, Lung-MAP and its operational leadership has since been expanded to include all of the adult cooperative groups and currently operates at over 700 hospitals and clinics across the country (6, 7).

Key Attributes of Lung-MAP

Innovative designs

To be successful, clinical trials must be able to evolve with the fast-paced treatment and drug development landscape. Lung-MAP has shown that successful trial design and execution require collaboration between drug companies and clinical researchers, and between different companies as well. Multiple different options for substudies have been designed and implemented over time, which allows for Lung-MAP to be tailored to each new drug and clinical scenario as needed. For some studies, this includes a phase II/III design in which an interim analysis is performed, and successful drugs seamlessly proceed into the latter phase of the study. This introduces efficiencies in terms of trial conduct. In this case, the use of patient data from the earlier phases and alleviation of time gaps in trial initiation is able to reduce the time and total number of patients needed to accrue as compared with the conduct of individual phase II and phase III studies. Pharmaceutical companies have utilized this process and in numerous cases found it to be complementary to their approach as evidenced by the recent Pragmatica lung trial which evolved out of SWOGs 1800a (8).

Culture coordination

Implementing and conducting a traditional clinical protocol is resource intensive and requires significant project coordination. This can be a time-consuming process to fulfill and manage procedures set by individual institutions. A protocol that contains multiple substudies and regular modifications can exacerbate the challenges and amplifies the need for cohesive operations and project management. A mutual commitment among all project participants to create a culture for maximizing efficiencies, even when that requires modifying normal processes, is needed to overcome the complexities of a master protocol. Lung-MAP works with advice from the FDA and NCI, as well as an Oversight Committee, Executive Operations Group, and Project Management Office to facilitate trial design and operation. This transparent and strong governance structure has helped to improve the inclusion of a wide variety of expertise and to overcome barriers that can slow typical clinical studies. Strong, cross-sector leadership is needed to identify and implement best practices to maximize efficiency and facilitate a joint commitment to address patient needs. This multi-stakeholder approach to project leadership has undertaken efforts to reduce overall trial start-up time for new substudies, assisted with the migration to a centralized Institutional Review Board, and often

aims to improve efficiencies for long-term sustainability and enrollment in the trial.

Drug selection

Maintaining a pipeline of new drug candidates is essential for the success of any master protocol. Lung-MAP operates with a Drug Selection Committee comprised of leading government and academic experts charged with identifying potential new drug targets and evaluating the applications of candidates as they are submitted. Lung-MAP has held over 30 formal drug selection committee meetings since 2013 assessing over 40 drugs from more than 20 companies (9–15). The selection process included additional *ad hoc* meetings to discuss pathways and targets, as well as monthly internal drug selection committee meetings. The ideal agents for this trial have been biomarker selected against specific driver targets which have shown activity in other settings and/or have limited activity in lung cancer. Conducting a trial with the molecular targets without the benefit to recruit patients with rare mutations from a large number of sites would be impossible.

Accelerating science

While genomic sequencing is used to determine which substudy patients are enrolled in, it also provides a wealth of information that can be used to identify additional genomic alterations as potential drug targets and be a molecular research tool for correlative studies in the future. These data are collected as part of the Lung-MAP database and activities related to their use are overseen by the Translational Medicine Committee. In addition to the baseline genomic sequencing, liquid biopsy collection has been incorporated into Lung-MAP and enabled studies such as an assessment of the concordance between tissue and plasma-based tests to identify mutations. To date, over 20 studies have been completed through the Lung-MAP partnership. Many of the studies have eliminated drug candidates due to futility. While this is an indicator of the continued challenges to successfully treating NSCLC, rapid identification of unviable treatments and a “fail fast” mentality can help clear the queue and enable efficient progression of future candidates into clinical testing. In addition, by establishing a common biorepository, biospecimens from both positive and negative trials are contributing to valuable future molecular research, including analyses to better inform sensitivity to different compounds and mechanisms of treatment resistance.

Patient access

The most important attribute of Lung-MAP has been the impact on patient access on multiple fronts. At the outset of the project, genomic sequencing was not as readily available as it is today. Sequencing has been performed by Foundation Medicine, Inc. since the beginning of the project, and given the widespread availability of the trial in several hundred research facilities, sequencing, and subsequently the associated targeted therapies, became available to many patients that otherwise may not have access. To date, over 60% of patients that have been enrolled in Lung-MAP have been from community-based centers which has enabled clinical trial access to more patients. Efforts to enable timely access for patients have also been successful. Prescreening procedures are used to help identify potential patients and preregister them to Lung-MAP. Over 4,600 patients have been screened since 2014, nearly half of which were prescreened during their frontline treatment enabling seamless access to a substudy if subsequent treatment was needed.

Lung-MAP has also shown that providing drugs at the point of care in diverse communities can improve trial diversity. An evaluation of the representativeness of patients enrolled found Lung-MAP improved access for patients of older age, from rural areas, and from neighborhoods with higher social needs compared with other NSCLC trials. The study also found Lung-MAP participants to be younger and less racially and ethnically diverse than patients with NSCLC in the United States, showing there is still work to be done, particularly in Latino populations. To further promote diversity and representativeness in its trials, Lung-MAP has formed a diversity, equity, and inclusion (DEI) subcommittee and is supporting community sites conducting a DEI Gap Analysis and engaging with lung cancer advocacy groups (16).

Future innovation

A recent study through the Lung-MAP partnership showed a significant improvement in overall survival in immune therapy refractory non-small cell lung cancer using the combination of pembrolizumab and ramucirumab as compared with standard of care alone (8). This study provides foundational evidence for what could ultimately advance standard-of-care treatment, and as acknowledgment of the transformative potential, the combination has received Breakthrough Therapy Designation by the FDA. A subsequent phase III study is underway to confirm the initial observations from Lung-MAP but given the already existing experience with both drugs in the combination it offers a subsequent opportunity to optimize the research paradigm. The primary investigators and sponsors, in collaboration with the NCI and FDA, have designed a protocol focused on collecting the core evidence needed to confirm the survival benefit, while minimizing extraneous datapoints that would complicate study conduct and be unnecessary in this situation. This follow-up study is part of the FDA's Project Pragmatica, which aims to identify and help implement studies that reduce the burden of data collection, maximize site and patient participation, and enable efficient research through the use of pragmatic trial designs. This first pilot through Project Pragmatica will inform the design and utilization of pragmatic trials as a future clinical research tool.

References

1. American Cancer Society. Key statistics for lung cancer. Available from: <https://www.cancer.org/cancer/lung-cancer/about/key-statistics.html> Accessed 1/6/23.
2. Howlader N, Forjaz G, Mooradian MJ, Meza R, Kong CY, Cronin KA, et al. The effect of advances in lung-cancer treatment on population mortality. *N Engl J Med* 2020;383:640–9.
3. Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. *Nature* 2018;553:446–54.
4. Wang M, Herbst RS, Boshoff C. Toward personalized treatment approaches for non-small-cell lung cancer. *Nat Med* 2021;27:1345–56.
5. Malik S, Pazdur R, Abrams J, Socinski M, Sause W, Harpole D, et al. Consensus report of a joint NCI thoracic malignancies steering committee: FDA workshop on strategies for integrating biomarkers into clinical development of new therapies for lung cancer leading to the inception of "master protocols" in lung cancer. *J Thorac Oncol* 2014;9:1443–8.
6. Herbst RS, Gandara DR, Hirsch FR, Redman MW, LeBlanc M, Mack PC, et al. Lung master protocol (Lung-MAP)—a biomarker-driven protocol for accelerating development of therapies for squamous cell lung cancer: SWOG S1400. *Clin Cancer Res* 2015;21:1514–24.
7. Redman MW, Papadimitrakopoulou VA, Minichiello K, Hirsch FR, Mack PC, Schwartz LH, et al. Biomarker-driven therapies for previously treated squamous non-small-cell lung cancer (Lung-MAP SWOG S1400): a biomarker-driven master protocol. *Lancet Oncol* 2020;21:1589–1601.

Conclusion

Lung-MAP is a unique public-private partnership that has developed a standing infrastructure that new drugs can be rapidly incorporated into for efficient evaluation of their safety and efficacy. It has facilitated genomic sequencing and enabled treatment of more than 1,000 patients on the basis of the molecular profile of their cancer. Master protocols can serve as a more efficient approach, particularly for smaller, molecularly defined patient subsets, than individual trials. Issues regarding patient quality of life and the implications of participation in genomic studies have also been explored (17, 18). However, implementation can be complicated. Oversight, implementation, and project management are more laborious than for a single study, or even the same number of independent studies. Upfront planning and regular communication are critical. A shared goal and creation of a culture toward constant innovation, aggressive timelines, and teamwork are essential. As science continues to evolve, the models for timely research also need to advance to efficiently and successfully meet patients' needs.

Authors' Disclosures

R.S. Herbst reports other support from Southwest Oncology Group during the conduct of the study and from American Association for Cancer Research, International Association for the Study of Lung Cancer, and Society for Immunotherapy of Cancer outside the submitted work; personal fees from Abbvie, Bristol Myers Squibb, Candel Therapeutics, Cybrexa Therapeutics, DynamiCure Biotechnology, LLC, eFFECTOR Therapeutics, EMD Serono, Gilead, HiberCell, I-Mab Biopharma, Immune-Onc Therapeutics, Janssen, Johnson and Johnson, Loxo Oncology, Mirati Therapeutics, Nextcure, Novartis, Ocean Biomedical, Inc., Oncocyte Corp, Oncternal Therapeutics, Pfizer, Regeneron Pharmaceuticals, Revelar Biotherapeutics, Ribbon Therapeutics, Roche, Sanofi, Seattle Genetics, and Xencor, Inc; and personal fees and other support from Immunocore, Junshi Pharmaceuticals, AstraZeneca, Bolt Biotherapeutics, Checkpoint Therapeutics, Eli Lilly and Company, Genentech, Merck and Company, and Normunity. C.D. Blanke reports grants from NIH during the conduct of the study. E.V. Sigal reports personal fees from EQRx, Grail, and AstraZeneca outside the submitted work. No other disclosures were reported.

Received September 8, 2023; revised October 3, 2023; accepted October 26, 2023; published first October 30, 2023.

8. Reckamp KL, Redman MW, Dragnev KH, Minichiello K, Villaruz LC, Faller B, et al. Phase II randomized study of ramucirumab and pembrolizumab versus standard of care in advanced non-small-cell lung cancer previously treated with immunotherapy-lung-MAP S1800A. *J Clin Oncol* 2022;40:2295–306.
9. Leigh NB, Redman MW, Rizvi N, Hirsch FR, Mack PC, Schwartz LH, et al. Phase II study of durvalumab plus tremelimumab as therapy for patients with previously treated anti-PD-1/PD-L1 resistant stage IV squamous cell lung cancer (Lung-MAP substudy S1400F, NCT03373760). *J Immunother Cancer* 2021;9:e002973.
10. Gettinger SN, Redman MW, Bazhenova L, Hirsch FR, Mack PC, Schwartz LH, et al. Nivolumab plus ipilimumab vs nivolumab for previously treated patients with stage IV squamous cell lung cancer: the lung-MAP S1400I Phase 3 randomized clinical trial. *JAMA Oncol* 2021;7:1368–77.
11. Owonikoko TK, Redman MW, Byers LA, Hirsch FR, Mack PC, Schwartz LH, et al. Phase 2 study of talazoparib in patients with homologous recombination repair-deficient squamous cell lung cancer: Lung-MAP substudy S1400G. *Clin Lung Cancer* 2021;22:187–94.
12. Borghaei H, Redman MW, Kelly K, Waqar SN, Robert F, Kiefer GJ, et al. SWOG S1400A (NCT02154490): a phase II study of durvalumab for patients with previously treated stage IV or recurrent squamous cell lung cancer (Lung-MAP Sub-study). *Clin Lung Cancer* 2021;22:178–86.

13. Edelman MJ, Redman MW, Albain KS, McGary EC, Rafique NM, Petro D, et al. SWOG S1400C (NCT02154490)-A Phase II study of palbociclib for previously treated cell cycle gene alteration-positive patients with stage IV squamous cell lung cancer (Lung-MAP substudy). *J Thorac Oncol* 2019;14:1853-9.
14. Aggarwal C, Redman MW, Lara PN Jr, Borghaei H, Hoffman P, Bradley JD, et al. SWOG S1400D (NCT02965378), a Phase II study of the fibroblast growth factor receptor inhibitor AZD4547 in previously treated patients with fibroblast growth factor pathway-activated stage IV squamous cell lung cancer (Lung-MAP Substudy). *J Thorac Oncol* 2019;14:1847-52.
15. Langer CJ, Redman MW, Wade JL 3rd, Aggarwal C, Bradley JD, Crawford J, et al. SWOG S1400B (NCT02785913), a Phase II study of GDC-0032 (Taselisib) for previously treated PI3K-positive patients with stage IV squamous cell lung cancer (Lung-MAP sub-study). *J Thorac Oncol* 2019;14:1839-46.
16. Vaidya R, Unger JM, Qian L, Minichiello K, Herbst RS, Gandara DR, et al. Representativeness of patients enrolled in the lung cancer master protocol (Lung-MAP). *JCO Precis Oncol* 2023;7:e2300218.
17. Unger JM, Qian L, Redman MW, Tavernier SS, Minasian L, Sigal EV, et al. Quality-of-life outcomes and risk prediction for patients randomized to nivolumab plus ipilimumab vs nivolumab on LungMAP-S1400I. *J Natl Cancer Inst* 2023;115:437-46.
18. Roth JA, Trivedi MS, Gray SW, Patrick DL, Delaney DM, Watabayashi K, et al. Patient knowledge and expectations about return of genomic results in a biomarker-driven master protocol trial (SWOG S1400GEN). *JCO Oncol Pract* 2021;17:e1821-9.

Incorporating Pragmatic Elements in Study Designs to Enhance Oncology Randomized Clinical Trials

Introduction

There has been a trend towards increased complexity in cancer clinical trials due to various factors resulting in burden to patients, research staff, and sponsors alike. While novel investigational therapies will require more frequent safety assessments and often a host of primary and secondary efficacy endpoints to characterize risks and benefits, other study contexts where more is known about the therapies under investigation may not necessitate this assessment intensity. Reducing the complexity of trials, where appropriate, may lead to reduced burden on patients, improved enrollment, reduced attrition, and expansion of the number of sites (e.g., site selection) that may be used to generate data on broader patient populations.

Efforts to streamline data collection and simplify clinical trial designs through introduction of pragmatic clinical trial (PCT) elements, where appropriate, are underway. Pragmatic elements range from recruitment, to broadening eligibility criteria and selection of routine clinical practice sites,¹ to flexibility in delivery and monitoring of therapy, to streamlined design, endpoints and data collection including follow-up. The Pragmatic-Explanatory Continuum Indicator Summary (PRECIS) ²² tool is one example of the types of pragmatic elements that can be considered to reduce complexity and make trials more reflective of routine clinical practice (See Appendix 1 for examples). Appropriate integration of pragmatic elements into clinical trial designs will vary depending on the clinical context of the trial and how the results will be used (e.g., inform clinical practice, regulatory intent), and should be incorporated in a manner that ensures study data integrity and patient safety.

Incorporating pragmatic elements can reduce the burden of trial participation. Reduced patient burden holds promise to facilitate enrollment of potentially more diverse trial populations, enable quicker enrollment, and reduce attrition. The lower burden of participation can benefit patients and potential trial sites. Such sites may be in community settings interested in performing research integrated within clinical care and sharing valuable clinical insights, especially outside of academic medical centers in areas that may be enriched for patient populations typically underrepresented in clinical trials. Further, broadening eligibility criteria provides the opportunity to assess efficacy and safety of therapeutics in additional patient populations not usually included

Thank You to Our Contributors

Ashita Batavia, Johnson & Johnson Innovative Medicine

Amy Cramer, Johnson & Johnson Innovative Medicine

Carsten Goessel, GSK Oncology

Nafsika Kronidou Horst, F. Hoffmann-La Roche

Erin Larkins, U.S. Food and Drug Administration

Boris Kin Lin, Eli Lilly and Company

Sumithra Mandrekar, Mayo Clinic/ The Alliance for Clinical Trials in Oncology

Kristin McJunkins, Patient Advocate

Amy McKee, Parexel International

Brittany Avin McKelvey, Friends of Cancer Research

Margaret Mooney, National Cancer Institute

Monika Patre, F. Hoffmann-La Roche

Sheila Prindiville, National Cancer Institute

Donna Rivera, U.S. Food and Drug Administration

Kelly Shanahan, Patient Advocate

This document reflects discussions that occurred among stakeholder groups on various topics. This document should not be construed to represent FDA's views or policies.

in clinical trials, such as those with significant organ dysfunction or reduced performance status. This ultimately enables an improved understanding of a treatment’s effectiveness and safety in a population more representative of the heterogeneous populations that are affected by the disease.

Within the continuum of trial designs, trials can include various pragmatic elements and study objectives. The prospective nature of pragmatic trial designs is critical to address challenges typically seen in observational studies using real-world data which may include data quality, missingness, and heterogeneity of endpoints and outcomes when incorporating data collection more reflective of real-world practices and settings. While trials may be designed with pragmatic elements in various prospective settings, this white paper will focus on randomized interventional PCTs (Figure 1). Randomized PCTs can be categorized as “a type of clinical trial designed to compare an intervention and a comparator in participants who are more similar to those affected by the condition(s) under study in routine clinical practice settings.”³ While not the focus herein, non-randomized pragmatic studies may also be valuable for signal seeking in novel indications, such as the Targeted Agent and Profiling Utilization Registry (TAPUR) Study.⁴

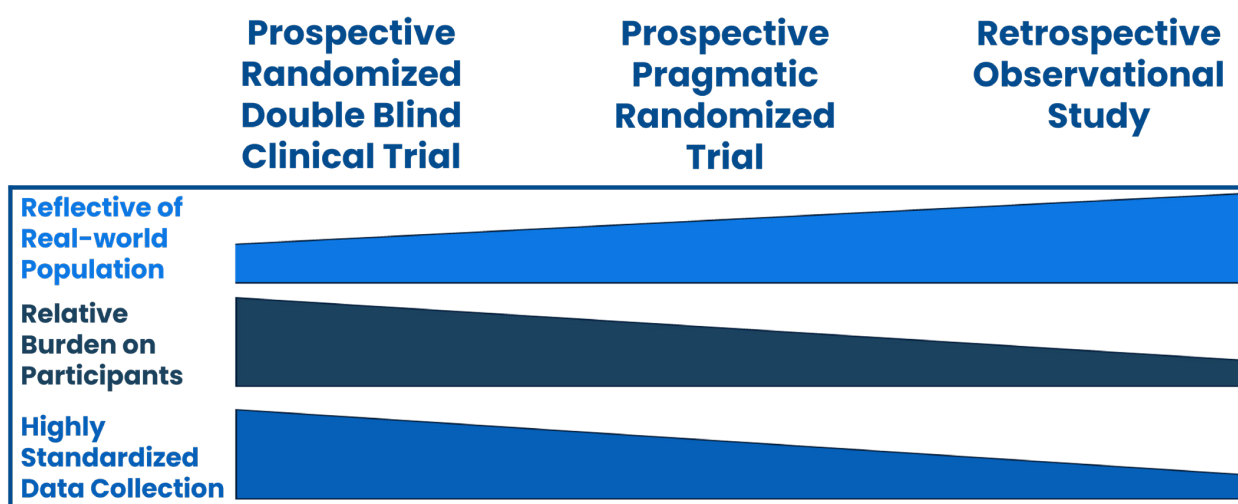


Figure 1. Spectrum of clinical trial designs and characteristics. *Prospectively designed randomized trials with pragmatic elements may include a broader patient population than in traditional clinical trials, with less overall burden and simplified data collection. However, these trials often require more structure and participant burden than traditional observational studies. Adapted from Bevan A, et al. Pragmatic randomized trials considerations for design and implementation, 2019 white paper.*

Outside of oncology, the pragmatic United Kingdom RECOVERY trial,⁵ which randomized treatments for patients hospitalized with COVID-19, allowed for minimal patient eligibility criteria, and streamlined follow-up monitoring through a single online follow-up form recording when each patient was discharged, died, or at 28 days after randomization, whichever occurred first. To date, the trial has provided evidence supporting four treatments for severe COVID-19. These findings highlight the benefits of incorporating pragmatic elements into clinical trial designs to

reach a broader patient population¹, which provides valuable translational lessons for oncology. The FDA Oncology Center of Excellence (OCE) is identifying opportunities to incorporate pragmatic elements into oncology randomized clinical trials as evidenced by the OCE's Project Pragmatica³. Incorporating pragmatic elements into clinical trials may not be appropriate for every drug, stage of development, disease setting, and clinical question. Friends of Cancer Research (*Friends*) convened a multi-stakeholder group of experts including members from the FDA and National Cancer Institute (NCI), drug developers (sponsors), patient advocates, and academic clinicians representing the NCI National Clinical Trial Network (NCTN) to lay out considerations for determining the appropriateness of incorporating pragmatic elements into randomized clinical trials and to outline potential innovative trial designs that can support a shift to streamlining the data collection plan for studies.

Opportunities to Leverage Clinical Trials with Pragmatic Elements

Randomized clinical trials with pragmatic elements could generate evidence to inform clinical practice and reimbursement (e.g., inform NCCN guidelines or payor decisions) as well as regulatory decision-making. Pragmatic trials may be conducted by a variety of entities. For example, pharmaceutical companies may be more likely to conduct trials with regulatory intent, while cooperative groups or academic centers may be more likely to conduct trials to generate evidence to support clinical practice. While trials may initially be designed as research focused only, evidence may ultimately support regulatory decision-making. Therefore, data should be collected in a manner amenable to regulatory submission where appropriate. For trials with regulatory intent, drug developers should meet with the FDA early to share the trial design and understand requirements for data collection, including methodological and evidentiary standards.

In certain circumstances, studies with pragmatic elements may be used to support a regulatory submission. Some examples include fulfilling a post-marketing commitment (e.g., additional safety information), supporting label updates to address evidence gaps, modifying treatment regimens (e.g., adding information on subpopulations not studied in the pivotal study, such as older patients or patients with worse performance status), or supporting a supplemental approval or expanded indication. As efficacy and safety evidence accumulate through the lifecycle of a drug, this expanded knowledge base may allow for the introduction of pragmatic elements to encourage continued evidence generation in an efficient manner through reduced data collection and expanded sources of data (e.g., EHR, registries, Digital Health Technologies) (Figure 2). Conversely, it is unlikely that a highly pragmatic trial design would support the registration of a new molecular entity, given the lack of previous safety and efficacy data.

