July 20, 2018

COST DRIVERS IN THE DEVELOPMENT AND VALIDATION OF BIOMARKERS USED IN DRUG DEVELOPMENT

FINAL REPORT

Submitted to

Erin Rubens
Office of the Assistant Secretary of Planning and Evaluation
U.S. Department of Health and Human Services
200 Independence Ave. S.W.
Washington DC 20201

Submitted by

Booz Allen Hamilton
8283 Greensboro Drive
McLean, VA 22102

Contract No HHSP233201500132I
Task Order HHSP23337004T

TABLE OF CONTENTS

LIST OF EXHIBITS

ACKNOWLEDGEMENTS

We gratefully acknowledge Erin Rubens (ASPE) and Chazeman Jackson (ASPE) for their leadership, guidance, and feedback throughout this project. We also would like to thank the Project Advisory Group, including Shashi Amur (FDA), Zeid El-Kilani (ASPE), Amber Jessup (ASPE), Kirk Kerr (FDA), Michael A. Pacanowski (FDA), Anuradha Ramamoorthy (FDA), and Casey Sullivan (ASPE).

The core Booz Allen Hamilton team consisted of Matt Aldag, Kristin Baxter, Diane Epperson, Rebecca Hutchins and Joanna Yeh. We also would like to acknowledge the contributions of our subject matter experts, Leora Schiff, and Glenn Miller (independent consultants to Booz Allen Hamilton) to the data gathering and analysis of this project.

Many biomarker experts provided valuable information for the study. We are grateful to all of them for sharing their expertise and experiences with us.

EXECUTIVE SUMMARY

Despite the increasing investment and interest in drug development, the amount of time and resources needed to develop a new drug continues to rise. Biomarkers[1] are an important tool with the potential to decrease the time, cost, and failure rate of drug development. Given the role of biomarkers in drug development, it is important to understand how biomarkers are identified, developed, and validated,[2] so that the costs of these activities can be assessed. To promote development and validation of biomarker candidates, a better understanding of associated costs can encourage dedication of resources and investments by supporting organizations. Currently, there is limited published information and cost data describing the financial and other resources required for biomarker development that examines discrete phases of development and differences between biomarkers.

The overall objective of this project was to develop a framework for estimating the cost of biomarker development and validation for use in drug development. To achieve this objective, a comprehensive understanding of the biomarker validation process, activities that drive the process, and factors that drive cost was needed. This project was also designed to address the following five research questions:

1. How are biomarkers identified for use in drug development?
    a. How do these processes differ depending on the category of biomarker?
2. Given that regulatory acceptance of a biomarker is a data-driven process, what are the sources of data necessary for biomarker development?
    a. What is the range of costs associated with a given component of a biomarker development program (e.g., validation of a biomarker's analytical method and its associated platform)?
    b. What factors influence the cost of developing and validating biomarkers for use in drug development?
3. How do these factors vary across the selected biomarker categories?
4. How do these factors contribute to the overall cost of developing and validating biomarkers within a particular category?
5. How do the cost drivers compare across biomarker categories?

The methods of this project included an environmental scan, case studies, and cost framework development. The environmental scan gathered data using a literature review of publicly available information on biomarkers and consultations with experts in the field of biomarker development. The case studies examined the trajectories of six specific biomarkers and focused on the factors that affected the process and investment needed to develop and validate the biomarkers. The cost framework was designed to identify the key challenges and drivers influencing the cost of biomarker development, outline the variabilities associated with these challenges and drivers, and provide a methodology for estimating development cost across the development and validation lifecycle, including costs associated with failures.

[1] Biomarker: A defined characteristic that is measured as an indicator of normal biological processes, pathogenic processes, or responses to an exposure or intervention, including therapeutic interventions.

[2] Validation: A process to establish that the performance of a test, tool, or instrument is acceptable for its intended purpose.

This project focused primarily on four categories of biomarkers: predictive biomarkers,[3] prognostic biomarkers,[4] safety biomarkers,[5] and surrogate endpoints.[6] Additionally, it included more detailed findings for two of these biomarker categories: predictive biomarkers and surrogate endpoints.

Overall findings from this project identified and characterized four phases of the biomarker development and validation process: 1) Identification and Feasibility; 2) Develop Assay and Intended Use; 3) Analytical Validation; and 4) Clinical Validation and Utility