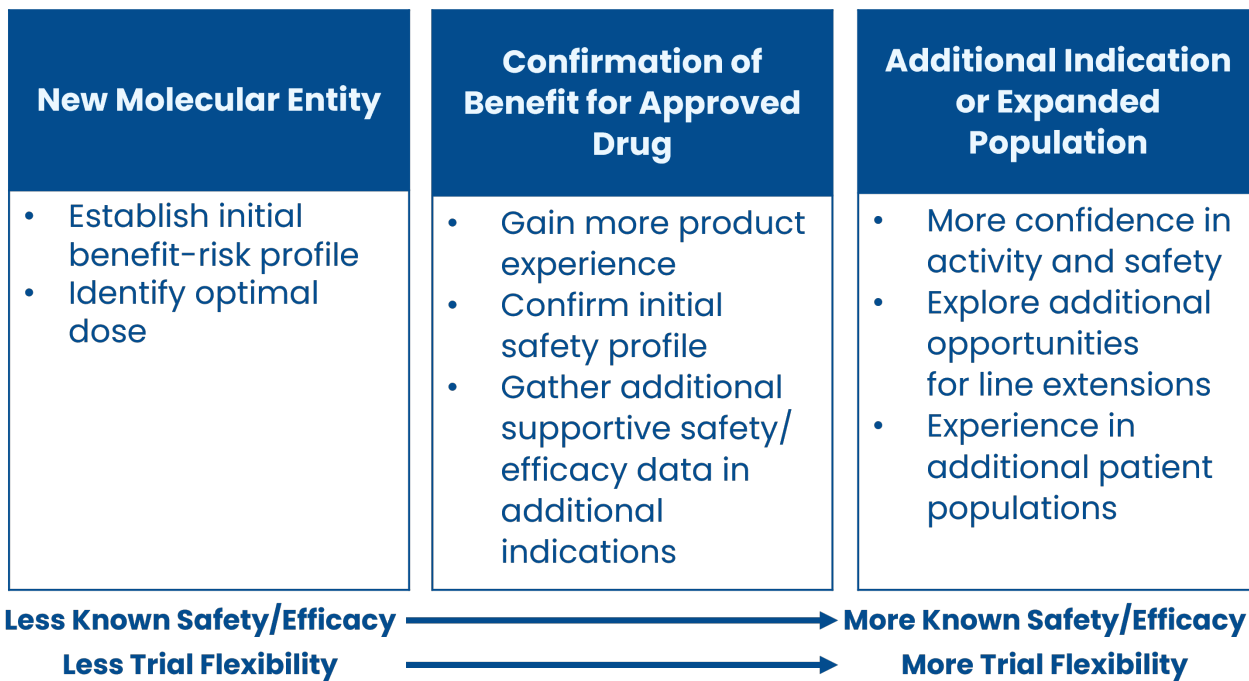


Figure 2. Key objectives across stages of evidence development

Considerations for Including Pragmatic Elements in Clinical Trial Designs

Including pragmatic elements may not be appropriate for every scenario. To aid in identifying characteristics of drug development scenarios that may be amenable to incorporating pragmatic elements, two ongoing oncology trials were assessed, Pragmatica Lung⁶ and the Radiotherapy Comparative Effectiveness (RadComp) trial⁷ (Table 1).

Table 1. Examples of pragmatic study designs and characteristics of scenarios amenable to pragmatic design			
Trial Characteristics	Pragmatica Lung	Rad Comp	Characteristics Amenable to Pragmatic Designs
Purpose of Evidence Generation	Regulatory intent to support a supplemental approval for a new indication	Inform clinical practice and guidelines	Evidence generation from trials with pragmatic elements may inform both clinical and regulatory decision-making
Study Population	Patients with stage IV non-small cell lung cancer (NSCLC)	Patients with locally advanced breast cancer	Disease biology well understood with well understood treatments available

Trial Design	Prospectively designed randomized Phase 3 trial with registrational intent to evaluate overall survival	Prospectively designed randomized trial to evaluate major cardiovascular events	Prospective design, randomized trials, objective endpoint that is meaningful to patients, clearly defined and able to be ascertained in the clinical setting
Previous Supporting Data- Safety and Efficacy	Phase 2 randomized controlled trial reported positive efficacy results for the combination therapy with novel safety concerns not expected, individual agents have known safety profiles	Both therapies are considered standard of care with known efficacy for the intent to treat population with known safety profiles	Previous clinical trials/ SOC clinical practice in disease setting support efficacy and safety
Intervention	Combination of two previously FDA approved agents (ramucirumab and pembrolizumab) in NSCLC, albeit not FDA approved in combination or for the specific treatment setting under investigation	Standard of care proton therapy	Agents are FDA approved in relevant cancer type
Control	Standard of care chemotherapy, physician and patient choice	Standard of care photon therapy	Standard of care treatment available for control
Pragmatic Study Design Elements	<ul style="list-style-type: none"> • Overall survival efficacy endpoint • Minimal adverse event (AE) reporting- only severe (Grade 3 or higher) AEs <ul style="list-style-type: none"> • Broader Eligibility: Enrollment of patients with lower performance status 	<ul style="list-style-type: none"> • Patient-centric health-related quality of life (HRQOL) measurements • Eligibility is minimally restricted (not excluding pre-existing comorbidities) • Treatment is flexible in dosing and techniques • Treatment decisions are at the discretion of the local providers and patients 	Validated clinically meaningful endpoints that are not overly burdensome for data collection (e.g., focused, minimal, and amenable to real-world data collection) and patient centric

A few key characteristics emerged from the two trials. As seen in **Table 1**, the therapies under investigation were FDA approved agents. For Pragmatica Lung however, the drug approvals were for a different patient population/indication than the one investigated, but the novel combination of agents had been studied in a Phase 2 trial in the specific patient population. In each case, there were previous data supporting the safety and efficacy profile of the interventions, such that it was reasonable for data collection to be streamlined. Additionally, the endpoints used in the trials are clinically meaningful, important to patients, validated for the disease setting, and amenable to capture in a routine clinical practice setting. Such trials may need to be larger to accommodate for the potential heterogeneity that can occur in a more flexible trial design. A large effect size can support the use of pragmatic elements as it increases confidence that there would be sufficient statistical power to be able to delineate outcomes, even if there is more heterogeneity in the data due to pragmatic elements, such as a broader patient population and flexibility in design based on routine clinical practice.

Considerations for Incorporating Pragmatic Elements into Study Designs

Trials incorporating pragmatic elements may have a more streamlined design, endpoints, and/or targeted data collection. In all cases, the acceptability of pragmatic elements should be justified by the clinical and regulatory context. The specific scientific question, intent (e.g., inform regulatory decision or treatment guidelines), indication, and drug(s), as well as the totality of evidence previously generated from clinical trials and observational studies, will dictate the elements that may be simplified or streamlined. PCTs may include specific pragmatic elements, though incorporation of all elements may not be feasible. For example, a trial may broaden eligibility criteria and streamline safety evaluation, while maintaining the rigor of primary efficacy endpoints such as radiographic progression endpoints. These elements should be prospectively defined, and patient consultation can add value to the design and planning of the trial. A standardized data collection template for use across all clinical sites to support streamlined data collection and for ease of analysis should be used. Highlighted below are a few pragmatic dimensions to consider for incorporation into a pragmatic trial.

Eligibility Criteria

One pragmatic element that should be considered across most cancer clinical trial contexts is eligibility criteria. Eligibility criteria may be broadened to enable the enrollment of a patient population that is more reflective of the real-world population affected by the disease. There is a sustained effort¹ to encourage broadening eligibility criteria in all oncology clinical trials and pragmatic designs¹ offer the opportunity to study patient populations that may have been excluded from prior trials. Prior evidence will be important in determining the appropriate degree of pragmatism. Broadening the patient population can be nuanced and only specific criteria may be broadened instead of multiple criteria. For example, the performance status may be broadened, but patients with chronic kidney disease may still be excluded if the drug is renally cleared. The totality of available clinical data, including historical trial data, should support the rationale for broadening specific eligibility criteria. Another important consideration is the safety profile of the investigational therapy; there should be enough evidence that there is no safety concern overall in the additional patient population (i.e., known toxicities associated with the therapy are not expected to worsen or be exacerbated by pre-existing conditions included in the broader patient population). If there are concerns with the safety of the agent in the broader patient population that is planned to be included in the pragmatic trial, then additional safety data should be collected and approaches to ameliorate adverse events should be prospectively identified.

Some examples of eligibility criteria that may be relaxed include:

- **Performance Status:** Enroll patients with varying performance statuses, such as patients with an Eastern Cooperative Oncology Group (ECOG) score of 2 in addition to 0-1 scores. Evidence generated from this expanded patient population may inform clinical practice.
- **Organ Dysfunction:** Include patients with pre-specified organ dysfunction, particularly if there is not significant concern from prior clinical data, and the drug's mechanism of action and side effects are known and pose minimal risk. Evidence generated from this expanded patient population may support labeling changes to modification of treatment regimens or optimization of dosing for specific patient subpopulations or inform clinical practice guidelines. Additional safety and clinical pharmacology data may be necessary to support label modifications.
- **Comorbidities:** Include patients with comorbidities such as those diagnosed with HIV, Hepatitis B and/or C, or those that may be immunocompromised if there is no concern for additional patient risk or side effects. Evidence generated from this patient population may inform clinical practice or labeling changes.

Some examples of eligibility criteria that may be specified to ensure adequate representation⁸ include:

- **Age:** Enroll older patients than may have been underrepresented in the pivotal trial but are known to be impacted by the disease. Evidence generated from this expanded patient population may inform clinical practice.
- **Race and Ethnicity:** Enroll patients who may have been underrepresented in the pivotal trial (e.g., non-white and/or Hispanic patients). Evidence generated from this expanded patient population may inform clinical practice or may satisfy a post-marketing commitment or requirement.
- **Gender:** Enroll patients who may have been underrepresented in the pivotal trial (e.g., females) but are known to be impacted by the disease. Evidence generated from this expanded patient population may inform clinical practice.

Efficacy Outcomes

Efficacy data collection may be simplified to reduce patient and site burden by decreasing the number of patient visits/assessments while still providing meaningful information to inform patient treatment. Efficacy endpoints suitable for a pragmatic approach should be clinically meaningful, patient-centric (i.e., meaningful to patients), and amenable to measurement in routine clinical practice, such as overall survival (OS).

The choice of endpoint will depend on the clinical context and trial intent (i.e., how the trial results will be used). When considering efficacy endpoints, it is important to determine if the endpoint measurement would be influenced if the trial design is not double-blinded (both patients and/or investigators are blinded to the treatment the patient receives on the trial). For example, Pragmatica Lung allows investigator's choice of standard of care therapy as the control agent. While objective endpoints such as OS would not be affected by unblinding, endpoints such as disease progression and time to treatment discontinuation (TTD) may be impacted by a patient's or investigator's knowledge of being assigned to control or investigational therapy.

Some examples of specific efficacy endpoints that may be amenable to incorporate in pragmatic trials include:

- **Overall Survival (OS):** OS is a validated clinically meaningful endpoint that is not overly

burdensome for data collection, is patient centric, not subject to bias, and encompasses an understanding of both safety and efficacy. While the trial protocol may only specify collection of survival status, disease assessment will likely also occur based on standard of care. Trials may require the collection of additional efficacy endpoints depending on the disease setting and indication and the intent to support regulatory submission, especially since OS is influenced by subsequent lines of treatment. Further, collecting the cause of death (e.g., disease-related or not) may provide additional context.

- **Response Endpoints:** Response endpoints, such as objective response rate and progression-free survival (PFS), that require strict adherence to assessment criteria (e.g., Response Evaluation Criteria in Solid Tumors- RECIST and the International Myeloma Working Group response criteria for multiple myeloma), central review and evaluation, and a strict schedule of assessment may not be amenable to a simplified approach. Without this strict adherence, heterogeneity and bias in evaluation may be introduced due to variability in the timing of scans, non-biased objective review of scans, or lack of adherence to the strict assessment criteria. Endpoints that rely on tumor assessments may lead to surveillance bias, and consideration should be given to the schedule of data collection to reduce biases. While the criteria for assessment may be more rigid and reflective of a traditional clinical trial, there may be opportunity to relax the schedule of assessments. For example, less frequent assessments with a wider window (e.g., an assessment every 12 weeks +/- 7 days versus a traditional 4 weeks +/- 3 days) may allow a more pragmatic approach to response assessment.
- **Time to Treatment Discontinuation (TTD) and Time to Next Treatment (TTNT):** The inclusion of endpoints that may be captured more easily in clinical settings, such as TTD and TTNT may be considered. However, these endpoints are not routinely used in clinical trials, and therefore may be challenging to standardize and establish thresholds for success/failure. Further, there is difficulty discerning the cause for treatment discontinuation, which may be due to AEs or tolerability, a lack of efficacy, or may be due to a therapy shortage, insurance lapse, or other interruption due to circumstances unrelated to the disease. Past studies have shown patient-level association between TTD and PFS in clinical trials of NSCLC patients across therapeutic classes, and further work is needed to strengthen the evidence of association, including the association with OS. These endpoints are subject to bias of the investigator and patient's clinical circumstance. Thus, the need for randomization of the trial minimizes potential biases. While such endpoints may be appropriate for trials intended to inform clinical practice, at this time they would not be appropriate for trials intended to support regulatory decision-making.

Safety Evaluation

Safety data collection may also be streamlined to reduce patient and trial site burden.⁹ Data collection should focus on signals that may cause physicians to modify or discontinue treatment or pose significant concerns.¹⁰ Fewer patient assessments may be used, such as only evaluating a patient's vital signs and completing study AE forms once per cycle, to streamline safety collection. In addition, attribution has been shown to have minimal value and thus collection of attribution should be minimized or eliminated.^{11,12}

Some examples of specific safety data collection that may be amenable to incorporate in pragmatic trials include:

- **Grade 3 or Higher AEs:** If there is a well-established safety profile and expectation that an expanded population would tolerate the treatment in a similar fashion, the trial may only need to report AEs that are serious and unexpected. Currently, most NCTN Phase 3 trials do not collect these Grade 1 and 2 AEs.
- **Targeted Safety Event Collection:** If a trial incorporates a reduced safety data collection method, the mechanism of action of the drug and prior clinical data will be critical to determine if additional targeted safety data is needed. For example, in study of a novel combination, if there is overlapping toxicity or concerns for specific safety events with the combination, additional data may be needed. Further, if there is a concern for a specific adverse event in a specific patient population included in the pragmatic trial due to previous data, additional data collection for the specific AE may be warranted. This additional data collection may be imperative to support regulatory decision-making.
- **Patient Reported Outcomes (PROs):** PROs could be considered to capture the safety and tolerability events relevant to patients, for more patient-centric data. The Patient Reported Outcome Common Terminology Criteria for Adverse Events (PRO-CTCAE) item library evaluates the symptom attributes of frequency, severity, interference, amount, and presence/absence for patients. Additionally, digital health technologies (DHTs) may be used to collect long-term longitudinal data on patients' symptoms. With all patient assessments of symptoms, consideration should be given to the items and frequency of data collection to reduce patient burden, and patient advocates should be included in the decision-making process for PRO inclusion. PRO data must be well designed, adequately collected, and carefully measured such that data integrity is maintained. Additionally, the intent of PRO inclusion for the overall trial objective is important. A primary endpoint using PROs may be used to inform clinical practice, however incorporation of PROs into a trial intended for regulatory decision-making with other primary endpoints may add additional data collection burden and not support a streamlined approach.

Operational Aspects of Implementing Trial Designs

While this white paper does not go into depth regarding operational aspects to consider when designing trials incorporating pragmatic elements, including simplified informed consent, considerations related to site selection and data sources, these elements are critical to successful implementation of the trial design. Work by the Clinical Trials Transformation Initiative¹³ on embedding clinical trial elements into clinical practice highlights operational aspects to consider, as well as the white paper on point of care trials by Duke Margolis Center for Health Policy.¹⁴ Careful consideration is needed to determine the appropriate research infrastructure and clinical setting in which to conduct these trials; this will have a large impact on data collection and quality, patient population, and overall evidence generation.

Innovative Study Designs to Incorporate Pragmatic Elements

There are numerous approaches to incorporating pragmatic elements into clinical trial design, depending on the specific scenario. To encourage consideration for innovative study designs, a few case studies highlight pragmatic trial design considerations amenable to each scenario. These considerations may inform the inclusion of pragmatic elements into a development program. However, each development program is unique, and the trial design, data collection, evidentiary needs will be different for each scenario. Sponsors should meet early and often with FDA to discuss possible trial designs for their specific indication and therapy.

Case Study 1: Evaluating two well-characterized, FDA approved drugs in a novel combination

Pragmatica Lung is a pragmatic clinical trial including multiple pragmatic elements and is an example of targeted data collection that was acceptable for regulatory decision-making given the prior data available.

Trial Design: Randomized trial comparing a novel combination therapy to control arm of physician's choice of standard of care (following NCCN guidelines).

Trial Design Aspects	Prior Data Available to Support Pragmatic Elements	Pragmatic Element(s)	Operationalizing Pragmatic Elements in Trial Design	Considerations for Including Pragmatic Elements
Patient Eligibility	<ul style="list-style-type: none"> Phase II randomized study of combination in patients with ECOG 0-1 Two standard of care (SOC) agents with known safety/efficacy profile 	<ul style="list-style-type: none"> Expanded eligibility 	<ul style="list-style-type: none"> Lower performance status (ECOG 0-2) <ul style="list-style-type: none"> Stratification factor (ECOG 0-1 vs. 2) All patients with the ability to safely receive the regimens, per FDA label and investigator's discretion (e.g., includes reduced organ function, etc.) 	<ul style="list-style-type: none"> The totality of evidence in higher performance status patients and early FDA input led to acceptable probability of technical and regulatory success
Efficacy Evaluation	<ul style="list-style-type: none"> Phase II randomized study of combination versus SOC with a signal for improved OS 	<ul style="list-style-type: none"> Reduced efficacy data collection Patient-centric endpoint 	<ul style="list-style-type: none"> Overall survival as primary endpoint No protocol required disease assessment (e.g., CT, imaging) No protocol required lab tests, specimen collection Collect primary cause of death, but not contributor causes or source of information 	<ul style="list-style-type: none"> The disease setting/indication (e.g., disease stage, existing therapies, etc.) may require the need to collect additional efficacy endpoints
Safety Evaluation	<ul style="list-style-type: none"> Well-known safety profile of individual agents (both FDA approved) Safety profile in combination (Phase II randomized study of combination) showed no new events 	<ul style="list-style-type: none"> Reduced safety data collection 	<ul style="list-style-type: none"> Serious Grade 3 or higher AEs (Grade 5 or unexpected Grade 3/4 treatment related AE) Fewer patient assessments <ul style="list-style-type: none"> Only vital status and AE form (once per cycle) 	<ul style="list-style-type: none"> If there has not been extensive study of the combination (e.g., not yet studied or in a small number of patients that may not be representative of the broader patient population), additional safety data will be needed If there is overlapping toxicity, or concerns for specific safety events with the combination, additional data may be needed

Case Study 2: Evaluating an FDA approved drug to optimize dosing in a specific patient population

The ASCO PCORI grant¹⁵ is studying dosing strategies of oral CDK4/6 inhibitors in older patients with metastatic breast cancer. This trial aims to collect more evidence on optimal dose for a patient population not well represented in registrational trials. The study design may be best suited to generate evidence to support changing clinical practice/guidelines to inform practitioners of dose modifications in a specific patient population. If there is regulatory intent (e.g., label modification for specific patient population, or to satisfy a post-marketing requirement for dose optimization), additional data will need to be collected.

Trial Design: Randomized trial comparing FDA approved dosing in the patient population to a titrated dosing approach using the same dose schedule but starting at a lower dose and escalating if tolerated.

Trial Design Aspects	Prior Data Available to Support Pragmatic Elements	Pragmatic Element(s)	Operationalizing Pragmatic Elements in Trial Design	Considerations for Including Pragmatic Elements
Patient Eligibility	<ul style="list-style-type: none"> FDA approvals for patients in the study indication 	<ul style="list-style-type: none"> Expanded eligibility (focused on specific patient) 	<ul style="list-style-type: none"> Enroll a specific population not included, or minimally included, in registrational trial (e.g., older adults >65) 	<ul style="list-style-type: none"> Initial evidence in the specific patient population to drive exploration of alternate dosing
Efficacy Evaluation	<ul style="list-style-type: none"> FDA approvals in study indication proving efficacy 	<ul style="list-style-type: none"> Patient-centric endpoint 	<ul style="list-style-type: none"> TTD as primary endpoint EFS, PFS, OS as secondary endpoints 	<ul style="list-style-type: none"> FDA does not commonly use TTD as a primary endpoint for regulatory decision-making, and would likely need additional data (e.g., response and durability of response, PFS) to support a label modification It may be valuable to collect the reason for treatment discontinuation

Case Study 3: Streamlined safety data collection for a pivotal trial investigating a new indication for a previously approved drug

This case study is theoretical and provides considerations for how one may incorporate pragmatic trial elements as part of the pivotal trial in the clinical development program for a targeted agent not yet approved in a new indication (e.g., new cancer type). In this case study there is strong early scientific evidence (e.g., strong scientific rationale for the mechanism of action and prior Phase I/II data that showed a large effect size with the safety profile expected from the approved indication) to support investigation in the new indication. The Phase III trial might be conducted with reduced safety data collection based on the supportive evidence of the earlier phase trial(s). This trial design could provide evidence to support regulatory decision-making by collecting the appropriate efficacy data while streamlining safety data. This reduction in safety data collection could ease burden enabling additional trial sites to participate and to reach additional patient populations.

Trial Design: The pivotal registrational clinical trial is conducted for an agent in a novel indication. The pivotal trial streamlines safety data collection while maintaining efficacy data collection reflective of a traditional explanatory trial.

Trial Design Aspects	Prior Data Available to Support Pragmatic Elements	Pragmatic Element(s)	Operationalizing Pragmatic Elements in Trial Design	Considerations for Including Pragmatic Elements
Safety Evaluation	<ul style="list-style-type: none"> • Registrational trial data, albeit limited in the specific patient population, supports the safety of the therapy; well known safety profile 	<ul style="list-style-type: none"> • Patient-centric endpoints • Reduced safety collection 	<ul style="list-style-type: none"> • Tolerability (Grade 3-4 AEs) • PRO-CTCAEs • Quality of life (PROMIS-29) and FACT-G single item GP5 • Healthcare utilization 	<ul style="list-style-type: none"> • To support a label modification, additional safety and PK data collection will likely be required • Consideration for the frequency of patient assessment for PROs and surveys to limit patient burden

Conclusions and Future Directions

Clinical trials with pragmatic elements have the potential to bridge clinical research and clinical practice by reducing the burden of trial participation. Potential advantages to a more pragmatic clinical trial include enrollment of a more diverse trial population, more rapid enrollment, and reduced attrition. The clinical and regulatory context will determine which scenarios are more appropriate for incorporating pragmatic elements. Approved drugs with established safety and efficacy data are amenable to a more highly pragmatic approach, but all trial contexts can benefit from evaluating how or if increased pragmatism is possible. Thoughtful consideration should be taken regarding whether including pragmatic elements is feasible early in the trial design process. Engagement with FDA will be crucial to determine the data collection, study design, and statistical analysis strategy, should those trials be intended to serve a regulatory purpose.

While the idea of pragmatic clinical trials has existed for decades,¹⁶ there are not many examples used in regulatory decision-making, particularly in oncology. Additional work is needed to encourage and enable the uptake of trials incorporating pragmatic elements with robust evidence generation. Beyond the study design elements discussed in this paper, additional considerations to enable the conduct of pragmatic trials include elements related to data sources and data quality and building local infrastructure at the point of care. Even with conduct in the routine practice setting, there are standards for acceptable data quality to generate evidence. All data may not exist in the electronic health record (EHR) in a structured or standardized way across sites, and data missingness is also of concern; prospectively defined data standards and templates may be needed. Sites that may not routinely conduct clinical trials who have interest in participating in these trials may be inexperienced or lack support staff or the infrastructure necessary to capture needed data to accurately assess endpoints. Therefore, initially there is likely to be some burden on these trial sites while they build their infrastructure and not all sites may be feasible for a trial. Efforts to increase the standardization and level of structured data in the EHR, such as mCODE,¹⁷ may eventually support data collection. Alignment between clinical care and clinical research on data collection standards is needed. In addition, resources and best practices are needed for engaging sites that are not large academic centers and may not regularly conduct clinical trials.

As the field gains more experience identifying ideal scenarios for incorporating pragmatic elements and conducting these trials, it will be important to evaluate whether the predicted benefits are realized and to develop best practices to encourage future use of trials with pragmatic elements to generate robust evidence to support regulatory decision-making.

References

1. Kim ES, Bruinooge SS, Roberts S, et al. Broadening eligibility criteria to make clinical trials more representative: American society of clinical oncology and friends of cancer research joint research statement. *Journal of Clinical Oncology*. 2017;35(33):3737–3744. [doi:10.1200/JCO.2017.73.7916](https://doi.org/10.1200/JCO.2017.73.7916)
2. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ*. 2015;350. [doi:10.1136/BMJ.H2147](https://doi.org/10.1136/BMJ.H2147)
3. Project Pragmatica | FDA. Accessed October 24, 2023. <https://www.fda.gov/about-fda/oncology-center-excellence/project-pragmatica>
4. Mangat PK, Garrett-Mayer E, Perez JK, Schilsky RL. The Targeted Agent and Profiling Utilization Registry Study: A pragmatic clinical trial. *Clin Trials*. Published online 2023. [doi:10.1177/17407745231182013](https://doi.org/10.1177/17407745231182013)
5. Dexamethasone in Hospitalized Patients with Covid-19. *New England Journal of Medicine*. 2021;384(8):693–704. [doi:10.1056/NEJMOA2021436/SUPPL_FILE/NEJMOA2021436_DATA-SHARING.PDF](https://doi.org/10.1056/NEJMOA2021436/SUPPL_FILE/NEJMOA2021436_DATA-SHARING.PDF)
6. Pragmatica-Lung Cancer Treatment Trial - NCI. Accessed October 24, 2023. <https://www.cancer.gov/types/lung/research/pragmatica-lung-cancer-trial>
7. Bekelman JE, Lu H, Pugh S, et al. Pragmatic randomised clinical trial of proton versus photon therapy for patients with non-metastatic breast cancer: the Radiotherapy Comparative Effectiveness (RadComp) Consortium trial protocol. *BMJ Open*. 2019;9(10). [doi:10.1136/BMJOPEN-2018-025556](https://doi.org/10.1136/BMJOPEN-2018-025556)
8. Enhancing the Diversity of Clinical Trial Populations-Eligibility Criteria, Enrollment Practices, and Trial Designs Guidance for Industry. Published online 2020. Accessed October 24, 2023. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/enhancing-diversity-clinical-trial-populations-eligibility-criteria-enrollment-practices-and-trial>
9. Mahoney MR, Sargent DJ, O'Connell MJ, Goldberg RM, Schaefer P, Buckner JC. Dealing with a deluge of data: An assessment of adverse event data on North Central Cancer Treatment Group trials. *Journal of Clinical Oncology*. 2005;23(36):9275–9281. [doi:10.1200/JCO.2004.00.0588](https://doi.org/10.1200/JCO.2004.00.0588)
10. Kaiser LD, Melemed AS, Preston AJ, et al. Optimizing Collection of Adverse Event Data in Cancer Clinical Trials Supporting Supplemental Indications. *Journal of Clinical Oncology*. 2010;28(34):5046. [doi:10.1200/JCO.2010.29.6608](https://doi.org/10.1200/JCO.2010.29.6608)
11. Hillman SL, Mandrekar SJ, Bot B, et al. Evaluation of the value of attribution in the interpretation of adverse event data: a North Central Cancer Treatment Group and American College of Surgeons Oncology Group investigation. *J Clin Oncol*. 2010;28(18):3002–3007. [doi:10.1200/JCO.2009.27.4282](https://doi.org/10.1200/JCO.2009.27.4282)
12. Le-Rademacher JG, Storrick EM, Jatoi A, Mandrekar SJ. Physician-Reported Experience and Understanding of Adverse Event Attribution in Cancer Clinical Trials. *Mayo Clin Proc Innov Qual Outcomes*. 2019;3(2):176–182. [doi:10.1016/J.MAYOCPHQ.2019.01.002](https://doi.org/10.1016/J.MAYOCPHQ.2019.01.002)
13. Clinical Trials Transformation Initiative - CTTI. Accessed October 24, 2023. <https://ctti-clinicaltrials.org/>
14. Point-of-Care Clinical Trials: Integrating Research and Care Delivery Point-of-Care Clinical Trials: Integrating Research and Care Delivery 2 Point-of-Care Clinical Trials: Integrating Research and Care Delivery. Published online 2022. Accessed October 24, 2023. <https://>

healthpolicy.duke.edu/projects/real-world-evidence-collaborative.

15. ASCO Awarded \$11M to Study Oral Drug Dosing Strategies in Older Adults with Metastatic Breast Cancer Receiving CDK4/6 Inhibitors | ASCO. Accessed October 24, 2023. <https://old-prod.asco.org/about-asco/press-center/news-releases/asco-awarded-11m-study-oral-drug-dosing-strategies-older>
16. Schwartz D, Lellouch J. Explanatory and Pragmatic Attitudes in Therapeutical Trials. *J Clin Epidemiol*. 2009;62(5):499-505. [doi:10.1016/j.jclinepi.2009.01.012](https://doi.org/10.1016/j.jclinepi.2009.01.012)
17. mCODE: Creating a Set of Standard Data Elements for Oncology EHRs | ASCO. Accessed October 24, 2023. <https://old-prod.asco.org/news-initiatives/current-initiatives/cancer-care-initiatives/mcode-standard-data-ehr>
18. White Paper: Pragmatic Randomized Trials: Considerations for Design and Implementation - Evidera. Accessed October 24, 2023. <https://www.evidera.com/pragmatic-randomized-trials-considerations-for-design-and-implementation/>

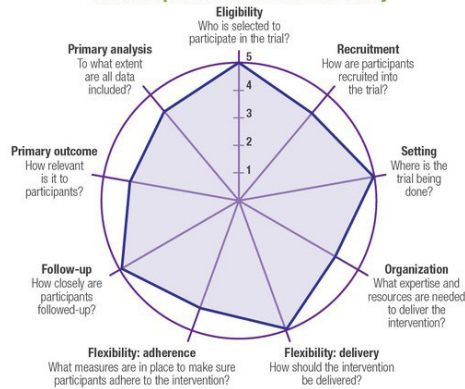
Abbreviations of Terms

AEs	Adverse Events
DHT	Digital Health Technologies
ECOG	Eastern Cooperative Oncology Group
FACT-G	Functional Assessment of Cancer Therapy- General
FDA	Food and Drug Administration
HRQOL	Health-Related Quality of Life
mCODE	minimal Common Oncology Data Elements
NCCN	National Comprehensive Cancer Network
NCTN	NCI National Clinical Trials Network
OCE	Oncology Center of Excellence
OS	Overall Survival
PCT	Pragmatic Clinical Trial
PFS	Progression-Free Survival
PRECIS	Pragmatic Explanator Continuum Indicator Summary
PRO-CTCAE	Patient-Reported Outcome Common Terminology Criteria for Adverse Events
PROMIS-29	Patient-Reported Outcomes Measurement Information System
TTD	Time to Treatment Discontinuation
TTNT	Time to Next Treatment

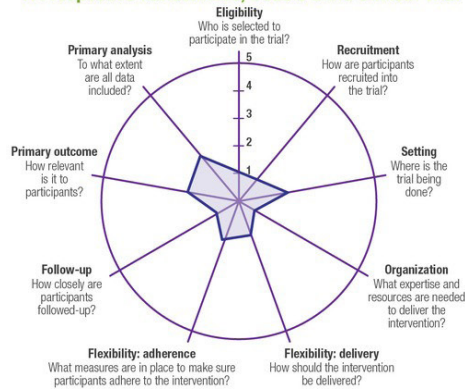
Appendix 1

The PRECIS 2 tool highlights the spectrum of elements that may be more or less pragmatic for a specific study, dependent on the regulatory and clinical context of the trial. A trial incorporating pragmatic elements (see Pragmatic Randomized Trial) may not utilize each element in the most pragmatic manner, or utilize every element.¹⁸

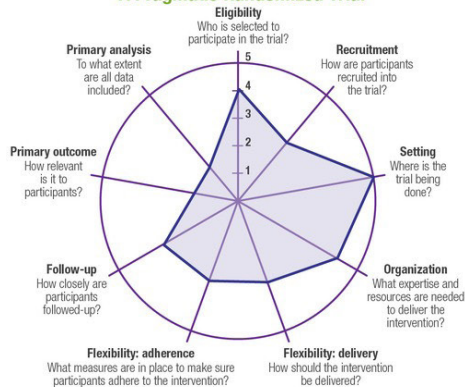
A Retrospective Observational Study



A Prospective Randomized, Double Blind Clinical Trial



A Pragmatic Randomized Trial



Impact of COVID-19 pandemic mitigation strategies on industry and NCI cancer treatment trials.

Joseph M. Unger, Caroline Schenkel, Hillary Stires, Laura Levit, Mark Stewart, Brittany Avin McKelvey, Beverly Canin, Emily Van Meter Dressler, Keith Flaherty, Peter Fredette, Lee Jones, Margaret McCann, Therica Miller, Adedayo A. Onitilo, Frances M. Palmieri, Timil Patel, Rocio Paul, Gary L. Smith, Suanna Steeby Bruinooge, Ajjai Shivaram Alva, ASCO Staff Authors; Fred Hutchinson Cancer Research Center, Seattle, WA; American Society of Clinical Oncology, Alexandria, VA; Friends of Cancer Research, Washington, DC; Breast Cancer Options, Rhinebeck, NY; Wake Forest University School of Medicine, Winston-Salem, NC; Massachusetts General Hospital, Boston, MA; EQRx, Cambridge, MA; Fight Colorectal Cancer, Arlington, VA; Merck & Co, Inc., Rahway, NJ; Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY; Wisconsin NCORP, Marshfield, WI; Sarah Cannon Research Institute, Nashville, TN; U.S. Food and Drug Administration, Silver Spring, MD; National Cancer Institute, Bethesda, MD; Cancer Therapeutics Evaluation Program, National Cancer Institute, Bethesda, MD; Department of Internal Medicine, Division of Hematology and Oncology, University of Michigan, Ann Arbor, MI














Background: The onset of the COVID-19 pandemic resulted in major disruptions in enrollment and conduct of cancer clinical trials. In response, regulators allowed modifications to traditional trial processes to enable clinical research and care to continue. The American Society of Clinical Oncology and Friends of Cancer Research established a task force to evaluate how sponsors perceived the impact of these mitigation strategies on data quality and overall trial conduct. **Methods:** The study used a survey and live interviews of industry and National Cancer Institute (NCI) cooperative group sponsors of cancer treatment trials. Sponsors with trials active in the United States from January 2015-May 2022 were eligible. We assessed perceived impacts of the pandemic on protocol deviations, types of mitigation strategies, trial closures, dropouts, adverse events (AEs), and data integrity. Descriptive statistics were used for survey data summaries. Key findings from semi-structured interviews were described by theme. **Results:** Of forty-one eligible sponsors, 20 (49%; 15/36 industry and 5/5 cooperative group) completed the survey; eleven (55%; 7/15 industry and 4/5 cooperative group) were interviewed. Sixty percent of sponsors reported large portfolios (>10 trials) of phase II and/or phase III trials. The most widely adopted mitigation strategies were remote distribution of oral anticancer therapies (70%), remote consenting (65%), and remote symptom monitoring for AEs (65%). The proportion of sponsors reporting a “substantial” increase in protocol deviations compared to the pre-pandemic period dropped from 42% in the initial wave (March-April 2020) to 16% thereafter. Sponsors primarily reported “no change” in trial drop-out rates (77%), the number of trials closed due to low accrual (90%), or rates of AEs (81%) at any point during the pandemic. Overall, most (83%) respondents reported the pandemic had “minimal” (14) or “no” impact (1) on data integrity. In interviews, many sponsors reported persistent time delays in data entry related to labor shortages at sites. **Conclusions:** This study represents the first systematic evaluation of clinical trial sponsors about the impact of the COVID-19 pandemic on the conduct of cancer clinical trials. Remote clinical trial conduct mitigation strategies were broadly adopted. Despite an observed increase in protocol deviations, most sponsors reported the pandemic had minimal or no impact on data integrity. The COVID-19 pandemic impacted how cancer clinical trials were performed and has likely accelerated a trend towards greater flexibility in trial conduct that was already emerging, with potential benefits of reduced burden and improved access for patients and sites. Future work is planned to further quantify the impact of the pandemic and trial mitigation strategies on the quality of trial data both overall and for historically underrepresented patient groups. Research Sponsor: None.

Scan to view poster



© 2023 by American Society of Clinical Oncology. Visit meetings.asco.org and search by abstract for disclosure information.
Downloaded from ascopubs.org by 69.138.219.233 on November 27, 2023 from 069.138.219.233
Copyright © 2023 American Society of Clinical Oncology. All rights reserved.

Sponsor Perspectives on the Impact of the COVID-19 Pandemic on Interventional Cancer Clinical Trial Protocols and Data Quality

Joseph M. Unger, PhD, MS¹ ; Hillary Stires, PhD² ; Laura A. Levit, JD³ ; Mark Stewart, PhD² ; Brittany Avin McKelvey, PhD² ; Beverly Canin⁴ ; Emily Dressler, PhD⁵; Keith Flaherty, MD⁶ ; Peter Fredette⁷; Lee Jones, MBA⁸ ; Peggy McCann, PhD⁹; Therica Miller, MBA, CCRP¹⁰ ; Adedayo A. Onitilo, MD, PhD, MSCR, FACP¹¹ ; Fran Palmieri, RN, MSN, OCN, CCRP¹²; Timil Patel, MD¹³; Rocio Paul, MSHS, CCRP, RAC¹⁴; Gary L. Smith, MT, MGA¹⁴; Suanna S. Bruinooge, MPH³ ; Elizabeth Garrett-Mayer, PhD, FSCT³ ; Xiudong Jennifer Lei, PhD¹⁵; Ajjai Alva, MD¹⁶; and Caroline Schenkel, MSc³ 

DOI <https://doi.org/10.1200/OP.23.00185>

ABSTRACT

PURPOSE The onset of the COVID-19 pandemic created major disruptions in the conduct of cancer clinical trials. In response, regulators and sponsors allowed modifications to traditional trial processes to enable clinical research and care to continue. We systematically evaluated how these mitigation strategies affected data quality and overall trial conduct.

METHODS This study used surveys and live interviews. Forty-one major industry and National Cancer Institute Network groups (sponsors) overseeing anticancer treatment trials open in the United States from January 2015 to May 2022 were invited to participate. Descriptive statistics were used for survey data summaries. Key themes from interviews were identified.

RESULTS Twenty sponsors (48.8%; 15 industry and five Network groups) completed the survey; 11/20 (55.0%) participated in interviews. Sponsors predominantly (n = 12; 60.0%) reported large (≥11 trials) portfolios of phase II and/or phase III trials. The proportion of sponsors reporting a moderate (9) or substantial (8) increase in protocol deviations in the initial pandemic wave versus the pre-pandemic period was 89.5% (17/19); the proportion reporting a substantial increase dropped from 42.1% (n = 8/19) in the initial wave to 15.8% (n = 3/19) thereafter. The most commonly adopted mitigation strategies were remote distribution of oral anticancer therapies (70.0%), remote adverse event monitoring (65.0%), and remote consenting (65.0%). Most respondents (15/18; 83.3%) reported that the pandemic had minimal (n = 14) or no impact (n = 1) on overall data integrity.

CONCLUSION Despite nearly all sponsors observing a temporary increase in protocol deviations, most reported the pandemic had minimal/no impact on overall data integrity. The COVID-19 pandemic accelerated an emerging trend toward greater flexibility in trial conduct, with potential benefits of reduced burden on trial participants and sites and improved patient access to research.

ACCOMPANYING CONTENT

[Data Supplement](#)

Accepted July 17, 2023
Published August 29, 2023

JCO Oncol Pract 19:907-916
© 2023 by American Society of
Clinical Oncology



[View Online Article](#)

INTRODUCTION

Clinical trials are key to advancing new treatments for patients with cancer. At the outset of the COVID-19 pandemic, patient enrollment and treatment in cancer clinical trials dropped dramatically.^{1,2} In response, the National Cancer Institute (NCI) and the US Food and Drug Administration (FDA) issued guidance statements that provided greater flexibility for sponsors overseeing clinical trial processes in oncology.^{3,4} The goal was to allow modifications to trial

protocols to mitigate the impact of pandemic-related disruptions on trial participants and clinical research. There is increasing evidence that widespread adoption of these modifications enabled a rebound of cancer treatment trial enrollment following an initial steep decrease.⁵⁻⁹ However, a knowledge gap remains about the impact of COVID-19-era protocol modifications on the quality of clinical trial data.¹⁰

To address this, ASCO and Friends of Cancer Research (*Friends*) convened a task force to evaluate the impact of the

CONTEXT

Key Objective

In response to the onset of the COVID-19 pandemic, major cancer clinical trial sponsors were allowed to modify traditional trial processes to enable clinical research and care to continue. Our aim was to evaluate how these mitigation strategies affected data quality and overall trial conduct.

Knowledge Generated

This study shows that major clinical trial sponsors widely adopted the recommended mitigation strategies to help maintain the conduct of clinical trials during the COVID-19 pandemic. Although a temporary increase in protocol deviations was reported by most sponsors, most also reported that the pandemic had minimal/no impact on overall data integrity.

Relevance

Our findings suggest that the strategies implemented during the pandemic to provide greater flexibility in trial conduct may reduce the burden of trial participation for patients and sites with limited adverse consequences for trial data.

COVID-19 pandemic on the conduct of cancer clinical treatment trials. This task force included representation from physician investigators and clinical trial operations executives from academic- and community-based sites, NCI Network group and pharmaceutical industry sponsors, FDA, the NCI Cancer Therapy Evaluation Program (CTEP), patient advocates, biostatisticians, and a contract research organization. The goal was to assess the extent to which trial sponsors perceived that changes to protocols adopted during the pandemic affected data quality, an important consideration when evaluating whether efforts to modernize trial processes may make trials more accessible to patients and speed their conduct without adverse consequences.^{10,11}

METHODS

This study combined surveys with live interviews (Data Supplement, online only). All pharmaceutical companies and NCI Network groups sponsoring at least one anticancer treatment trial before (January 2015–February 2020) and after (March 2020–May 2022) the COVID-19 pandemic were eligible to participate. ClinicalTrials.gov was queried to develop a list of eligible sponsors. The study protocol was reviewed and classified as exempt research by WCG Institutional Review Board (IRB) in April 2022.

Definitions

Trial Eligibility

All survey and interview questions referred to interventional anticancer treatment trials of any modality (eg, systemic therapy [cytotoxic, immune, hormonal, targeted, etc], surgery, or radiation) sponsored by the organization that were open in the United States. Although many industry trials are operated in multiple countries, sponsors were asked to restrict their observations to trial activities located in the United States.

Time Windows

We defined the following time periods to organize our evaluation:

1. Pre–COVID–19: January 2015–December 2019
2. Immediately pre–COVID–19: January–February 2020
3. Initial wave: March–April of 2020
4. Post–initial wave: May 2020–May 2022

Outcomes

The primary data quality metric was protocol deviations, interpreted to represent nonadherence to stated treatment and data collection processes defined prospectively within trial protocols.¹² To ensure consistency, the following definition of a protocol deviation was provided: any noncompliance with IRB–approved protocol, including prospectively approved deviations and waivers. Furthermore, a significant or serious protocol deviation was defined as a protocol deviation that increases the potential risk to participants or affects the integrity of study data.

Our terminology is premised on published and anecdotal evidence that the COVID–19 outbreak had both direct (ie, reduced patient willingness to participate in trials) and indirect (mediated through the declaration of a public health emergency [PHE]) effects on the conduct of cancer clinical trials.¹³ Thus, we generally refer to impact of the COVID–19 pandemic itself—the underlying causal mechanism—even if, in some instances, the PHE was the more proximal cause of a consequence.

Survey

The task force developed a 35–item REDCap questionnaire. The electronic survey collected pre–COVID–19 and COVID–19–era data related to number and types of active treatment trials; trial openings, holds, and closures; organizational

protocol deviation definitions; volume and types of protocol deviations collected; mitigation strategies implemented; impact on adverse events (AEs) collected (where AEs were categorized as physician-reported grades 1 [mild] or 2 [moderate] v grades 3 [serious] or 4 [life-threatening] using standard NCI definitions)¹⁴; and impact on overall data integrity. Sponsors were not asked to perform any analyses before participating in the survey.

Representatives from eligible sponsors were invited to participate in the survey. Sponsors were encouraged to engage multiple staff within their organization to inform responses. The survey was open from May 10 to August 22, 2022. Sponsors were offered 30 days to complete the survey, with 7- to 45-day extensions allowed by request.

Interviews

All sponsors that completed the survey were invited to be interviewed. Sponsor organizations were interviewed between August 11 and October 3, 2022. Two ASCO and *Friends* staff members (interviewers) alternated serving as primary interviewer and note-taker. Interviews were conducted via video conference and recorded for analysis purposes.

An interview guide was developed concurrently with the survey. Sponsors received the guide before the interview and were encouraged to select representatives with relevant knowledge of oncology clinical trial conduct and data to participate (one interview per organization). A semistructured interview approach was employed using the interview guide questions and appropriate follow-up questions on the basis of survey responses.

Statistical and Evaluation Methods

Survey data were summarized using descriptive statistics. To compare aggregate trends in the adoption of mitigation strategies between industry and NCI Network groups, each mitigation strategy was treated as an independent opportunity; the total was summed and compared using a Fisher's exact test. To describe how the volume of protocol deviations in the initial wave compared with the pre- and post-pandemic periods, we assessed the difference in paired Likert scale (1 = substantial increase, 2 = moderate increase, 3 = no change, 4 = moderate decrease, 5 = substantial decrease) scores, adjusted for organization type (industry v Network groups) using linear regression.

Interview data were evaluated using a three-step thematic analysis. First, interviewers classified the sponsor representatives' comments into three overarching categories that corresponded to the research objectives: (1) major protocol deviations collected during the pandemic, (2) key takeaways and impacts, and (3) future directions. Second, interviewers reconciled any points of discordance. Finally, interviewers agreed upon commonly occurring themes within the categories.

RESULTS

Forty-one eligible sponsor organizations were invited to participate; 21 (51.2%; all pharmaceutical company sponsors) did not participate, including 11 (26.8%) that did not respond, 8 (19.5%) that declined, and 2 (4.9%) that dropped out before completing the survey. Twenty sponsors (48.8%) completed the survey, including 15 pharmaceutical companies and all five NCI-sponsored Network groups. Representatives from 11/20 participating sponsors (55.0%) were interviewed. The median number of sponsor representatives per call was 3 (range, 1-8). Representatives who provided data for the survey and interviews were in data management, clinical development, regulatory science/affairs, statistics/biostatistics, and medical writing roles. Most (27/34) representatives were in director- or vice president-level roles (including associate, senior, and executive).

Where we did not receive a survey response from all sponsors, the denominator used for analysis is specified; otherwise, it is 20.

Survey Findings

Sponsor Characteristics

In January-February 2020, four sponsors (20.0%) had 0-5 open phase II trials (small portfolios), 4 (20.0%) had 6-10 trials (medium), and 12 (60.0%) had ≥11 (large; [Table 1](#)). The majority of all sponsors (60.0%) also had large portfolios of open phase III trials. Among industry sponsors, about half (46.7%) reported large phase II portfolios and about half (46.7) reported large phase III portfolios. NCI Network groups were more likely ($P = .05$) to have reported large portfolios of both phase II (100%) and phase III (100%) trials.

Protocol Deviations

Sponsors' definitions of significant or serious protocol deviations referenced the potential impact on participant safety and data and scientific integrity, similar to the definition provided in study materials. Before the COVID-19 pandemic, nearly all (≥90%) sponsors classified eligibility- or consent-related issues, treatment-related issues, and assessment-, lab-, or imaging-related issues (including missed and out-of-window visits) as protocol deviations, with minor exceptions. Sponsors were evenly divided (yes, 50.0%; no, 50.0%) in considering device-related issues as protocol deviations. A significant minority reported that lab/imaging/test/procedure after withdrawal of consent (20.0%) or imaging performed by a nonqualified site (25.0%) was not considered protocol deviations.

Nearly all sponsors (17/19; 89.5%) reported a moderate (9/19; 47.4%) or substantial (8/19; 42.1%) increase in volume

TABLE 1. Sponsor Responses to Selected Survey Questions

Survey Question	Sponsor Category, No. (%)		
	All	Industry	NCI Network Group
Baseline (January-February 2020) portfolio characteristics	N = 20	N = 15	N = 5
No. of phase I trials active			
0-5 (small)	7 (35.0)	4 (26.7)	3 (60.0)
6-10 (medium)	1 (5.0)	1 (6.7)	0 (0)
11 or more (large)	12 (60.0)	10 (66.7)	2 (40.0)
No. of phase II trials active			
0-5 (small)	4 (20.0)	4 (26.7)	0 (0)
6-10 (medium)	4 (20.0)	4 (26.7)	0 (0)
11 or more (large)	12 (60.0)	7 (46.7)	5 (100.0) ^a
No. of phase III trials active			
0-5 (small)	7 (35.0)	7 (46.7)	0 (0)
6-10 (medium)	1 (5.0)	1 (6.7)	0 (0)
11 or more (large)	12 (60.0)	7 (46.7)	5 (100.0) ^a
Rate impact to overall data integrity of protocol deviations during the pandemic	N = 18	N = 14	N = 4
No impact	1 (5.6)	1 (7.1)	0 (0)
Minimal impact	14 (77.8)	11 (78.6)	3 (75)
Somewhat negative impact	3 (16.7)	2 (14.3)	1 (25)
Very negative impact	0 (0)	0 (0)	0 (0)
Extremely negative impact	0 (0)	0 (0)	0 (0)
Changes to a patient's protocol-specified treatment plan that were typically defined as a PD in the pre-COVID-19 period (January 2015-December 2019)	N = 20	N = 15	N = 5
Patient did not meet eligibility criteria	19 (95)	15 (100)	4 (80)
Incorrect or incomplete informed consent process	19 (95)	15 (100)	4 (80)
Reconsent not obtained as required	19 (95)	15 (100)	4 (80)
Failure to follow treatment randomization	19 (95)	14 (93.3)	5 (100)
Failure to discontinue treatment	19 (95)	15 (100)	4 (80)
Administration of non-protocol-defined therapy to treat subject's disease or concomitant medication used was not permitted per protocol	19 (95)	15 (100)	4 (80)
SAE reported out of window	18 (90)	14 (93.3)	4 (80)
Agent-related issues	20 (100)	15 (100)	5 (100)
Device-related issues	10 (50)	8 (53.3)	2 (40)
Schedule-related issues	19 (95)	15 (100)	4 (80)
Physical assessment deviation	18 (90)	14 (93.3)	4 (80)
Patient does not have safety follow-up as required	19 (95)	15 (100)	4 (80)
Lab/imaging/test/procedure after withdrawal of consent	16 (80)	14 (93.3)	2 (40)
Lab, imaging, or other test/procedure not done	19 (95)	15 (100)	4 (80)
Imaging performed by a nonqualified site	15 (75)	13 (86.7)	2 (40)
Other imaging-related issues	16 (80)	12 (80)	4 (80)
Average volume of PDs in March-April 2020 compared with January 2015-December 2019	N = 19	N = 15	N = 4
Substantial increase after March 2020	8 (42.1)	8 (53.3)	0 (0)
Moderate increase after March 2020	9 (47.4)	7 (46.7)	2 (50)
No measurable change after March 2020	1 (5.3)	0 (0)	1 (25)
Moderate decrease after March 2020	1 (5.3)	0 (0)	1 (25)
Substantial decrease after March 2020	0 (0)	0 (0)	0 (0)
Average volume of PDs after May 2020-May 2022 compared with January 2015-December 2019	N = 19	N = 15	N = 4
Substantial increase after March 2020	3 (15.8)	2 (13.3)	1 (25)
Moderate increase after March 2020	13 (68.4)	11 (73.3)	2 (50)
No measurable change after March 2020	2 (10.5)	1 (6.7)	1 (25)

(continued on following page)

TABLE 1. Sponsor Responses to Selected Survey Questions (continued)

Survey Question	Sponsor Category, No. (%)		
	All	Industry	NCI Network Group
Moderate decrease after March 2020	0 (0)	0 (0)	0 (0)
Substantial decrease after March 2020	1 (5.3)	1 (6.7)	0 (0)
Percentage of trials delayed or otherwise impacted by holds at sites during March-April 2020	N = 17	N = 15	N = 2
0%-25%	8 (47.1)	7 (46.7)	1 (50)
26%-50%	1 (5.9)	1 (6.7)	0 (0)
51%-75%	5 (29.4)	4 (26.7)	1 (50)
>76%	3 (17.6)	3 (20.0)	0 (0)
Percentage of trials affected by closures at sites during March-April 2020	N = 18	N = 15	N = 3
0%-25%	15 (83.3)	13 (86.7)	2 (66.7)
26%-50%	1 (5.6)	0 (0)	1 (33.3)
51%-75%	0 (0)	0 (0)	0 (0)
>76%	2 (11.1)	2 (13.3)	0 (0)
Impact to cancer treatment trial dropout rates since March 2020	N = 17	N = 14	N = 3
Increased during the pandemic and have not returned to pre-pandemic levels	0 (0)	0 (0)	0 (0)
Increased during the pandemic but have returned to pre-pandemic levels	2 (11.8)	2 (14.3)	0 (0)
Decreased during the pandemic and have not returned to pre-pandemic levels	2 (11.8)	1 (7.1)	1 (33.3)
Decreased during the pandemic but have returned to pre-pandemic levels	0 (0)	0 (0)	0 (0)
No change observed	13 (76.5)	11 (78.6)	2 (66.7)
Change to rates of reported grade 3-4 AEs (ie, severe/medically significant or life-threatening/disabling events) during the pandemic ^b	N = 16	N = 13	N = 3
Increased during the pandemic and have not returned to pre-pandemic levels	1 (6.2)	1 (7.7)	0 (0)
Increased during the pandemic but have returned to pre-pandemic levels	1 (6.2)	1 (7.7)	0 (0)
Decreased during the pandemic and have not returned to pre-pandemic levels	0 (0)	0 (0)	1 (33.3)
Decreased during the pandemic but have returned to pre-pandemic levels	1 (6.2)	0 (0)	0 (0)
No change observed	13 (81.2)	11 (84.6)	2 (66.7)

Abbreviations: AEs, adverse events; NCI, National Cancer Institute; PDs, protocol deviations; SAE, serious adverse event.

^aStatistically significantly larger by Fisher's exact test, $P \leq .05$.

^bNo. (%) for changes to rates of reported grade 1-2 AEs (ie, mild/asymptomatic or bothersome but not dangerous events) were identical.

of protocol deviations in the first wave of the COVID-19 pandemic (March–April 2020; Fig 1). After the initial wave (beginning May 2020), the increase in volume compared with the pre-pandemic period was lower ($P = .03$ in linear regression), with only 3/19 (15.8%) describing the increase as substantial. However, an additional 13/19 (68.4%) reported a moderate increase in protocol deviations after the initial wave, indicating that the level of deviations had not returned to pre-pandemic levels.

Sponsors were also asked to assess whether more serious/significant protocol deviations were being reported at the time of the survey, compared with pre-COVID-19. Among 16 sponsors providing data, 10 (62.5%) stated that the average number of serious protocol deviations was stable relative to the number of minor protocol deviations. Five (31.3%) reported that the average number of serious protocol deviations had increased compared with the pre-pandemic period, and one sponsor reported a decrease.

Nearly all sponsors (19; 95.0%) collected protocol deviations attributable to the COVID-19 pandemic. Most respondents

(17; 85.0%) did not collect data regarding whether protocol deviations were attributable to study staff or participant decision making.

Mitigation Strategies

The most common mitigation strategies adopted between January and May 2020 were remote distribution of oral anticancer therapies (70.0%); remote symptom monitoring of AEs (65.0%); and e-consenting or remote informed consent (65.0%; Fig 2). Other commonly (ie, $\geq 50\%$ overall) adopted strategies included remote collection of patient-reported outcomes (55.0%), remote routine laboratory testing (50.0%), remote imaging (50.0%), and remote study-specific laboratory tests (50.0%). Few sponsors adopted the strategy of remote study-required biopsies (10%). All sponsors reported either *yes* or *no* for all 12 specified mitigation strategies; thus, across the 20 sponsors, there were a total of 240 opportunities to adopt the 12 mitigation strategies, and nearly half (110/240; 45.8%) were adopted. This proportion did not differ

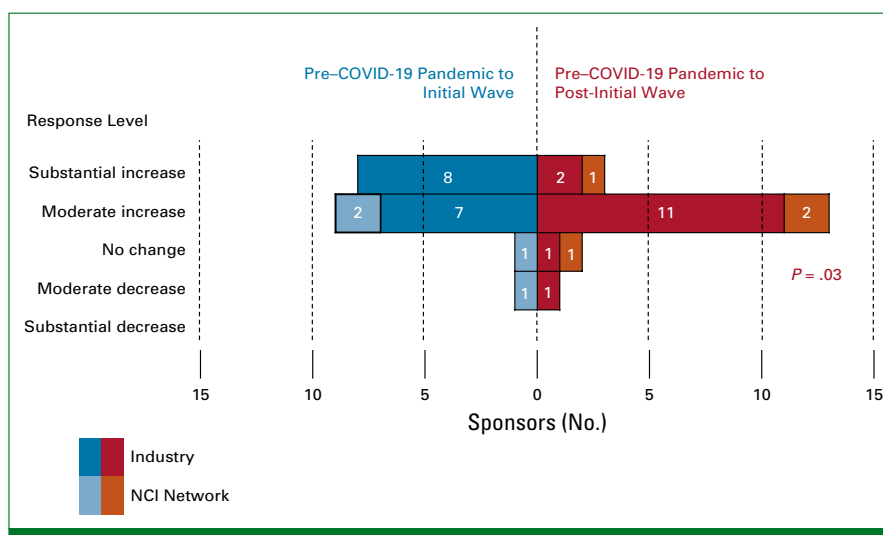


FIG 1. Change in volume of protocol deviations compared with the pre-pandemic period (N = 19; industry = 15, NCI Network groups = 4). The horizontal bars indicate the number of sponsors indicating the level of change in volume of protocol deviations between the initial wave versus the pre-COVID-19 pandemic period (in blue) and between the post-initial wave versus the pre-COVID-19 pandemic period (in red). NCI, National Cancer Institute.

between NCI Network groups (32/60 opportunities, 53.3%) compared with industry (78/180 opportunities; 43.3%; $P = .18$).

Impact of Trial Holds and Closures at Study Sites

Among 17 respondents, during the initial wave, trial holds were reported for none/few (0%–25%) sites by eight (47.1%) sponsors, some (26%–50%) sites by one (5.9%) sponsor, most (51%–75%) sites by five (29.4%) sponsors, and nearly all/all (>76%) sites by three (17.6%) sponsors. Among nine sponsors who encountered holds, average hold time at sites in March–April 2020 ranged from 2 to 12 weeks (mean, 7.3; standard deviation, 3.6).

Most sponsors (15/18; 83.3%) reported that closures affected none/few of their sites during the pandemic’s first wave.

Trial holds and closures at sites occurred less frequently after the initial wave. Among 17 respondents, trials delayed or affected by holds were reported as much or somewhat lower from May 2020 to May 2022 compared with March–April 2020 by 10 (58.8%) sponsors and the same by 6 (35.3%) sponsors. Similarly, among 16 respondents, trials affected by closures were reported as much or somewhat lower by 9 (56.3%) sponsors and the same by 7 (43.8%) sponsors.

Dropouts and Trial Closures

Among 17 respondents, most (n = 13; 76.5%) reported no change in patient dropout rates during the pandemic. Two

industry sponsors observed an increase in dropout rates that had since returned to the pre-pandemic level, and one industry sponsor and one NCI Network group sponsor observed a decrease in dropout rates that had not increased back to the pre-pandemic level.

Among 19 respondents, most (n = 17; 89.5%) observed no change in the number of trials closed because of low accrual during the pandemic, while two respondents (both NCI Network groups) reported a decrease.

AEs

Among 16 respondents, 13 (81.3%) reported no change in rates of both grades 1–2 and 3–4 AEs during the pandemic. One industry sponsor each indicated that rates of reported grades 1–2 and 3–4 AEs increased and have not returned or increased and have returned, respectively, while one NCI Network group reported that levels decreased and have returned.

Overall Impact on Data Integrity

Sponsors were asked to rate the impact of the level of protocol deviations on overall data integrity during the COVID-19 pandemic using a five-point Likert-type scale with undefined response anchors (ie, left up to respondent interpretation). Among 18 respondents, the majority (n = 15; 83.3%) reported a minimal impact (14) or no impact (1) on overall data integrity (Fig 3).

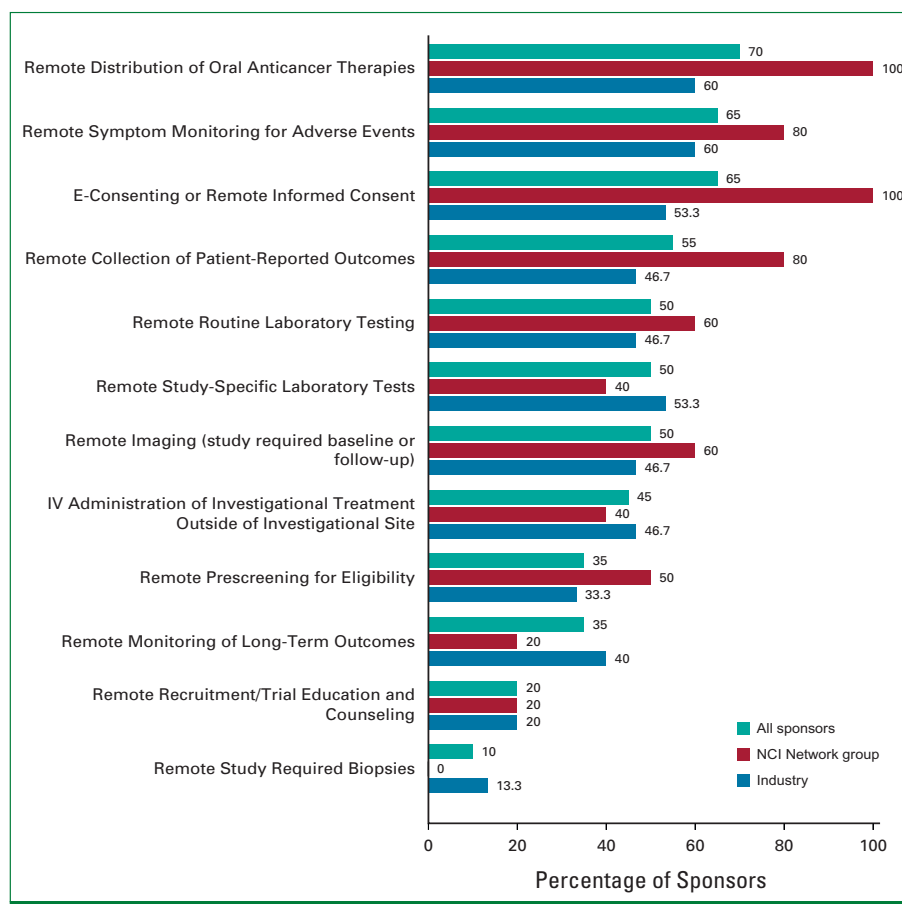


FIG 2. Mitigation strategies adopted by sponsors between January and May 2020 (%; N = 20; industry = 15, NCI Network groups = 5). E, electronic; IV, intravenous; NCI, National Cancer Institute.

Interview Findings

Follow-up interviews were conducted among 11 sponsors (55.0%), including seven industry sponsors and four NCI Network groups. Representatives from NCI CTEP were also interviewed, using a modified version of the sponsor interview guide.

Key findings that emerged from thematic analysis included the perception that the pandemic accelerated the inclusion of remote elements in protocols, especially e-consenting, remote distribution of oral investigative therapies, and virtual patient visits (Table 2). Sponsors reported that disruption to trial activities was mostly limited to March–April 2020 and was more likely to affect recruitment and enrollment rather than treatment continuity.

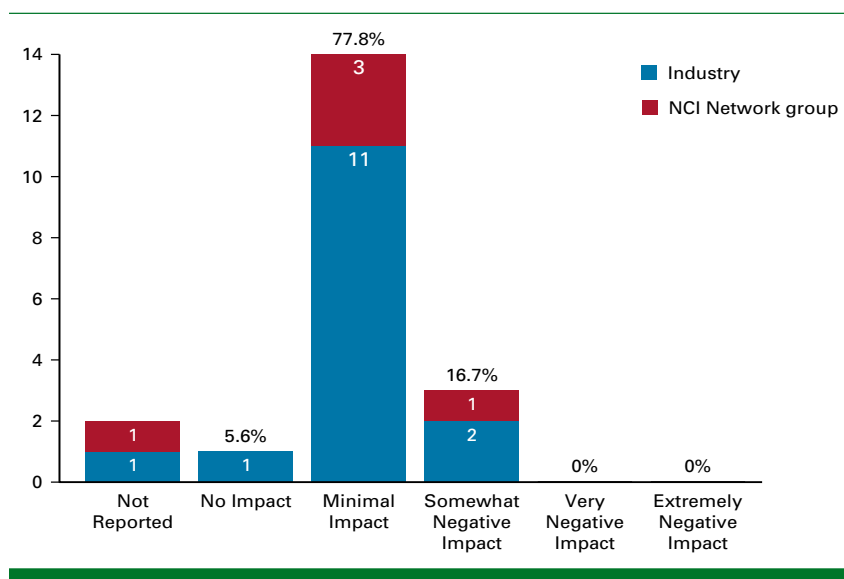
Additionally, sponsors relied upon the FDA and NCI guidance documents for establishing COVID-19–era procedures and indicated that these were essential to mitigating negative

effects on trials and patients, particularly during the initial wave. Most sponsors reported that substantial staff shortages and turnover at sites led to persistent data entry lags compared with pre-pandemic timelines, although nearly all sponsors perceived minimal impact of the pandemic on overall data integrity.

Finally, most sponsors reported the intention to allow the mitigation measures to continue after the expiration of the PHE, although some sponsors also reported concern about appropriate clinical oversight of remote treatment or assessment and the regulatory burden of remote auditing.

DISCUSSION

This study represents a systematic evaluation of major clinical trial sponsors about the impact of the COVID-19 pandemic and its associated PHE on the conduct of cancer clinical trials, and thus, provides critical evidence from key collaborators to fill an evidence gap.¹⁰ On the basis of survey and interviews, we found that most respondents observed an



Overall impact of protocol deviations on data integrity. Overall percentages among those with known data (ie, excluding not reported) are shown at the top of the bars. Counts by sponsor type are shown within the bars in white font. NCI, National Cancer Institute.

TABLE 2. Key Interview Findings

Theme	Details
Acceleration of existing trends for remote trial elements	The COVID-19 PHE accelerated the inclusion of remote elements in protocols, particularly e-consenting, remote distribution of oral investigative therapies, and virtual patient visits or remote assessments. This was true to a lesser extent for remote auditing. During the pandemic, changes to study conduct were often incorporated into trial protocols or categorized as operational or logistical changes, rather than documented as protocol deviations. Sponsors reported that inclusion of new flexibilities was left to the discretion of PIs rather than required.
Limited disruption to trial activities after the initial COVID-19 pandemic wave	Sponsors reported that disruption to trial activities was mostly limited to March-April 2020 and was more likely to affect recruitment/enrollment than treatment continuity for enrolled patients. Some sponsors indicated that patients in early-phase trials were more likely to continue visiting primary study sites in person than patients on later-phase trials, but the impact of the pandemic on different trial types (eg, phase, disease area) varied across sponsors.
Reliance on timely FDA and NCI guidance documents	Both industry and NCI Network group sponsors relied on FDA and NCI guidance documents as their primary reference points for establishing COVID-19-era procedures. Sponsors emphasized that the timeliness of those guidance documents was essential to mitigating the pandemic's negative effects on trials and patients, particularly during the first wave. Although nearly all sponsors reported flagging PDs that were specifically attributable to the pandemic, few used the data other than for regulatory/NCI submissions as required. All NCI Network groups reported that pre-existing NCI flexibilities enabled swift adaption to the COVID-19 pandemic. As described by the NCI, guidance documents were developed with input from all major branches and updated to incorporate feedback from the NCI Network groups. NCI Network groups reported that NCI guidance on documentation may have introduced consistency where procedures were previously inconsistent among NCI Network groups and sites (eg, regarding the documentation of minor PDs).
Persistent lags in data submission, but minimal impact on overall data integrity	Most sponsors observed that data entry lags pre-pandemic timelines and the NCI reported observing data missingness and delays in its audits. Sponsors perceived that staff shortages and turnover at sites was the primary cause of delays, but none had conducted any specific analysis of the effect. Despite those delays, nearly all sponsors perceived minimal impact of the COVID-19 pandemic on overall data integrity.
Ongoing assessment of whether and how decentralized elements will be incorporated permanently	Most sponsors indicated an intent to allow remote consenting, treatment, and monitoring options introduced during the PHE to continue after the expiration of the US public health emergency. Many representatives perceived that when implemented appropriately, these measures can ease patient burden while preserving data integrity. Some sponsors, however, reported that investigators are concerned about insufficient oversight of remote treatment or assessment.

Abbreviations: FDA, US Food and Drug Administration; NCI, National Cancer Institute; PHE, public health emergency; PIs, principal investigators.

increase in protocol deviations and many reported persistent lags in data collection >2 years later. However, the majority (83.3%) reported that the pandemic had minimal/no impact on overall data integrity. Sponsors indicated that remote elements were broadly implemented to minimize disruptions to enrollment and care of trial participants.

The COVID-19 pandemic was associated with severe interruptions in routine care for patients with cancer and for patients wishing to receive their care in clinical trials.¹⁵⁻²⁰ In part, this was related to fear of exposure to SARS-CoV-2. In one study, among the one fifth of patients reporting they were less likely to participate in a clinical trial during the pandemic, most reported they were fearful of contracting SARS-CoV-2.¹³ Patients with cancer are often immunocompromised because of their cancer or its treatment; as such, becoming infected with SARS-CoV-2 while receiving care at clinics is likely to exacerbate their existing clinical risks.²¹

Given these challenges, enrollment to cancer clinical treatment trials dropped precipitously during the initial COVID-19 pandemic wave.^{2,8} In response, NCI and FDA provided early guidance to sponsors about mitigation strategies that could help overcome the difficulties of conducting cancer clinical research during a PHE. Many of these mitigation strategies had been previously considered but not widely adopted.²² A focus has been on allowing protocol procedures and processes to be conducted outside of traditional specialized academic centers where the majority of trials are conducted. Proposals to decentralize clinical care outside of trials have included increased use of telemedicine for monitoring and evaluation, with accompanying documented benefits for reducing treatment and participation burdens on patients and their caregivers.^{23,24} Such proposals can be extended to the conduct of clinical trials as part of a broader effort to modernize clinical research.

Concerns about the potential impacts on data quality of decentralized approaches to clinical trial conduct have previously prevented their widespread adoption.^{25,26} However, the onset of the COVID-19 pandemic forced their rapid adoption, thus serving as a natural experiment to evaluate their impact. To our knowledge, this study for the first time demonstrates that these mitigation strategies were widely adopted by major sponsors with minimal or no perceived impact on overall data integrity. Many calls to further

evaluate whether to permanently incorporate decentralized elements into the conduct of clinical trials on the basis of the experience of the COVID-19 pandemic have been made.^{11,27-29} Importantly, since the aim of these strategies is also to reduce the burden of trial participation for patients, their adoption may have the salutary effect of improving representation for diverse patient populations. Recently, the FDA highlighted how decentralized trials may reduce barriers in access to trials and improve representation of historically underrepresented patient groups.¹⁰

The study is strengthened by high representation of industry organizations and NCI Network groups that sponsor most cancer treatment trials in the United States. The study is limited, however, by its reliance on voluntary survey and interview data alone. Sponsors were not asked to perform analyses before participating in the survey or interviews, although some conducted data aggregation/analysis before reporting findings. The number of days that sponsors spent completing the survey and number of representatives participating in the interviews varied and may have also influenced the results. Thus, the findings rely on variable levels of sponsors' internal analysis and on representatives' perceptions. "Furthermore, sponsors may have been less likely to report negative impacts of the pandemic on their data, leading to a potential bias." To help mitigate this possibility, a presurvey/interview confidentiality document informed sponsors that their responses would be wholly anonymized in all data presentations. On the basis of hypotheses generated by this work, the ASCO-Friends task force is leading a quantitative evaluation of clinical trial data to provide greater insight into the impact of the pandemic on data quality.

The COVID-19 pandemic affected the conduct of cancer clinical trials and accelerated a trend toward greater flexibility. The strategies implemented during the pandemic to provide greater flexibility in the execution of interventional clinical trial procedures, patient evaluation, and data ascertainment may improve clinical trial access and reduce the burden of participation for sites and patients without compromising trial data. Sponsors continue to include flexibilities in new protocols while still following regulatory requirements and guidance. Future work to quantify the impact of the pandemic on the quality of trial data is vital for informing recommendations about whether more flexible processes may become permanent fixtures in the conduct of oncology clinical trials.

AFFILIATIONS

¹Fred Hutchinson Cancer Center, Seattle, WA

²Friends of Cancer Research, Washington, DC

³American Society of Clinical Oncology (ASCO), Alexandria, VA

⁴Breast Cancer Options, Kingston, NY

⁵Wake Forest School of Medicine, Winston-Salem, NC

⁶Massachusetts General Hospital, Boston, MA

⁷EQRx (employed by IQVIA during study design phase), Cambridge, MA

⁸Fight Colorectal Cancer, Arlington, VA

⁹Merck, Rahway, NJ

¹⁰Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY

¹¹Marshfield Clinic Health System, Weston, WI

¹²Sarah Cannon Research Institute, Nashville, TN

¹³U.S. Food and Drug Administration, Center for Drug Evaluation and Research, Silver Spring, MD

¹⁴National Cancer Institute, Cancer Therapy Evaluation Program, Clinical Trials Monitoring Branch, Bethesda, MD

¹⁵University of Texas MD Anderson Cancer Center (employed by ASCO during study design phase), Houston, TX

¹⁶University of Michigan Medical Center, Ann Arbor, MI

CORRESPONDING AUTHOR

Joseph M. Unger, PhD, MS, SWOG Statistics and Data Management Center, Fred Hutchinson Cancer Center, 1100 Fairview Ave N, M3-C102, Seattle 19024, WA; e-mail: junger@fredhutch.org.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Disclosures provided by the authors are available with this article at DOI <https://doi.org/10.1200/OP.23.00185>.

AUTHOR CONTRIBUTIONS

Conception and design: Joseph M. Unger, Hillary Stires, Laura A. Levit, Mark Stewart, Brittany Avin McKelvey, Beverly Canin, Keith Flaherty, Lee

Jones, Peggy McCann, Therica Miller, Adedayo A. Onitilo, Timil Patel, Suanna S. Bruinooge, Elizabeth Garrett-Mayer, Ajjai Alva, Caroline Schenkel
Financial support: Joseph M. Unger

Administrative support: Joseph M. Unger, Caroline Schenkel

Provision of study materials or patients: Caroline Schenkel

Collection and assembly of data: Joseph M. Unger, Hillary Stires, Mark Stewart, Brittany Avin McKelvey, Emily Dressler, Peggy McCann, Caroline Schenkel

Data analysis and interpretation: Joseph M. Unger, Hillary Stires, Laura A. Levit, Mark Stewart, Brittany Avin McKelvey, Emily Dressler, Keith Flaherty, Peggy McCann, Therica Miller, Fran Palmieri, Timil Patel, Gary L. Smith, Suanna S. Bruinooge, Xiudong Jennifer Lei, Caroline Schenkel

Manuscript writing: All authors

Final approval of manuscript: All authors

Accountable for all aspects of the work: All authors

ACKNOWLEDGMENT

The authors thank the participating sponsor organizations and their individual representatives for their essential contributions to this work.

REFERENCES

- Unger JM, Blanke CD, LeBlanc M, et al: Association of the coronavirus disease 2019 (COVID-19) outbreak with enrollment in cancer clinical trials. *JAMA Netw Open* 3:e2010651, 2020
- Bakouny Z, Labaki C, Bhalla S, et al: Oncology clinical trial disruption during the COVID-19 pandemic: A COVID-19 and cancer outcomes study. *Ann Oncol* 33:836-844, 2022
- National Cancer Institute: Coronavirus guidance, 2022. https://ctep.cancer.gov/investigatorResources/corona_virus_guidance.htm
- US Food and Drug Administration: FDA guidance on conduct of clinical trials of medical products during COVID-19 pandemic, 2020. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/fda-guidance-conduct-clinical-trials-medical-products-during-covid-19-pandemic>
- American Association for Cancer Research: AACR report on the impact of COVID-19 on cancer research and patient care, 2022. <https://www.AACR.org/COVIDReport>
- Boughey JC, Snyder RA, Kantor O, et al: Impact of the COVID-19 pandemic on cancer clinical trials. *Ann Surg Oncol* 28:7311-7316, 2021
- Prindiville SA, Sarosy GA, Loose D, et al: Patterns of enrollment in cancer treatment trials during the COVID-19 pandemic at National Cancer Institute-designated cancer centers. *Cancer J* 28:111-117, 2022
- Unger JM, Xiao H, LeBlanc M, et al: Cancer clinical trial participation at the 1-year anniversary of the outbreak of the COVID-19 pandemic. *JAMA Netw Open* 4:e2118433, 2021
- Waterhouse DM, Harvey RD, Hurley P, et al: Early impact of COVID-19 on the conduct of oncology clinical trials and long-term opportunities for transformation: Findings from an American Society of Clinical Oncology survey. *JCO Oncol Pract* 16:417-421, 2020
- US Food and Drug Administration: Advancing oncology decentralized trials: Modernizing evidence generation, 2022. <https://www.fda.gov/about-fda/oncology-center-excellence/advancing-oncology-decentralized-trials>
- Pennell NA, Dillmon M, Levit LA, et al: American Society of Clinical Oncology road to recovery report: Learning from the COVID-19 experience to improve clinical research and cancer care. *J Clin Oncol* 39:155-169, 2021
- Galuchie L, Stewart C, Meloni F: Protocol deviations: A holistic approach from defining to reporting. *Ther Innov Regul Sci* 55:733-742, 2021
- Fleury ME, Farnier AM, Unger JM: Association of the COVID-19 outbreak with patient willingness to enroll in cancer clinical trials. *JAMA Oncol* 7:131-132, 2021
- National Cancer Institute: Cancer Therapy Evaluation Program. Common terminology criteria for adverse events (CTCAE), 2023. https://ctep.cancer.gov/protocoldevelopment/electronic_applications/ctc.htm
- Walker MJ, Wang J, Mazuryk J, et al: Delivery of cancer care in Ontario, Canada, during the first year of the COVID-19 pandemic. *JAMA Netw Open* 5:e228855, 2022
- Patt D, Gordan L, Diaz M, et al: Impact of COVID-19 on cancer care: How the pandemic is delaying cancer diagnosis and treatment for American seniors. *JCO Clin Cancer Inform* 4:1059-1071, 2020
- Caston NE, Lawhon VM, Smith KL, et al: Examining the association among fear of COVID-19, psychological distress, and delays in cancer care. *Cancer Med* 10:8854-8865, 2021
- Dhada S, Stewart D, Cheema E, et al: Cancer services during the COVID-19 pandemic: Systematic review of patient's and caregiver's experiences. *Cancer Manag Res* 13:5875-5887, 2021
- Sutcuoglu O, Yazici O, Ozet A, et al: Harmful consequences of COVID-19 fear in patients with cancer. *BMJ Support Palliat Care* 10.1136/bmjspcare-2020-002628 [epub ahead of print on December 21, 2020]
- Vanni G, Materazzo M, Pellicciaro M, et al: Breast cancer and COVID-19: The effect of fear on patients' decision-making process. *In Vivo* 34:1651-1659, 2020 (suppl 3)
- Zhang L, Zhu F, Xie L, et al: Clinical characteristics of COVID-19-infected cancer patients: A retrospective case study in three hospitals within Wuhan, China. *Ann Oncol* 31:894-901, 2020
- Khazin S, Coravos A: Decentralized trials in the age of real-world evidence and inclusivity in clinical investigations. *Clin Pharmacol Ther* 106:25-27, 2019
- Cox A, Lucas G, Marcu A, et al: Cancer survivors' experience with telehealth: A systematic review and thematic synthesis. *J Med Internet Res* 19:e11, 2017
- Pennell NA, Dicker AP, Tran C, et al: mHealth: Mobile technologies to virtually bring the patient into an oncology practice. *Am Soc Clin Oncol Educ Book* 37:144-154, 2017
- Coert RMH, Timmis JK, Boorsma A, et al: Stakeholder perspectives on barriers and facilitators for the adoption of virtual clinical trials: Qualitative study. *J Med Internet Res* 23:e26813, 2021
- Polhemus AM, Kadhim H, Barnes S, et al: Accelerating adoption of patient-facing technologies in clinical trials: A pharmaceutical industry perspective on opportunities and challenges. *Ther Innov Regul Sci* 53:8-24, 2019
- Adams DV, Long S, Fleury ME: Association of remote technology use and other decentralization tools with patient likelihood to enroll in cancer clinical trials. *JAMA Netw Open* 5:e2220053, 2022
- Fu S, Gerber DE, Beg MS: Decentralized clinical trials in oncology: Are we ready for a virtual-first paradigm? *J Clin Oncol* 41:181-185, 2023
- Goodson N, Wicks P, Morgan J, et al: Opportunities and counterintuitive challenges for decentralized clinical trials to broaden participant inclusion. *NPJ Digit Med* 5:58, 2022

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Sponsor Perspectives on the Impact of the COVID-19 Pandemic on Interventional Cancer Clinical Trial Protocols and Data Quality

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

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

Hillary Stires

Consulting or Advisory Role: Avalere Health

Emily Dressler

Research Funding: Omada Health

Other Relationship: ICF

Uncompensated Relationships: Society for Clinical Trials

Keith Flaherty

Stock and Other Ownership Interests: Clovis Oncology, Loxo, X4 Pharma, Strata Oncology, PIC Therapeutics, Apricity Health, Oncoceutics, FOGPharma, Tvardi Therapeutics, Checkmate Pharmaceuticals, Kinnate Biopharma, Scorpion Therapeutics, ALX Oncology, xCures, Monopteros Therapeutics, Vibliome Therapeutics, Transcode Therapeutics, Soley Therapeutics, Nextech Invest, Alterome Therapeutics

Consulting or Advisory Role: Novartis, Lilly, Oncoceutics, Tvardi Therapeutics, Takeda, Debiopharm Group, OmRx Oncology, Quanta Therapeutics, Immagine

Peter Fredette

Employment: IQVIA, EQRx

Travel, Accommodations, Expenses: IQVIA, EQRx

Lee Jones

Honoraria: Bayer, Eisai

Peggy McCann

Employment: Merck

Stock and Other Ownership Interests: Merck, Pfizer, CVS Health, GE Healthcare, GlaxoSmithKline, Bristol Meyers Squibb Company

Therica Miller

Consulting or Advisory Role: AOC Oncology

Other Relationship: University of Southern California, AOC Oncology, Alliance Foundation Trials, Children's Hospital Los Angeles, George Washington University, Alphasights

Fran Palmieri

Employment: Sarah Cannon Research Institute

Gary L. Smith

Stock and Other Ownership Interests: Mallinckrodt, Medtronic

Ajjai Alva

Consulting or Advisory Role: AstraZeneca, Merck, Pfizer, BMS

Research Funding: Genentech (Inst), Bristol Myers Squibb (Inst), Merck Sharp & Dohme (Inst), Prometheus (Inst), Mirati Therapeutics (Inst), AstraZeneca (Inst), Roche (Inst), Bayer (Inst), Astellas Pharma (Inst), Arcus Biosciences (Inst), Progenics (Inst), Celgene (Inst), Janssen (Inst)

Travel, Accommodations, Expenses: Merck, BMS

No other potential conflicts of interest were reported.





Complex Biomarkers: Aligning Best Practices to Support Future Utilization

A017 An analysis of 13 independently performed assays to measure homologous recombination deficiency using 90 freshly extracted high grade serous ovarian tumors: findings from the friends of cancer research hrd harmonization project.

Hillary Stires¹, Lisa M. McShane², Rebecca C. Arend³, Alyssa K. Chapman⁴, Li Chen⁴, Tommaso Coletta⁵, Yuan Ding⁶, Mohit Gupta⁷, Nikita Kotlov⁸, Alexander J. Lazar⁹, Ming-Chung Li², Yi-Hsuan Lucy Lai¹⁰, Wenjie Li¹¹, Brittany A. McKelvey¹, Jerod R. Parsons¹², Ethan S. Sokol¹³, Elizabeth R. Starks¹⁴, Mark D. Stewart¹, Peihua Wang¹⁵, Zhiwei Zhang², Yingdong Zhao², ShiPing Zou¹⁶, Jeff Allen¹.

¹Friends of Cancer Research, Washington, DC, ²National Cancer Institute, Rockville, MD, ³University of Alabama at Birmingham, Birmingham, AL, ⁴Frederick National Laboratory for Cancer Research, Frederick, MD, ⁵SOPHiA GENETICS, Rolle, Switzerland, ⁶Illumina Inc., San Diego, CA, ⁷Thermo Fisher Scientific, South San Francisco, CA, ⁸BostonGene, Corp, Waltham, MA, ⁹The University of Texas MD Anderson Cancer Center, Houston, TX, ¹⁰ACT Genomics Co., Ltd., Taipei, Taiwan, ¹¹Burning Rock Biotech, Guangzhou, China (People's Republic of China), ¹²Tempus Labs, Chicago, IL, ¹³Foundation Medicine, Cambridge, MA, ¹⁴Invitae, San Francisco, CA, ¹⁵Amoy Diagnostics Co., Ltd., Shanghai, China (People's Republic of China), ¹⁶Pillar Biosciences, Natick MA.

Homologous recombination deficiency (HRD) assays determine patient eligibility for treatment with PARP inhibitors and other DNA repair targeting drugs; therefore, understanding variability in how these assays measure and report HRD is critical for patients and providers. HRD assays measure various factors to determine HRD status including causes (i.e., inactivation in HR pathway genes) and consequences (i.e., genomic scarring). Methodological variability across HRD assays has led to a suggestion that the assays may not agree on a per patient basis. An empirical assessment of assay variability may guide our understanding of how to implement “HRD status” as a biomarker. Friends of Cancer Research (*Friends*) initiated a unique research partnership to assess levels of agreement and variability across HRD assays. We previously presented an analysis of the HRD status of 348 TCGA ovarian cancer samples across 11 assays. Concordance across assays was analyzed by measuring all possible pairings of samples and assays leading to a median (and IQR) positive percent agreement (PPA) of 74% (51-89%) and negative percent agreement (NPA) of 81% (64-92%). The median percent positivity (percent of patients with HRD) was 49% (range 9-67%). However, some groups modified their HRD pipelines to analyze the *in-silico* data and we were unable to interrogate the influence of patient and sample characteristics on HRD calls. To establish a more comprehensive dataset, we identified 142 archival specimens from the University of Alabama at Birmingham from patients with stage III-IV high grade serous ovarian cancer diagnosed between 2011 and 2022. Full clinical information is available including response to platinum therapy. FFPE tumor from debulking surgery was sectioned for the 99 samples with adequate tissue and underwent DNA and RNA extraction by the Molecular Characterization Laboratory (MoCha) at the NCI Frederick National Laboratory. MoCha shipped identical aliquots of DNA and/or RNA from the 90 samples that passed QC for independent sequencing and HRD measurement by the 13 different assays. Statisticians from the NCI performed statistical analyses to assess concordance across assays. Three of the 13 assays considered mutations in non-*BRCA1/2* HR pathway genes and 7 measured gLOH as determinants for HRD status among other factors. HRD calls resulted in a median pairwise PPA of 81% (69-91%) and a median pairwise NPA of 74% (61-89%). The median percent positivity was 53% (range 23-74%). Ongoing analyses will consider how each of the HRD assays predict responsiveness to platinum-based chemotherapy. Additional analyses will consider assay, patient, and sample characteristics that may drive variation in HRD calls. Preliminary findings demonstrate variability in HRD calls across HRD assays, similar to the *in-silico* analysis. These findings will help characterize the variability of HRD assays and inform best practices for measuring and reporting HRD.

Scan to view poster

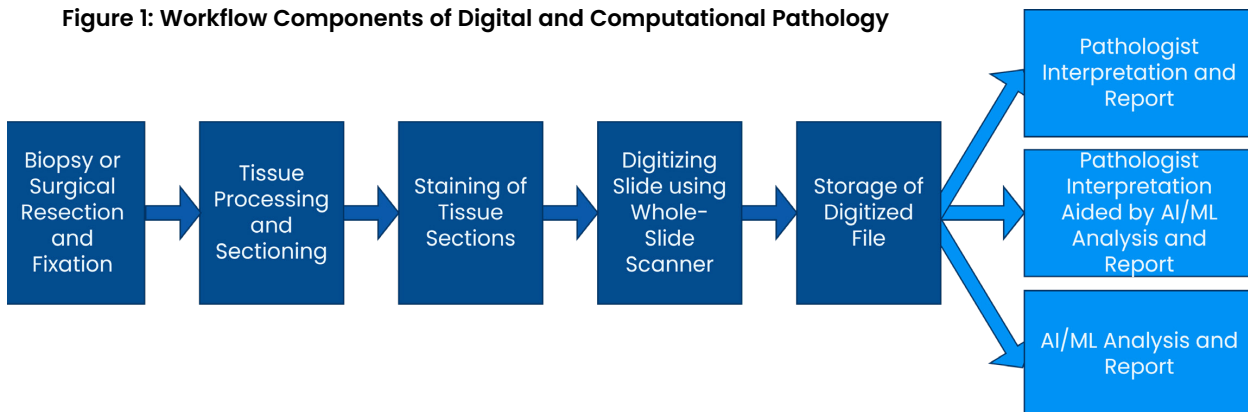


Supporting the Application of Computational Pathology in Oncology

Introduction

Biological heterogeneity of cancers causes tumors to respond differently to the same treatments. Thus, there is a compelling need to appropriately diagnose patients and identify relevant biomarkers for oncology treatments in both clinical practice and trials. Digital pathology is an emerging application in oncology drug development and clinical care, which allows for whole-slide image creation for storage, viewing, analyses, and interpretation. Digitized images are used directly by pathologists for biomarker interpretation, cellular annotation, and diagnosis. These images can also be used to support development of computational pathology platforms that utilize techniques such as artificial intelligence (AI) and machine learning (ML) to analyze and measure specific image elements, such as subvisual morphological patterns and phenotypes, identify features, and generate reproducible and structured data. These AI and ML platforms referred to in aggregate as computational pathology, may establish novel biomarkers, aid in quantifying prognostic and predictive biomarkers currently assessed or categorized by a pathologist, and expedite diagnosis or pathological scoring, all of which may go towards identifying and selecting patients for oncology treatments. Digital and computational pathology encompass several linked workflow components including both the digitization of the whole slides as well as the platforms for analysis (Figure 1).

Figure 1: Workflow Components of Digital and Computational Pathology



Thank You to Our Contributors

Nick Anderson, PathAI

J. Carl Barrett, formerly AstraZeneca, University of North Carolina at Chapel Hill

Anne-Laure Bauchet, Sanofi

Vipul Baxi, Bristol Myers Squibb

Nike Beaubier, Tempus Labs, Inc.

Daniel N. Cohen, GlaxoSmithKline

Suzana Couto, Neomorph, Inc.

Tathagata Dasgupta, 4D Path Inc.

Megan Doyle, Amgen

Dorothy French, Amgen

Brandon D. Gallas, U.S. FDA Center for Devices and Radiological Health

Emre Gulturk, Paige AI

Mark Gustavson, AstraZeneca

Sarah Hersey, Bristol Myers Squibb

Kimary Kulig, Kulig Consulting and My Biomarker Navigator

Alexander Lazar, MD Anderson Cancer Center

Jochen K. Lennerz, Massachusetts General Hospital

Mike Montalto, PathAI

Thomas Mrowiec, The Healthcare business of Merck KGaA, Darmstadt, Germany operates as EMD Serono in the U.S. and Canada

Satabhisa Mukhopadhyay, 4D Path Inc.

E. Tom Richardson, Merck & Co., Inc.

Roberto Salgado, University Hospital of Antwerp

Alain Silk, Tempus Labs, Inc.

Nino Sireci, Loxo@Lilly

Martin Stumpe, Tempus Labs, Inc.

Jian Wang, AstraZeneca

This paper reflects discussions that occurred among stakeholder groups, including the U.S. FDA. The topics covered in the paper, including recommendations, therefore, are intended to capture key discussion points. The recommendations provided should not be used in lieu of FDA published guidance or direct conversations with the Agency about a specific development program. This paper should not be construed to represent FDA's views or policies.

Objectives

Computational pathology has the potential to generate novel insights and biomarkers, and provide greater accuracy, reproducibility, and standardization of pathology-based features to aid in oncology drug development. Friends of Cancer Research (*Friends*) assembled a multi-stakeholder group of experts including government officials, computational pathology platform developers, academic pathologists and researchers, and biopharmaceutical industry members to outline proposals that facilitate robust development of computational pathology platforms for oncology drug development. The objectives of this group were to:

- Characterize current and future uses of computational pathology in oncology drug development and how they can facilitate clinical research.
- Identify the challenges in current drug and diagnostic co-development and articulate lessons learned to circumvent these in computational pathology.
- Provide proposals to facilitate robust development of computational pathology platforms for oncology drug development, including:
 1. Outline input and platform performance characteristics to report for optimized transparency.
 2. Establish a risk classification framework to inform evidentiary needs and performance criteria.
 3. Establish common reference standards and repositories of reference materials to support future platform development and cross-validation of platforms.

Uses of Digital and Computational Pathology in Oncology Drug Development

Digital pathology currently aids oncology drug development in operational and logistical tasks by supporting remote sharing of slides, storage of data for future analyses, and promoting efficient training of pathologists (Table 1). However, this white paper will focus on the use of AI/ML and other (image-based) computational pathology methods into a digital pathology workflow. Computational pathology can identify and quantify features from image data beyond human analytic capability. As such, computational pathology can establish novel biomarkers and improve current assessment of pathological features that would not otherwise be produced through conventional pathological evaluation. While this white paper focuses on the use of computational pathology in oncology, there is promise in other applications such as in non-alcoholic steatohepatitis (NASH)¹, inflammatory bowel disease (IBD)², and other diseases, and the proposals described herein may be relevant to these other applications.

Computational Pathology Applications in Oncology Drug Development

There is a spectrum of applications for digital and computational pathology throughout oncology drug development, including early discovery, pre-clinical and translational research, early phase trials, registrational trials, post-market/clinical use (Table 1). While some applications are currently in use in oncology drug development (e.g., digitization of tumor slides for future biomarker correlation to outcomes), others are currently in various stages of development (e.g., prediction of biomarker status) or are not yet ready for trials or clinical use (e.g., exploratory endpoints). Further, while each phase of development is depicted as distinct, the long-term goal for an integrated computational pathology workflow should be considered as it will determine the types of evidence and validation necessary for the platform. For example, a computational pathology platform used in exploratory translational research or early phase trials may not be intended for use in later phase trials or clinical care, while the goal for a platform used in a late phase trial may be to develop a companion diagnostic (CDx) for use in the post-market setting. Therefore, as some platforms may be used in several phases of drug development, developers should consider the various validation needs of these uses early in the platform development process.

Table 1. Examples of Potential Uses of Digital Pathology Workflows in Oncology Drug Development

Drug Development Phase					
Use of Digital Pathology	Pre-Clinical/Research	Exploratory/Translational	Early Phase Prospective Trials (I/IIa)	Late Phase Prospective Trials (IIb/III)	Post-Launch/Clinical Use
Digital Pathology Workflow	<ul style="list-style-type: none"> Reading/interpretation of pharmacology models and toxicology studies from digital images Peer review of toxicology digital slides from non-GLP or GLP studies Qualitative assessment of multiplex biomarker assays Evaluation of drug distribution (PD) 	<ul style="list-style-type: none"> Pathologist manual annotation and semi-quantitative scoring of images for biomarker prevalence, discovery and validation Exploration of manual/human interpretable features as biomarkers (i.e., mitosis count, IHC scores, % immune cell infiltrate) 	<ul style="list-style-type: none"> Archiving slides and retrospective qualitative or semi-quantitative biomarker analysis Review of biomarker status for trial enrollment Pharmacodynamic biomarker measurements using digital images Promote pathologist peer review, document decision-making process, and improve quality 	<ul style="list-style-type: none"> Visual review for trial enrollment: biomarker measurement as inclusion criteria Foster pathologist peer review, document decision process, and improve quality Measure clinical trial endpoint (e.g., pCR) using digital images 	<ul style="list-style-type: none"> Pathologist remote disease diagnosis, second opinion consult, tumor boards, pathologist education/board certification Store images for future analysis Pre-screen for selection of treatment Promote efficient training of pathologists
Computational Pathology Analyses Based on Digital Pathology Workflow	<ul style="list-style-type: none"> Toxicology/ veterinary pathology read/ count assistance on digital images Investigation of novel biomarkers, spatial characterization Evaluation and quantitation of drug distribution (PD) Quantitative assessment and interpretation of multiplexed biomarkers 	<ul style="list-style-type: none"> Retrospective analysis of clinical trial data to discover new biomarker spatial correlations or image features with prognostic or predictive value Enables quantification of histologic feature to create a unique biomarker(s) Data driven biomarker scoring 	<ul style="list-style-type: none"> PD biomarker quantification Tool for clinical trial enrollment Exploration of predictive biomarkers with more precise and continuous cutoffs Support, guide, and monitor pathologist scoring 	<ul style="list-style-type: none"> Trial enrollment: Biomarker measurement as inclusion criteria Evaluating exploratory endpoints Clinical trial outcomes assessment (e.g., pCR, RCB) Discovery of biomarkers in TME that may correlate to efficacy and/or safety Support, guide, and monitor pathologist scoring 	<ul style="list-style-type: none"> Biomarker assessment for targeted treatment identification Pre-screen followed by confirmatory testing for treatment selection Support, guide, and monitor pathologist scoring

GLP: Good Laboratory Practice, IHC: immunohistochemistry, PD: pharmacodynamics, pCR: pathologic complete response, RCB: residual cancer burden, TME: tumor microenvironment.

Platform Description and Use

For each computational pathology application, it will be useful to have a clear description of what it does and how it will be used, including the level of reliance on the output. This description will impact the evidentiary needs for validation. Some platforms may improve existing manual processes and assist the pathologist by enhancing or providing efficiencies (e.g., image quality control and low-level tasks like object or feature recognition, counting, and segmentation). Results generated from platforms that assist the pathologist in routine tasks or workflow support rely on the pathologist's final judgment and "sign-off."

However, computational pathology platforms are likely to provide novel insights that go beyond traditional histopathology assessments of pathologists, such as novel quantitative biomarker discovery or detection of spatial relationships between multiple biomarkers. These platforms may be further divided into those that produce an output that can be independently validated by a pathologist or other orthogonal method (e.g., DNA/ RNA sequencing) and those with an output that cannot be independently generated by a pathologist or other mechanism (i.e., "black box"). The ability to verify a platform's output by an alternate method may impact the level of evidence necessary to support its use. For example, in a clinical setting, a platform used as a pre-screen for a biomarker followed by confirmatory testing with a gold standard methodology (e.g., sequencing) may have different evidentiary needs for validation than if the output is the sole determinant for a patient receiving treatment.

Challenges in the Current Diagnostic and Drug Development Landscape

Currently, oncology diagnostic development for a predictive biomarker generally follows the paradigm where a single test or assay defines a single biomarker for a specific drug in a drug-diagnostic co-development model. This paradigm usually results in the U.S. Food and Drug Administration (FDA) approval of a CDx, which provides information that is essential for the safe and effective use of the corresponding drug or biological product.³ However, in clinical practice, additional assays, including laboratory developed tests, are often independently developed for the same biomarker and may be used in lieu of the approved CDx. As a result, a diverse set of assays with varying performance and predictive ability will be in use to detect the same biomarker to assist with treatment selection. Without robust data about performance and comparability across assays, this may result in confusion and lack of confidence in the diagnostics. This concern is reflected in FDA's recently released final guidance: Oncology Drug Products Used with Certain In Vitro Diagnostic Tests: Pilot Program.⁴ The pilot aims to increase transparency regarding performance characteristics for tests used to identify biomarkers for selection of oncology drug products.

Previous biomarker alignment and concordance demonstration projects on programmed death ligand 1 (PD-L1) immunohistochemistry (IHC)^{5,6} and tumor mutational burden (TMB),⁷ highlight disparate methodologies in biomarker assessment across available assays, with various clinical cutoffs used for reporting results and supporting treatment decision-making, possibly leading to disparate care for patients. The discordance seen in these projects provides lessons learned for improved prospective harmonization and transparency in the pre-market stage for computational pathology.

Disparate platforms and methodologies for biomarker assessment may make comparing computational pathology platforms challenging unless harmonization efforts exist. Currently,

there is not a simple mechanism for comparing the performance of the multiple available computational pathology platforms assessing the same biomarker. However, addressing this gap could support broader clinical use of computational pathology derived biomarkers, in addition to supporting broader regulatory authorizations outside of the single-platform, single-drug paradigm. Outlining best practices for validation studies, identifying and reporting key input and platform performance characteristics, and establishing standards to support the consistent performance of different computational pathology platforms can address concerns around test accuracy, reliability, and comparability.

Proposals for Robust Use of Computational Pathology in Drug Development

The following proposals for computational pathology development and use in oncology drug development will help to ensure the development of robust and well characterized platforms while enabling innovation.

Proposal 1: Input and Platform Performance Characteristics Reported for Optimized Transparency

Transparent methodology, input requirements, output scale and units, and performance characteristics will aid drug developers in identifying platforms that are appropriate for a given use case and aid platform developers and regulatory agencies in validating and evaluating robustness of platforms.

To increase transparency of the platform's methodology, the design and testing of the algorithm should be described, as well as the types of data used as training and validation sets, how the datasets were used, and how the datasets are related to the distribution of outputs. This information can support critical evaluation of the algorithm development and validation process, ensuring that datasets capture real-world parameters and are representative of the heterogeneity of treatment settings, patients, and tumor characteristics. Transparency in the baseline performance characteristics of a computational pathology platform for specific use cases can also help harmonize future development efforts resulting in high quality performance irrespective of the platform and developer.

Input Parameters to Consider in Development and Reporting

Given the multiple workflow components involved in computational pathology (Figure 1), it is important to clearly state and define the multiple input parameters that can influence the platform's robustness and performance. Defining the input parameters encourages more robust and transparent platform development and use and can be used to develop quality metrics, which can be applied across platforms. In turn, this can aid in the development of pathology practice standards to ensure consistent practice irrespective of where tissue is collected and processed, scanning devices used, and what platform is used. The two relevant categories of input parameters to define for computational pathology platforms are tissue processing (slide preparation) and image acquisition (scanning). Within these categories, key parameters to consider when evaluating input quality and the robustness of a platform for a given intended use are listed in Table 2 and are informed by FDA guidance on the technical performance assessment of digital pathology whole slide imaging devices.⁸ Each input parameter can be described or measured, and the appropriate specifications and quality metrics required

will depend on the platform’s application. Certain input parameters may be easier to control for quality (e.g., slide age) than others (e.g., tissue artifacts) and future work is needed to define quality metrics. The input parameters described in Table 2 are intended to help computational pathology developers directly by describing the specifications of their platform with regards to variation in the input parameters. This can also help drug developers understand and evaluate the capabilities and limitations of algorithms when considering their potential use in supporting drug development.

Table 2. Input Parameters to Define and Evaluate*

Parameter	Definition	Considerations
Tissue and Slide Processing		
Means of Tissue Acquisition	The type of tissue sample (e.g., excisional, core needle, fine needle aspirate, cytology, etc.)	Relevant to sampling bias and potentially algorithm performance; Relevant to valid use of platform per sample type
Tissue Sample Origin	The origin of the tissue sample (e.g., primary tumor vs. metastatic lesion vs. lymph node) as well as organ site	Relevant to valid use of platform per tissue origin
Tissue Processing	Specific steps for processing tissue (e.g., freezing, type of fixative, fixation time, etc.)	May impact tissue quality or usability with the platform; Some artifacts are specific to tissue processing and may affect the quality of the stain applied; Platforms may work differently on fresh frozen vs. FFPE tissue, etc.
Glass Slide Type	A description of the slide including thickness and slide material	Slide type may impact coloring and depth of the tissue that is measured and may affect opacity
Tissue Thickness	Acceptable range of tissue thickness in microns	May affect image quality and characteristics such as color and optic density of features as well as the number of cells analyzed
Tissue Area	Minimum and maximum tissue area recommended for reliable and reproducible analysis, including tissue area alone as well as tumor content (as a percentage of total area)	There may be a minimum amount of tumor tissue/tumor cells required for the analysis
Tissue Folds/Tears	Description of any tissue folds or tears in the tissue, and how these are handled	Presence of tissue folds/tears which may cause out-of-focus digitization, in addition to the reduced usability of areas that are directly affected
Surgical Ink/Pigments	Presence of surgical ink or other markings, and how these are handled	Markings may impact the software, and may result in false counts and misidentification of features
Other Tissue Artifacts	Other relevant artifacts (e.g., tissue lifting, incomplete decalcification, dust or surgical glove powder, bubbles, over fixation, improper dehydration, tissue bloating, etc. ⁹) and how they are handled	Various artifacts may impact analyses when present
Tissue Age	The recommended duration between when slides are cut and stained	May impact the stability of some features and affect stain characteristics such as intensity and color
Slide Age	The recommended duration between tissue staining and scanning	The time post-staining may impact intensity and quality of the slide (e.g., chromogen stability, diffusion of chromogenic dyes, fading of fluorescent dyes, etc.)

*Concepts in this table may be specific to currently existing technologies (e.g., IHC). As emerging technologies evolve (e.g., multiplex immunofluorescence, RNA mass spectrometry, etc.) the input parameters may also evolve depending on the technology.

Table 2. Input Parameters to Define and Evaluate* (con't)

Parameter	Definition	Considerations
Antibody Used	The antibody used for staining including clone, company, and catalog number	The type, batch, and age of antibody used may impact staining results
Staining Conditions	The staining conditions, such as incubation, blocking, etc.	Staining conditions may alter the staining intensity and results ¹⁰
Slide Storage	The manner and environment in which the physical slides are stored	Storage conditions (e.g., oxygen, humidity, sun or heat exposure) may impact staining results and/or tissue
Image Acquisition		
Scanner Hardware and Software Versions	Description of scanner hardware and software versions	Differences in hardware (e.g., optics) as well as software (e.g., pre/post-processing, color normalization, or application programming interface) can impact the algorithm performance
Scanner Software Configurable Parameters	Description of configurable parameters in the scanner software and the actual values, or acceptable ranges, which should be used during the scanning operation	Differences in scanner software configurable parameters (e.g., exposure and saturation) can impact the algorithm performance
Slide Viewer Used	Software and version used for slide viewing	Relevant to ability to use platform with different slide viewers and screens
Type of Image Files	Description of acceptable file formats and compression, and use of single plane images or image stacks	Relevant to whether image files can be appropriately processed by the algorithm
Region of Interest Selection	Information on whether the whole tissue, whole tumor area, or specific fields of view (including size) are used by the algorithm	The type of region may affect how algorithms are trained and their applicability to different tissue types
Magnification	The acceptable range of magnification of the digitized slide	Relevant to the use of the platform at different magnifications
Resolution	Specified magnification for image acquisition (e.g., 100x, 200x, 400x) and any requirements related to pixel resolution (expressed as micrometers per pixel)	Algorithms may require specific magnification during image acquisition and specific pixel density/resolution to identify features
Color	Details of the color processing, such as white-balance or contrast settings, which result in hue, saturation, brightness of the image; metrics for acceptable color settings and characteristics should be reported (with ranges of acceptability or a description of the color normalization procedure if used)	Algorithms can be sensitive to variations in color and contrast
Focus Quality	The focus quality required by the algorithm and a metric for acceptable focus quality	Focus quality can impact algorithms and should be quantified globally or locally as appropriate

Further, specifying performance/operating boundaries for the preprocess components of the workflow (e.g., scanners) will support use within the validated workflow. An appropriate description of the performance/operating boundaries may enable evaluation of the extent and conditions under which two different platforms used for the same purpose might produce similar results. The scanner model(s) and specific scanner configuration and acquisition protocol used for the training and testing of the computational pathology platform should be explicitly stated.

The specific parameters and acceptable ranges and values will depend on the computational pathology application. This includes the interaction of a human operator with the platform's output. For example, acceptable ranges may be wider when a human operator can independently check the output of the software or if it is being used to help direct a pathologist to examine certain slide areas, and narrower if the results cannot be independently verified by a human user.

Appendix 1 applies the reporting of input parameters to hypothetical use cases of computational pathology platforms. Some parameters, such as slide age, may be common across different use cases, whereas other input parameters may vary depending on the use case. Understanding the commonality or variability across use cases can also inform prioritization, by identifying parameters that may be relevant for model development and performance assessment for all studies.

Performance Characteristics and Assessment

Identifying and reporting key performance characteristics for computational pathology platforms will increase transparency, provide study designs and assessment methods for others to follow, and inform performance expectations for other quality and robust platforms. This may also increase confidence in using independently developed and validated platforms for a common purpose. Alignment is needed on standardized methods to report these characteristics to aid in transparency and the comparison.

Guidelines for establishing performance of AI or image analysis methods in computational pathology are limited. The FDA Center for Devices and Radiological Health (CDRH) has cleared one computational pathology device under a regulation that defines a broad intended use: "A software algorithm device to assist users in digital pathology [...] to provide information to the user about presence, location, and characteristics of areas of the image with clinical implications."¹¹ The special controls provided outline what information should be included in a Class II marketing submission for performance assessment, and the decision summary of the FDA-authorized device includes a summary of the scientific evidence that served as the basis for FDA's decision.¹² Other relevant FDA resources to understand key performance characteristics include regulations, reclassification orders, decision summaries, guidance documents, and other written works on the regulation of software as a medical device (SaMD) in areas other than pathology.¹³

The platform description, what it does and how it will be used, will impact its key performance characteristics. The College of American Pathologists published recommendations for the validation of whole slide imaging systems in clinical practice¹⁴ and further provides resources related to the validation of image analysis platforms in clinical practice.¹⁵ The Digital Pathology Association also broadly noted both hurdles and solutions for implementing computational

pathology and validating these platforms.¹⁶ Appendix 2 provides considerations for validation study designs, as well as examples of how the design elements were met in a few computational pathology validation studies.

Performance should be assessed on a dataset not used in the platform's development or training and is representative of the clinical population the platform is intended to evaluate to offer an unbiased assessment of performance.¹⁷ Performance characteristics may be influenced by details such as true biomarker prevalence in the study population, as well as training and testing data sources and sampling. These details and their impact on performance should be described. The input parameters highlighted in Table 2 will also impact performance and should be considered. Key performance characteristics should be evaluated in a manner consistent with what the platform does and how it will be used. This may include evaluation by standalone performance, a measure of the platform performance with little to no input or interpretation from the clinical end user, multi-reader multi-case study performance, and/or a measure of performance with interaction from the clinical end user or multiple end users. The end user involved in validation should be different than the user(s) involved in training.

Further, focusing on "explainable AI" (i.e., methods allowing for a representation of the input parameters used by the algorithm such as overlays of high attention areas or cell segmentation), may aid in the interpretability of "black box" algorithms. This interpretability could have two functions: allow for review of the impact of preanalytical variables, such as those detailed in Table 2, on the quality of the results, and bring additional confidence in the results to the end user.

Establishing Performance Comparisons

When performance comparisons to a "ground truth" or reference standard are possible and desirable, various study designs can be employed and careful consideration should be given to the method for establishing ground truth. Several methods exist for using pathologist interpretations as the reference standard, including using the original sign-out diagnosis, single readers, or consensus panels. Additionally, the concordance within pathologists should be considered when comparing concordance between a pathologist's interpretation and the platform's output, as there is also heterogeneity within pathologists' readings. Poor concordance within pathologists may indicate that multiple pathologists are needed to determine the reference standard. Also, the within-pathologist concordance may provide a performance criterion for model-pathologist concordance, assuming they are measured the same way.

In cases where comparison to a pathologist score/interpretation is not desired or possible, orthogonal methods that generate biological outputs such as gene or protein expression may be an acceptable comparison. For novel biomarkers, or in other cases where no orthogonal methods exist, native or contrived reference materials with a known or well-characterized status may be used as a comparison. Ultimately, establishing performance in relation to clinical characteristics or outcomes may be highly desirable, but is not always practical for certain use cases.

To compare the performance of several different computational platforms that report the same output, establishing a reference dataset with defined ground truth and pre-defined analysis methods is recommended. There is precedent for such approaches, such as the CAMELYON16

grand challenge,¹⁸ in which several model developers created models to detect lymph node metastatic disease and then tested the performance of their models on a single validation dataset.

Reporting Quantitative Measurements

Platforms may measure or define the biomarker of interest differently and direct cross comparisons may be challenging, especially with binary outputs. Although dichotomization of continuous biomarkers to a binary reading (e.g., high vs. low) by establishing a cutoff correlated to a clinical feature or outcome is frequently used in registrational trials for drug-CDx approval, the quantitative biomarker value (i.e., continuous scale) is often provided by computational pathology platforms and should be retained. Binary readings are often clinically desirable for ease of interpretation. How cutoffs are defined and derived should be encouraged. As part of the effort to establish an adequate cut off, there should be clear understanding of the variability in measurement surrounding the cutoff and reporting of the relevant range of quantitative measurements, their use within a final platform, and their relationship (if any) to outcomes in clinical trial data.

Proposal 2: Establish a Risk Classification Framework to Inform Evidentiary Needs and Performance Criteria

Adequate evidence generation, in the form of analytical and clinical validation, is needed to support the use of computational pathology platforms in oncology drug development. Further, a risk-based framework can support and inform this evidence generation and establishment of performance criteria across platforms and intended uses. Regulatory flexibilities are critical to encourage innovation and applying a risk-based approach will build an understanding of when flexibility is appropriate, what types of evidence are needed for computational pathology use in clinical trials and supporting regulatory approval, and regulatory pathways associated with a given platform.

Current Regulatory Classification and Pathways for Marketing

Regulatory agencies have applied existing risk classification systems for medical devices and diagnostics to digital pathology platforms. This paper focuses on the U.S. regulatory pathways, but depending on the intended use outside of the U.S., additional regulatory requirements should be considered in development (e.g., IVDR regulations). Diagnostic tests and digital pathology platforms are regulated based on their risk classification (i.e., Class I–III FDA designations), which helps inform the performance and reporting requirements.

Certain digital pathology platforms have been regulated as “Whole Slide Imaging” systems. In the U.S., these have largely been regulated as moderate-risk, class II devices requiring clearance of a 510(k) to be marketed.¹⁹ FDA issued recommendations regarding technical performance testing that should be completed to support a marketing submission for a whole slide imaging system.⁸ FDA has also regulated some AI/ML platforms as moderate-risk, class II devices, and issued special controls for these.¹² Further, the FDA, Health Canada, and United Kingdom’s Medicines and Healthcare products Regulatory Agency (MHRA) have put forth 10 guiding principles to inform Good Machine Learning Practices (GMLP) for medical devices using AI/ML, which could be applicable to computational pathology devices.²⁰ Further, FDA’s Drug Development Tool program²¹ and Medical

Device Development Tool program²² offers opportunities for public health stakeholders to pursue FDA qualification of digital and computational pathology tools.

Table 3 summarizes the current U.S. regulatory pathways and some applicable regulatory controls for specific defined use cases. Of note, this is not an exhaustive list of all regulatory controls that apply to developing and marketing a computational pathology platform. In addition to existing use cases and regulatory controls, a risk-based approach should be applied to future, not yet established use cases. Table 4 suggests example future use cases and a potential risk-based approach to regulating them. However, it is important to note there are currently no cleared or approved devices for these uses, and the FDA may not agree with the relationships between use cases and regulatory controls. Readers are encouraged to engage the FDA early and often, including with a Q-submission or a pre-IND to inquire about use cases and regulatory pathways.²³

Table 3: Potential Regulatory Pathways and Regulatory Controls for Marketing Digital Pathology Platforms by Intended Use²⁴		
Device Name, Risk Classification, Regulatory Pathway	Intended Use Summary	Potential Development and Evidence Generation Expectations
Software Algorithm Device To Assist Users In Digital Pathology ²⁵ Class II, 510(k)	Intended to aid a healthcare provider in determining a pathology diagnosis, provide information to the user about presence, location, and characteristics of areas of the image with clinical implications	Design Controls 21CFR820 Quality Management System ISO13485 Demonstrate substantial equivalence to a predicate Good ML Practices See Special Controls for Evidence Generation Expectations
Digital Pathology Image Viewing And Management Software ²⁶ Class II, 510(k)	Intended for viewing and management of digital images of scanned surgical pathology slides, as an aid to the pathologist to review and interpret these digital images for the purposes of primary diagnosis	Design Controls 21CFR820 Quality Management System ISO13485 Demonstrate substantial equivalence to a predicate Bench testing ⁸ Clinical Validation Study comparing to reference standard or manual read
Digital Pathology Display ²⁷ Class II, 510(k)	Intended for in vitro diagnostic use to display digital images of histopathology slides acquired by whole-slide imaging scanners that are used for review and interpretation by trained pathologists	Design Controls 21CFR820 Quality Management System ISO13485 Demonstrate substantial equivalence to a predicate Good ML Practices Bench testing ⁸ Display Equivalency Study
Whole Slide Imaging System ²⁸ Class II, 510(k)	Intended to aid the pathologist in review and interpretation of digital images of surgical pathology slides by automating digital slide creation, viewing, and management	Design Controls 21CFR820 Quality Management System ISO13485 Demonstrate substantial equivalence to a predicate Good ML Practices Bench testing ⁸ Clinical Validation Study Human factors study

Table 4: Example Future Use Cases with Potential Regulatory Pathways and Controls*

Potential Intended Use	Anticipated Risk Classification & Regulatory Pathway	Possible Regulatory Controls	Potential Development and Evidence Generation Expectations
Intended for use as a companion diagnostic	De Novo (Class II) or PMA (Class III) based on risk	Premarket Approval	Design Controls 21CFR820 Quality Management System ISO13485 Good ML Practices Demonstrate reasonable assurance of safety and effectiveness (AV, CV) Analytical validation studies (e.g., sensitivity, specificity, precision, accuracy, limit of detection, etc.) Clinical validation studies
Prescreening (with confirmation by another central test or CDx)	De Novo (Class I or II), due to lack of existing product code or classification regulation	Special Controls	Design Controls 21CFR820 Quality Management System ISO13485 Demonstrate substantial equivalence to a predicate Bench testing ⁸ Other data and controls, as requested by regulators, e.g.: Good ML practices Clinical Validation or Concordance Study
Automated computational digital pathology system for scanning, converting, reading, and detecting/measuring a biomarker on a pathology slide, with oversight and confirmation of output by a physician.	De Novo (Class I or II), due to lack of existing product code or classification regulation	Special Controls	Design Controls 21CFR820 Quality Management System ISO13485 Good ML Practices Demonstrate reasonable assurance of safety and effectiveness (AV, CV) Analytical validation studies (e.g., sensitivity, specificity, precision, accuracy, limit of detection, etc.) Clinical validation studies Usability study

*These are suggestions for a risk-based approach but have not been formally established via FDA classification decisions, clearances, or approvals to date.

Use of Computational Pathology Platforms in Clinical Trials

Currently, FDA guidance has not specifically addressed the use of computational pathology methods in clinical trials. Although others have published information and recommendations that may be helpful,²⁹ regulatory expectations for use of computational pathology platforms in oncology trials are still nascent. Recent publications have highlighted regulatory considerations for medical imaging AI/ML devices, including the existing regulatory pathways.^{30,31} Gaps in knowledge and test methods, and the novelty, pose a challenge for identifying regulatory expectations.³² To this end, this proposal seeks to build on prior regulatory resources by providing suggestions for a risk-based approach to these items, to advance the use of computational pathology platforms in oncology drug development.

While FDA has not opined specifically on the use of computational pathology in clinical trials, the agency has issued guidance on use of diagnostics and CDx in drug trials, as well as the use of

digital health technologies (DHTs) for remote data acquisition in clinical trials.^{33–35} Depending on the intended use, computational pathology platforms could be considered a diagnostic device as well as a type of DHT. Similar to when using a diagnostic device, or when using a DHT, trial sponsors should demonstrate that the platforms are fit-for-purpose (i.e., that the level of validation and performance characteristics are sufficient to support its use and interpretability) prior to use in the trial. Of note, evidence needed to demonstrate the platform is fit-for-purpose may not be commensurate with what would be expected to support regulatory authorization. Verification and validation would be expected, although the extent is not clearly defined. Additionally, there is an open question as to which quality and design principles to apply when developing a computational pathology platform for clinical trial use. With uncertainty in the regulatory pathway, the best course of action is to engage the FDA with a pre-IND submission in which one describes the computational pathology platform, the verification and validation results and plans, and how it will be used in the clinical trial.

Considerations that may be relevant to determining the level of evidence and design principles needed to demonstrate a computational pathology platform is fit-for-purpose could include:

1. The intended use of the platform;
2. Risk to patient safety;
3. Intent to support a marketing application for the platform or a drug; and
4. Business and trial operational risks.

For example, regarding intended use and risk to patients, computational pathology platforms used for pre-screening and confirmation with another medically established method, or to enrich for biomarker positive patient enrollment, may not require testing that is as robust as a platform used as the sole method for selecting participants for a trial or treatment arm, given these use cases pose less risk to patient safety. However, it is important to understand the concordance between the computational pathology platform and the confirmatory method, to avoid biases. Similarly, platforms used in an early phase study for biomarker discovery or exploration of disease biology likely require less stringent levels of validation and technical performance testing than a platform being used in a registrational trial where the data will inform patient management and support marketing authorization of the platform. Good software engineering practices and state-of-the-art software validation practices may be sufficient, from a quality and design perspective, for these lower risk use cases. Meanwhile, platforms being developed as a CDx and with an intent to market should be developed in accordance with design controls, AI/ML GMLPs, and would likely need to generate technical performance results, as well as robust evidence of analytical and clinical validity, among other data, to support a marketing submission. Further, it is imperative that the algorithm used in the clinical trial is predefined and locked in prior to use, including establishment of a cutoff.

The International Medical Device Regulators Forum (IMDRF) has published a SaMD risk categorization framework with four risk categories (I–IV) based on significance of the information to the healthcare decision (e.g. whether output from a SaMD is used to treat or diagnose, drive or inform clinical management) and the severity of the health condition.³⁶ Given the serious nature of cancer, using this risk categorization to inform evidence generation would be largely influenced by the intended use (e.g., inform management vs. treat or diagnose) and sponsors may find value

in applying this approach to the use of DHTs in clinical trials.

In addition to assessing the level of evidence needed to demonstrate a computational pathology platform is fit-for-purpose, sponsors should ensure compliance with other applicable regulatory requirements for the clinical trial. For example, for deployment into a clinical trial in the U.S., sponsors must follow 21 CFR part 812 to assess whether the platforms are considered to pose a significant risk to participants and/or seek an investigational device exemption (IDE) as needed.³⁷

In addition to patient safety risk, and unrelated to regulatory expectations, the operational risks to a clinical trial (e.g., logistics of incorporating new technology and costs) are also important considerations when determining the required level of performance testing of computational pathology platforms that will be used in a clinical trial. For example, a platform may present very little, if any, risk to patient safety, but may have an impact on important business drug development decisions such as a go/no go decision to proceed from an early safety/dose escalation trial to a registrational trial. Additionally, there are various operational models for implementing computational pathology in a clinical trial and commercial use, which may raise different risks for trial operations/business decisions. For example, implementation could use a centralized model (similar to central lab testing for a trial or a single-site PMA for a marketed diagnostic) or a distributed model (similar to a distributed IVD kit). Therefore, sponsors may want to assess the risks to trial operations/business decisions, when deciding whether the level of evidence is sufficient to use a computational pathology platform in a clinical trial.

Below are sample questions and considerations when determining fit-for-purpose performance testing of computational pathology platforms in oncology drug development. If the answers to these questions indicate a high risk to patient safety, then an organization should employ a high level of testing and quality oversight during development (e.g., strong engineering practices and/or design controls). Alternatively, if the answers to these questions suggest less impact to patient safety, then a less stringent level of performance testing or quality oversight may be acceptable.

Questions to Consider When Determining Fit-for-Purpose Performance Testing

1. How will the platform be used?
 - Will it be used prospectively to select patients for a trial or a treatment?
 - Will it be used retrospectively for biomarker discovery, disease biology, or other exploratory purposes?
 - Will it be used for assessment of a primary or secondary endpoint?
 - Will it be used for futility analyses or other analyses for decision-making on the trial?
 - Will it be used in conjunction with one or more confirmatory tests?
2. What is the risk to patients of an inaccurate result?
 - Will patient management change?
 - Could patients be exposed to treatment toxicities?
 - Will the dosing of patients be modified inappropriately?
 - Could a patient forgo the standard of care or be enrolled when little benefit is to be expected?
 - Could a patient be falsely excluded from receiving care with expected benefit?
3. Will the platform be the subject of a marketing authorization application?

- Will the platform be used to generate data in support of a marketing application for a drug?
 - Will the platform itself be the subject of a medical device marketing application?
 - Are both drug and device marketing applications intended?
4. What are the business risks of an inaccurate result?
- Will implementation of the platform be using a centralized model?
 - Will implementation of the platform be using a distributed model?

Proposal 3: Establish Common Reference Standards

Establishing common reference standards and repositories will support future platform development and cross-validation. As multiple platforms are developed for the same biomarker, utilizing common datasets to validate and develop these platforms can support 1) wider access to biomarker testing across multiple platforms showing similar performance characteristics that may already be in place in testing labs, 2) platform developers producing concordant or comparable platforms, and 3) clinician end-users making informed decisions because they will understand the comparability of different platforms. This may help prevent future situations such as that observed with the various PD-L1 follow-on tests, in which multiple PD-L1 IHC assays were independently developed as follow-ons for different therapies without an understanding of how these different assays and scoring methodologies were related.³⁸

While a single computational pathology platform may be used in a registrational trial for biomarker identification, additional, “follow-on” platforms measuring the same biomarker may be developed. Where available, the original slides could be used to ensure new platforms developed have high concordance with the originally approved platform, in addition to the other datasets used for validation of the follow-on platforms. However, institutional definitions of images as biospecimens versus de-identified data will impact the ease with which the images may be stored, shared, or used. Further, there are existing country-specific requirements and regulations regarding maintaining control of patient-level data that may impact the feasibility of sharing trial images. If the images cannot be shared, the platforms could be made available to the sponsor to evaluate performance across platforms using digital images from registrational studies, assessing the comparability of the performance of multiple platforms on its own dataset without sharing the slides. Although it would benefit drug developers to assess performance across platforms to identify a biomarker of interest, the scalability and management of such research is uncertain. The burden would be on drug developers to ensure proper consent for this future use and to conduct this work, as well as add potential regulatory or commercial risk to be involved with validation of third-party platforms outside of the CDx, which may limit the viability of this approach.

Unlike the banking of tissue and/or blood samples in which there is limited supply, banking slide images with proper informed consent for future use may be more attainable. However, criteria to define the appropriate number of images, or size of the training dataset will vary according to the platform being developed and the intent of use. Additionally, the storage, back-up, and auditing of the images are not negligible undertakings. The memory storage size and cost of databases needed to hold the images and associated metadata are substantial and should be considered when developing datasets. Furthermore, the workflow for digitization and interpretation of the

images involves many different people, roles, as well as potentially different locations (e.g., where the slide is cut, digitized, and image analysis conducted). Therefore, developing robust reference datasets must encompass the relevant stakeholders (e.g., sponsors, pathology labs, platform developers).

There are additional opportunities to develop reference datasets outside of a single sponsor in a pre-competitive manner. Commercially acquired digitized images, or those collected through a consortium, could provide access to images that could be analyzed using the same platform and algorithm deployed for the registrational trial of interest as a comparator and reference for other platforms. Consortia have previously used a commissioned third-party to securely hold and analyze data from drug and/or diagnostic developers and share results with the community. Alternatively, a federated model for a reference dataset could be implemented, with those in control of the images maintaining control over their critical datasets (either a sponsor or a source institution) but allowing a model to run on the images without the images themselves leaving the virtual workspace. This federated model would allow for concordance testing both between different datasets as well as different algorithms. Depending on the intended use of the reference dataset, linked outcomes data may not be necessary, which may increase the comfort level of sponsors to share data. Lastly, existing infrastructure may be leveraged to share digital pathology images, including the National Cancer Institute's (NCI) Imaging Data Commons,³⁹ a cloud-based repository of publicly available cancer imaging data, as well as the precisionFDA⁴⁰ platform, a secure, cloud-based environment permitting collaborative research and data sharing on a secure platform.

A common reference set of slides are needed to support generating robust data repositories. Recommendations for establishing a reference dataset (also see these references^{41,42}):

- Slides are digitized shortly after staining to minimize the impact of storage on the quality of slides. A timing threshold could be established and reported.
 - If slides are not digitized shortly after, such as when archived samples are imaged, detailed reporting of the slide age is needed.
- Images are stored appropriately and in the same file format to ensure the greatest amount of interoperability.
 - There are current initiatives to expand the DICOM standard to pathology imaging and could be one mechanism to enable alignment.
- Access to stored documents is secure and controlled, but not cumbersome.
- Relevant preprocess metadata including input parameters (Table 2) are linked to the images.
- Clinical metadata is ideally included, containing orthogonal information such as genomic and proteomic data, treatment regimens, and outcomes.
- Relevant characteristics of the intended patient population and measurement inputs are sufficiently represented in a sample of adequate size.
- All metadata are reported in a standardized format and of a given quality.
- Dataset represents the heterogeneity of real-world clinical/laboratory practices and patient populations, including slide preparation, scanning, patient characteristics, and tumor characteristics.

When platform developers leverage reference standards to perform comparisons and assess

performance, it is important to consider what the platform does and how it will be used, as well as the purpose of the reference dataset to ensure the intentions are aligned and the reference dataset has the appropriate data. This includes considerations on the types of tissue and slide processing, diseases, digitization methodology, and relevant metadata (Table 2). Reference datasets should also be diverse in the relevant patient and tumor characteristics, preferably from multiple centers to be more generalizable to real-world patient and clinical practice populations. It is imperative that reference datasets have data reported in a standardized format, including reporting the input parameters for digitization, patient and tumor characteristics, treatment and outcome data, and platform performance metrics and output. As noted in Proposal 2, computational pathology biomarker measurements should be reported as continuous variables in addition to binary results even if performance metrics dichotomize the data.

Conclusions

This white paper highlights the promise of computational pathology to aid oncology drug development, as well as the possible future challenges to evaluating the robustness of these platforms to support their validation and use in drug development. As such, the proposals outlined support identification and reporting of key input and platform performance characteristics, a framework to inform evidentiary needs and performance criteria, and opportunities for establishing standards and common reference datasets. Computational pathology can be used across the spectrum of oncology drug development, from early discovery to registrational trials, and the intended use for each computational pathology application will impact the evidentiary needs to validate the platform. Computational pathology is an evolving field with evolving technologies, and as such, the possible applications and validation of these platforms will grow.

In addition to this working group, there are many ongoing consortia and efforts surrounding the use of digital and computational pathology platforms and their validation, and collaboration is needed to tackle outstanding questions.^{43–45} Future efforts are needed to align on recommendations and benchmarks for quality metrics of preprocess input parameters to support transparency in platform development. Further, to support aligned data deposition into reference datasets, the development of standardized methodologies and data dictionaries is also needed. Alignment regarding data storage (e.g., on premises versus cloud solutions, ensuring data integrity and security, data transfer, redundancy/backups) is critical to ensure robust datasets for future use.

Formal guidance from regulatory bodies and relevant interest groups is needed to set regulatory expectations and establish performance metrics for computational pathology in drug development. FDA has signaled⁴⁶ their consideration of AI/ML in aiding drug development, with discussion ongoing. Clarity in the regulatory expectations for use of computational pathology in clinical trials would be valuable, including the evidence to demonstrate a platform is fit-for-purpose and the quality and design principles to apply when developing these platforms.

While this white paper demonstrates the potential promise of use of these platforms, there are currently regional differences in capabilities for using this technology. Many laboratories do not have digitization capabilities, due to lack of infrastructure, training, adequate funding, or other barriers. Additionally, if digitization capabilities are available, most have only access to one scanner type, which may impact the ability to use various platforms if they are not developed

in an agnostic way to the digitization workflow. Significant uptake of robust digital pathology is needed to realize the promise of these platforms and future work should address these barriers to enable broader uptake.

Lastly, there is an opportunity to leverage existing data (e.g., pathology slides, metadata) from various stakeholders to generate an accessible digital pathology dataset to cross-evaluate different computational pathology platforms measuring the same biomarker to support the concepts in this white paper. There is a precedent in the AI development industry to conduct “Challenges” to evaluate the variability of AI models using standard datasets for training and testing, and precisionFDA also hosts challenges.⁴⁷ Further, *Friends* has conducted previous harmonization efforts to support aligned biomarker measurement and use, including the Tumor Mutational Burden (TMB) Harmonization and Homologous Recombination Deficiency (HRD) Harmonization Projects, and is poised to support a harmonization effort in computational pathology. Future work will focus on building out an appropriate use case to test the proposals herein, clarify workflows, and provide concrete data to support guidance efforts.

References

1. Naoumov N V., Brees D, Loeffler J, et al. Digital pathology with artificial intelligence analyses provides greater insights into treatment-induced fibrosis regression in NASH. *J Hepatol.* 2022;77(5):1399-1409. doi:10.1016/J.JHEP.2022.06.018
2. Najdawi F, Sucipto K, Mistry P, et al. Artificial Intelligence Enables Quantitative Assessment of Ulcerative Colitis Histology. *Modern Pathology.* 2023;36(6):100124. doi:10.1016/j.modpat.2023.100124
3. Companion Diagnostics | FDA. Accessed August 30, 2023. <https://www.fda.gov/medical-devices/in-vitro-diagnostics/companion-diagnostics>
4. Oncology Drug Products Used with Certain In Vitro Diagnostics Pilot Program | FDA. Accessed August 30, 2023. <https://www.fda.gov/medical-devices/in-vitro-diagnostics/oncology-drug-products-used-certain-in-vitro-diagnostics-pilot-program>
5. Tsao MS, Kerr KM, Kockx M, et al. PD-L1 Immunohistochemistry Comparability Study in Real-Life Clinical Samples: Results of Blueprint Phase 2 Project. *Journal of Thoracic Oncology.* 2018;13(9):1302-1311. doi:10.1016/J.JTHO.2018.05.013
6. Rimm DL, Han G, Taube JM, et al. A Prospective, Multi-Institutional Assessment of Four Assays for PD-L1 Expression in NSCLC by Immunohistochemistry. *JAMA Oncol.* 2017;3(8):1051. doi:10.1001/JAMAONCOL.2017.0013
7. Vega DM, Yee LM, McShane LM, et al. Aligning tumor mutational burden (TMB) quantification across diagnostic platforms: phase II of the Friends of Cancer Research TMB Harmonization Project. *Annals of Oncology.* 2021;32(12):1626-1636. doi:10.1016/j.annonc.2021.09.016
8. Technical Performance Assessment of Digital Pathology Whole Slide Imaging Devices Guidance for Industry and Food and Drug Administration Staff Preface Public Comment. Published online 2016. Accessed August 30, 2023. <http://www.regulations.gov>
9. Taqi SA, Sami SA, Sami LB, Zaki SA. A review of artifacts in histopathology. *J Oral Maxillofac Pathol.* 2018;22(2):279. doi:10.4103/JOMFP.JOMFP_125_15
10. Bagchi A, Madaj Z, Engel KB, et al. Impact of Preanalytical Factors on the Measurement of Tumor Tissue Biomarkers Using Immunohistochemistry. *J Histochem Cytochem.* 2021;69(5):297-320. doi:10.1369/0022155421995600
11. CFR - Code of Federal Regulations Title 21. Accessed August 30, 2023. <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/cfrsearch.cfm?fr=864.3750>
12. Class II De Novo Paige Prostate FDA Letter. https://www.accessdata.fda.gov/cdrh_docs/pdf20/DEN200080.pdf.
13. Petrick NA, Chen W, Delfino JG, et al. Regulatory considerations for medical imaging AI/ML devices in the United States: concepts and challenges. <https://doi.org/10.1117/1.JMI.10.5.051804>. 2023;10(5):051804. doi:10.1117/1.JMI.10.5.051804
14. Evans AJ, Brown RW, Bui MM, et al. Validating Whole Slide Imaging Systems for Diagnostic Purposes in Pathology. *Arch Pathol Lab Med.* 2022;146(4):440-450. doi:10.5858/ARPA.2020-0723-CP
15. How to Validate AI Algorithms in Anatomic Pathology | College of American Pathologists. Accessed August 30, 2023. <https://www.cap.org/member-resources/clinical-informatics-resources/how-to-validate-ai-algorithms-in-anatomic-pathology>
16. Abels E, Pantanowitz L, Aeffner F, et al. Computational pathology definitions, best practices,

- and recommendations for regulatory guidance: a white paper from the Digital Pathology Association. *J Pathol.* 2019;249(3):286. doi:10.1002/PATH.5331
17. Moons KGM, Kengne AP, Grobbee DE, et al. Risk prediction models: II. External validation, model updating, and impact assessment. *Heart.* 2012;98(9):691-698. doi:10.1136/HEARTJNL-2011-301247
 18. Bejnordi BE, Veta M, Van Diest PJ, et al. Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. *JAMA.* 2017;318(22):2199-2210. doi:10.1001/JAMA.2017.14585
 19. Evaluation of Automatic Class III Designation- De Novo Request . https://www.accessdata.fda.gov/cdrh_docs/pdf16/DEN160056.pdf.
 20. Good Machine Learning Practice for Medical Device Development: Guiding Principles | FDA. Accessed August 30, 2023. <https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles>
 21. Drug Development Tool (DDT) Qualification Programs | FDA. Accessed September 4, 2023. <https://www.fda.gov/drugs/development-approval-process-drugs/drug-development-tool-ddt-qualification-programs>
 22. Medical Device Development Tools (MDDT) | FDA. Accessed September 4, 2023. <https://www.fda.gov/medical-devices/medical-device-development-tools-mddt>
 23. CDRH. Requests for Feedback and Meetings for Medical Device Submissions: The Q-Submission Program Guidance for Industry and Food and Drug Administration Staff Preface Public Comment. Accessed August 30, 2023. <https://www.reginfo.gov>.
 24. Lennerz JK, Green U, Williamson DFK, Mahmood F. A unifying force for the realization of medical AI. *npj Digital Medicine* 2022 5:1. 2022;5(1):1-3. doi:10.1038/s41746-022-00721-7
 25. Product Classification. Accessed August 30, 2023. <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpdc/classification.cfm?id=QPN>
 26. Product Classification. Accessed August 30, 2023. <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpdc/classification.cfm?ID=QKQ>
 27. Product Classification. Accessed August 30, 2023. <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpdc/classification.cfm?id=5328>
 28. Product Classification. Accessed August 30, 2023. <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpdc/classification.cfm?ID=PSY>
 29. Pell R, Oien K, Robinson M, et al. The use of digital pathology and image analysis in clinical trials. *J Pathol Clin Res.* 2019;5(2):81-90. doi:10.1002/CJP2.127
 30. Petrick N, Chen W, Delfino JG, et al. Regulatory considerations for medical imaging AI/ML devices in the United States: concepts and challenges. *J Med Imaging (Bellingham).* 2023;10(5). doi:10.1117/1.JMI.10.5.051804
 31. Chen W, Krainak D, Sahiner B, Petrick N. A Regulatory Science Perspective on Performance Assessment of Machine Learning Algorithms in Imaging. Published online 2023:705-752. doi:10.1007/978-1-0716-3195-9_23
 32. Digital Pathology Program: Research on Digital Pathology Medical Devices | FDA. Accessed August 30, 2023. <https://www.fda.gov/medical-devices/medical-device-regulatory-science-research-programs-conducted-osel/digital-pathology-program-research-digital-pathology-medical-devices>
 33. Principles for Codevelopment of an. Accessed August 30, 2023. <http://www.fda.gov/Drugs/>

GuidanceComplianceRegulatoryInformation/Guidances/de

34. Investigational IVDs Used in Clinical Investigations of Therapeutic Products Draft Guidance for Industry, Food and Drug Administration Staff, Sponsors, and Institutional Review Boards DRAFT GUIDANCE. Published online 2017. Accessed August 30, 2023. <http://www.fda.gov/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/G>
35. Digital Health Technologies for Remote Data Acquisition in Clinical Investigations. Accessed August 30, 2023. <https://www.fda.gov/vaccines-blood-biologics/guidance-compliance-regulatory-information-biologics/biologics-guidances>
36. Global Approach to Software as a Medical Device | FDA. Accessed August 30, 2023. <https://www.fda.gov/medical-devices/software-medical-device-samd/global-approach-software-medical-device>
37. CFR - Code of Federal Regulations Title 21. Accessed August 30, 2023. <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfCFR/CFRSearch.cfm?CFRPart=812>
38. Salgado R, Bellizzi AM, Rimm D, et al. How current assay approval policies are leading to unintended imprecision medicine. *Lancet Oncol*. 2020;21(11):1399-1401. doi:10.1016/S1470-2045(20)30592-1
39. Fedorov A, Longabaugh WJR, Pot D, et al. NCI imaging data commons. *Cancer Res*. 2021;81(16):4188. doi:10.1158/0008-5472.CAN-21-0950/674291/AM/NCI-IMAGING-DATA-COMMONS/NCI-IMAGING-DATA-COMMONS
40. precisionFDA - Overview. Accessed August 30, 2023. <https://precision.fda.gov/>
41. Homeyer A, Geißler C, Schwen LO, et al. Recommendations on compiling test datasets for evaluating artificial intelligence solutions in pathology. *Modern Pathology*. 2022;35(12):1759-1769. doi:10.1038/S41379-022-01147-Y
42. Wahab N, Miligy IM, Dodd K, et al. Semantic annotation for computational pathology: multidisciplinary experience and best practice recommendations. *J Pathol Clin Res*. 2022;8(2):116-128. doi:10.1002/CJP2.256
43. Digital and Computational Pathology Committee | College of American Pathologists. Accessed August 30, 2023. <https://www.cap.org/member-resources/councils-committees/digital-pathology-committee/>
44. Pathology Innovation Collaborative Community - MDIC. Accessed August 30, 2023. <https://mdic.org/program/picc/>
45. DPA: Digital Pathology Association. Accessed August 30, 2023. <https://digitalpathologyassociation.org/>
46. Cder Cber F. Using Artificial Intelligence & Machine Learning in the Development of Drug and Biological Products. Accessed August 30, 2023. <https://www.imdrf.org/documents/machine-learning-enabled-medical->
47. Challenges - precisionFDA. Accessed August 30, 2023. <https://precision.fda.gov/challenges>
48. Wolff AC, Somerfield MR, Dowsett M, et al. Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer: American Society of Clinical Oncology–College of American Pathologists Guideline Update. *Arch Pathol Lab Med*. Published online June 7, 2023. doi:10.5858/ARPA.2023-0950-SA/493531/HUMAN-EPIDERMAL-GROWTH-FACTOR-RECEPTOR-2-TESTING
49. Dawson H. Digital pathology – Rising to the challenge. *Front Med (Lausanne)*. 2022;9:888896. doi:10.3389/FMED.2022.888896/BIBTEX
50. Farahmand S, Fernandez AI, Ahmed FS, et al. Deep learning trained on hematoxylin and eosin

- tumor region of Interest predicts HER2 status and trastuzumab treatment response in HER2+ breast cancer. *Mod Pathol.* 2022;35(1):44-51. doi:10.1038/S41379-021-00911-W
51. Shamaï G, Livne A, Polónia A, et al. Deep learning-based image analysis predicts PD-L1 status from H&E-stained histopathology images in breast cancer. *Nat Commun.* 2022;13(1). doi:10.1038/S41467-022-34275-9
 52. Baxi V, Lee G, Duan C, et al. Association of artificial intelligence-powered and manual quantification of programmed death-ligand 1 (PD-L1) expression with outcomes in patients treated with nivolumab ± ipilimumab. *Mod Pathol.* 2022;35(11):1529-1539. doi:10.1038/S41379-022-01119-2
 53. Luchini C, Pantanowitz L, Adsay V, et al. Ki-67 assessment of pancreatic neuroendocrine neoplasms: Systematic review and meta-analysis of manual vs. digital pathology scoring. *Mod Pathol.* 2022;35(6):712-720. doi:10.1038/S41379-022-01055-1

Appendix I: Hypothetical Use Cases for Considering and Reporting Input Parameters

Below are examples of hypothetical use cases to aid in how one may consider what parameters/conditions should be evaluated and characterized for a specific use case:

- Mitosis counting
 - Counts the number of mitoses/mm² in a sample using an algorithmic method for identifying the region of interest (ROI) or allows pathologists to select the ROI to be analyzed.
 - Requires a minimum area of sufficient quality for analysis.
 - Can tolerate slides with large regions that are inadequate for analysis.
 - Prostate cancer Gleason grading
 - Algorithm to assign a Gleason score to prostate cancer samples.
 - Provides primary, secondary, and tertiary grades, and overall Gleason score by analyzing large scale histological patterns within a specimen.
 - Requires a minimum, representative total area of sufficient quality and with accurate location information for different prostatic regions; high magnification not required.
 - Less tolerant to slides with large regions that are inadequate for analysis or have artifacts.
 - Metastases detection
 - Algorithm that detects the presence of metastatic cells within a biopsy.
 - High sensitivity task requiring a minimum total area of high-quality tissue and images.
 - Intolerant of slides with large regions that are inadequate for analysis or have artifacts.
-

Appendix Table 1: Reporting Input Parameters for Hypothetical Use Cases.
This table highlights how certain input parameters listed in Table 2 may be common considerations across use cases or may be different dependent on use case.

Parameter	Mitosis Counting	Metastasis Detection for Diagnostic Aid	HER2 Status Prediction for CDx
Tissue Sample Origin	fine needle aspirate not appropriate, other biopsy types acceptable	no specific requirements regarding biopsy type	Breast biopsies, breast resections, or specimens from metastatic sites (if applicable)
Tissue Processing	standard FFPE preparation, standard H&E staining		No frozen tissue Cold ischemic and fixation times within range as stipulated by interpretive guidelines ⁴⁸
Slide Type	standard glass slide, 1mm thick		
Tissue Thickness	3–5um sections		
Tissue Area	2mm ² of tissue is analyzed, no tumor percentage minimum; region of interest based on algorithmic methods (details of method specified; i.e., random selection of X fields)	>1mm ² of sufficient quality tissue area required; no tumor percentage minimum; entire tissue area evaluated	>1mm ² of sufficient quality tissue area required and 20% minimum tumor content
Tissue Folds/ Tears	no tissue folds or tears in any analyzed region; algorithmic selection of regions of interest will prevent the presence of tissues folds/tears in analyzed regions	folds or tears and adjacent 5um distance will be excluded from analysis, excluded area must be less than 10% total analysis area	According to specific manufacturer QC procedure developed for intended use of the CDx, including detailed methodology and criteria to identify and exclude slides with tissue folds/tears.
Slide Age	scanned within X years/months of staining		
Slide Storage	slides protected from light, stored at room temperature		
Magnification	400X	200X or 400X	400x only

Appendix 2: Examples of Computational Pathology Platform Validation Study Designs
 Below are examples of study designs to validate digital pathology platforms, provided based on the considerations for validation study design put forth in *Frontiers in Medicine*.⁴⁹

	Ground Truth Definition	Case Selection	Acceptable Range of Output Values	Possible Confounding Effects	Identify Discrepant Cases
Her2⁵⁰	Algorithm compared to a ground truth to establish precision and recall	Case mix should reflect real-world setting in terms of morphological heterogeneity and complexity	Define acceptable range of deviation from the ground truth. This may depend on clinically relevant cutoffs that determine therapy	Consider any variables in image preparation. For example, compare effect of different scanners.	Output values outside the defined acceptable range are discrepant to the ground truth (can systematic reasons be identified).
	Herceptest IHC scored based on 2018 ASCO/CAP guidelines: Intense circumferential 3+ membrane staining in > 10% neoplastic cells are positive. The ground truth for the IHC results were defined as the consensus score reached by 3 pathologists for each case.	Trained a HER2 status predictor model on 188 HER2± H&E slides (93+/95-) and a test set of 187 HER2± H&E slides from The Cancer Genomic Atlas (TCGA) BRCA cohort.	The fully trained CNN model performance predicted the HER2 status with slide-level AUC of 0.90 (95% CI, 0.79–0.97). Model validation with an independent test set achieved an AUC of 0.80 (95% CI: 0.69–0.88) at the slide-level. Algorithm prediction of Trastuzumab clinical response is weak (sensitivity .56, specificity .58).	Utilized a deep learning based color normalization to remove batch effects and improve generalizability to independent datasets.	Borderline case confusion minimized by using only IHC 3+ cases training. Pathologist annotation improved model prediction.
PD-L1⁵¹	Breast Cancer PD-L1 expression was determined using the Ventana PD-L1 (SP142) assay as the proportion of tumor area occupied by PD-L1 expressing tumor-infiltrating immune cells (IC), and an expression in ≥1% IC was defined as PD-L1 positive status.	Tissue Microarrays: training set 2,516 (74.5%) cases; test set 860 (25.5%); and external test set 275.	CNN algorithm AUC performance with respect to the pathologist's binary PD-L1 status was 0.911 (95% CI: 0.891–0.925). Test set AUC performance was 0.915 (95% CI: 0.883–0.937). Independent test set AUC performance for PD-L1 prediction was 0.854 (95% CI: 0.771–0.908).	Data augmentation was performed to help the model deal with variability in staining methods and other differences between the cohorts.	Ground truth was not perfect; therefore, confirmed scores between 3 pathologists. Tissue microarrays used to train and test the algorithm may not have sufficient representation.

Appendix 2: Examples of Computational Pathology Platform Validation Study Designs
 Below are examples of study designs to validate digital pathology platforms, provided based on the considerations for validation study design put forth in *Frontiers in Medicine*.⁴⁹

	Ground Truth Definition	Case Selection	Acceptable Range of Output Values	Possible Confounding Effects	Identify Discrepant Cases
PD-L1⁵²	Algorithm compared to a ground truth to establish precision and recall Dako PD-L1 28-8 IHC with cutoffs of TPS $\geq 1\%$ and $\geq 5\%$.	Case mix should reflect real-world setting in terms of morphological heterogeneity and complexity 217 samples from patients with NSCLC, 600 from MEL, 400 from SCCHN, and 293 from patients with UC.	Define acceptable range of deviation from the ground truth. This may depend on clinically relevant cutoffs that determine therapy AI-based assessment was highly correlated with the median score from manual assessing of PD-L1-expressing TCs by 5 pathologists (r ranging from 0.73 to 0.85).	Consider any variables in image preparation. For example, compare effect of different scanners. Slides were scanned by two separate Aperio AT2 scanners across 5 days, two times per day (morning [AM] and afternoon [PM]).	Output values outside the defined acceptable range are discrepant to the ground truth (can systematic reasons be identified). A lower prevalence of PD-L1-positive patients was seen with AI-powered scoring (42.5% and 28.8%) compared with manual scoring (54.9% and 34.0%) at cutoffs of $\geq 1\%$ and $\geq 5\%$, respectively, though the difference was not significant. This could be due to the presence of artifacts or low PD-L1 membrane staining with cytoplasmic positivity (blush).
Ki67⁵³	Current guidelines for assessing Ki-67 recommended manual counting from a printed image that includes at least 500 neoplastic cells from tumor hotspots.	Review including 752 Pancreatic Neuroendocrine Neoplasms: G1 (55.3%), G2 (40.6%) and G3 (4.1%).	The pooled correlation estimate was 0.94 (95%CI: 0.83–0.98; I2 = 24.15%).	Risk of counting dividing non-neoplastic “contaminating” cells (endothelial cells, lymphocytes) and other brown pigment (hemosiderin).	Higher tumor grade generated due to overcounting “contaminating” cells or artifact.

Establishing Evidence: New Advancements Using ctDNA

A critical component of oncology clinical trials is evaluating the efficacy of new therapies and identifying which patients respond to therapy. A variety of endpoints are leveraged for measuring treatment efficacy, such as overall survival and progression-free survival. As the magnitude of benefit continues to improve with the advent of new therapies, clinical trials may take longer to assess efficacy based on currently available endpoints. Early endpoints that are reasonably likely to predict clinical benefit, such as response rate based on radiographic imaging, are used to evaluate treatment efficacy earlier than measuring overall survival. There is a need to identify, evaluate, and validate additional novel endpoints to assess efficacy earlier in the course of treatment that are predictive of long-term outcomes.

Circulating tumor DNA (ctDNA) is an emerging, new biomarker that can identify patients who respond to therapies by evaluating the presence and levels of ctDNA in a simple blood draw. Because of emerging data and growing excitement in the field, the U.S. Food and Drug Administration (FDA) released a draft guidance document that highlights the potential use of ctDNA as an early endpoint and emphasizes where additional evidence is needed for validation.¹

Project Overview

Recognizing the potential value of ctDNA as a novel endpoint in oncology drug development and the need for collaboration, Friends of Cancer Research (*Friends*) launched a unique multi-stakeholder research partnership to generate evidence and determine whether changes in ctDNA associate with long-term outcomes for patients with cancer on treatment. By combining efforts and aggregating data across multiple clinical trials, we will be able to generate the evidence necessary to characterize ctDNA as an indicator of response faster than if any single organization tried to do so alone. The ctDNA for Monitoring Treatment Response (ctMoniTR) Project is designed to answer the important question: Do changes in ctDNA reflect response to treatment? The ctMoniTR Project is taking a stepwise approach to analyze data across multiple trials to evaluate associations between changes in ctDNA and patient outcomes (**Figure 1**).

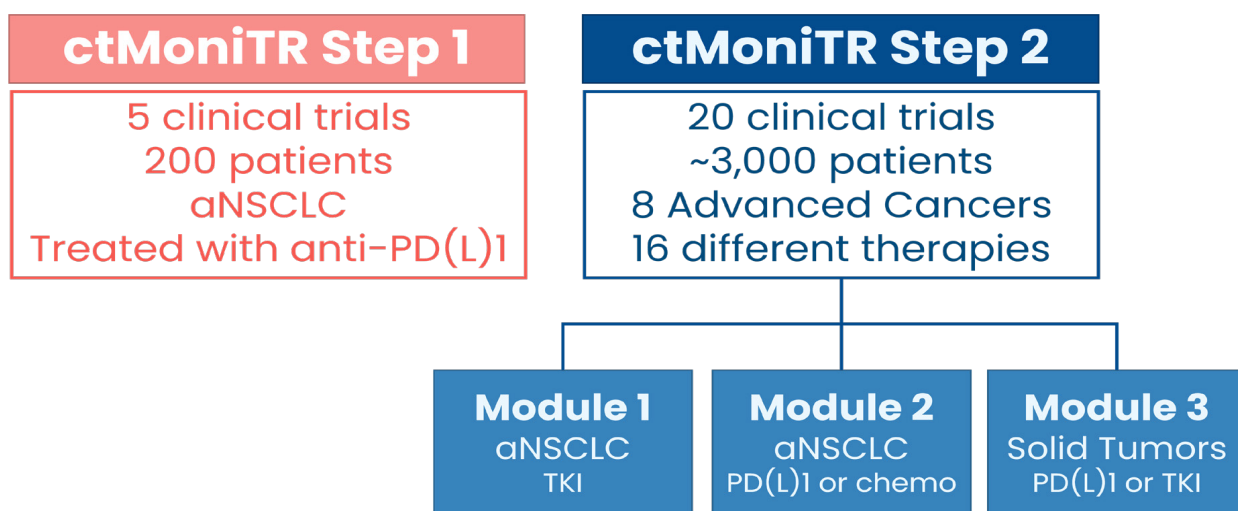


Figure 1. Overview of the ctMoniTR Project.

Findings from Step 1 showed robust and consistent associations between changes in ctDNA and patient outcomes for patients with advanced NSCLC (aNSCLC) receiving immunotherapy.² Step 2 of the ctMoniTR Project expands the scope of this research to study the associations between ctDNA and clinical outcomes across several clinical settings, drug classes, and cancer types. Data will be released throughout 2023 and 2024.

In addition to evaluating the use of ctDNA as an early endpoint, it is important to understand the impact assay technology and tumor biology may have on the use of ctDNA in oncology drug development. To establish evidence regarding baseline sensitivity metrics for ctDNA detection across cancer types, stages, and assays, *Friends* initiated a collaborative effort involving multiple diagnostic test developers called the Baseline ctDNA Project. A descriptive meta-analysis will be performed to compare trends in baseline ctDNA levels (ctDNA levels prior to a current cancer treatment) between cancer types and stages (**Figure 2**). A greater understanding of the biological landscape of baseline ctDNA levels will inform a conceptual framework for the use of ctDNA as an early endpoint predictive of long-term outcomes.

Baseline ctDNA	
<p><u>NSCLC</u></p> <ul style="list-style-type: none"> • 2,327 early-stage samples • 63,127 late-stage samples 	<p><u>Late-Stage</u></p> <ul style="list-style-type: none"> • 11,235 prostate • 10,532 breast • 1,359 bladder • 956 HNSCC
8 ctDNA assays	

Figure 2. Overview of the Baseline ctDNA Project.

Moving Forward

Both of these projects fill important data gaps outlined in an evidentiary roadmap created by key stakeholders to advance the use of ctDNA as an early endpoint.³ At the July 11th meeting “Establishing Evidence: New Advancements Using ctDNA” new data and insights will be shared regarding the use of ctDNA in oncology drug development, which will support ongoing research and regulatory discussions around its use as an early endpoint for regulatory processes.

Citations

1. FDA Draft Guidance Document. Use of Circulating Tumor Deoxyribonucleic Acid for Early-Stage Solid Tumor Drug Development. <https://www.fda.gov/media/158072/download>. (Accessed 6/20/23).
2. Vega D., et al. Changes in Circulating Tumor DNA Reflect Clinical Benefit Across Multiple Studies of Patients With Non-Small-Cell Lung Cancer Treated With Immune Checkpoint Inhibitors. *JCO Precis Oncol.* 2022 Aug;6:e2100372.
3. Friends of Cancer Research White Paper. Circulating Tumor DNA in Development of Therapies for Cancer: An Evidentiary Roadmap to an Early Endpoint for Regulatory Decision-Making. https://friendsofcancerresearch.org/wp-content/uploads/Circulating_Tumor_DNA_in_Development_of_Therapies_for_Cancer-Evidentiary_Roadmap.pdf

Changes in ctDNA levels as an early indicator of outcomes in advanced NSCLC treated with TKI: Initial findings from a retrospective aggregate analysis of 8 clinical trials.

Hillary Stires, Nevine Zariffa, Megan Eisele, Emily Goren, Carin R. Espenschied, Minakshi Guha, Jiannan Guo, Dilafruz Juraeva, Jean-Francois Martini, Brittany Avin McKelvey, Diana Merino Vega, Katherine K Nishimura, Gary Anthony Pestano, Sorena Rahmanian, Adam Rosenthal, Mark Stewart, Anna M. Szpurka, Antje Hoering, Jeff Allen, ctMoniTR Step 2 Working Group; Friends of Cancer Research, Washington, DC; NMD Group, Bala Cynwyd, PA; Cancer Research And Biostatistics, Seattle, WA; Cancer Research and Biostatistics, Seattle, WA; Guardant Health, Inc., Redwood City, CA; Takeda Pharmaceuticals, Cambridge, MA; Agilent, Kirkland, WA; Oncology Bioinformatics, The healthcare business of Merck KGaA, Darmstadt, Germany; Pfizer Inc., San Diego, CA; AstraZeneca, Gaithersburg, MD; Bodesix, Boulder, CO; Illumina, San Diego, CA; Loxo@Lilly, Eli Lilly and Company, Indianapolis, IN

Background: To determine whether changes in circulating tumor DNA (ctDNA) levels reflect treatment outcome, Friends of Cancer Research created the ctDNA to Monitor Treatment Response (ctMoniTR) Project with collaborators from industry, government, academia, and advocacy. A prior ctMoniTR effort analyzing 5 clinical trials (CT) showed an association between decreases in ctDNA levels and improved outcomes in patients with advanced non-small cell lung cancer (aNSCLC) treated with an anti-PD-(L)1. The current study expands that work and focuses on CT investigating tyrosine kinase inhibitors (TKI) treatment in oncogene-driven aNSCLC. **Methods:** We performed a retrospective analysis of patient-level clinical and ctDNA data across 8 CT representing 1,015 patients with aNSCLC treated with TKI (i.e., anti-EGFR, ALK, RET, or MET). Patients included in the analysis had a baseline ctDNA measurement (T0), an on treatment ctDNA measurement within 10 weeks of treatment initiation (for those with multiple ctDNA measurements within 10 weeks, we used the lowest measurement within 10 weeks) (T1), and overall survival (OS) data (n=749). CT used different ctDNA collection timepoints and assays. We randomly divided the dataset into training (2/3 of the data) and validation (1/3 of the data) datasets stratified by CT cohort (i.e., arm), age, tumor stage, and prior lines of therapy, then ran initial analyses on the training dataset (n=501; reported herein). ctDNA change was calculated as the percent change in maximum variant allele frequency (VAF) between T0 and T1 using tumor-derived variants provided by sponsors for each unique patient sample. CT used either ddPCR or an NGS assay. ctDNA limits of detection were assay specific and varied across CT. Multivariate analyses are ongoing and validation dataset analyses will be conducted. **Results:** At T0, 141 patients had non-detected (ND) ctDNA and 360 patients had detected (D) ctDNA. Of these, 27% (n=136) had ND ctDNA at both T0 and T1 (“ND/ND”), 52% (n=260) had changes from D at T0 to ND at T1 (“D/ND”), 12% (n=60) had at least a 50% decrease from T0 to T1 (“decrease”) and 9% (n=45) had an increase or a less pronounced decrease in ctDNA. In a univariate analysis, patients with ND/ND and D/ND were associated with improved OS compared to the decrease group. In addition to other characteristics, patients with max VAF \leq 0.5% or ND at T0 (n=214, 43%) had improved OS (HR=0.44, P<0.001) compared to those with max VAF >0.5% at T0 (n=287, 57%). **Conclusions:** In a retrospective aggregate analysis of 8 CT, ND ctDNA at T1 was associated with improved OS in patients with aNSCLC treated with TKI. Changes in ctDNA levels, particularly from D to ND, may provide an early indication of treatment benefit and predict long-term outcomes in this population. Additional ctMoniTR analyses are ongoing to validate the potential use of ctDNA as an early endpoint. Research Sponsor: Friends of Cancer Research.

Scan to view poster





ABOUT FRIENDS OF CANCER RESEARCH

Friends of Cancer Research (*Friends*) works to accelerate policy change, support groundbreaking science, and deliver new therapies to patients quickly and safely. We unite scientists, industry researchers, patient advocates and policy makers with shared trust and guide them toward meaningful cooperation.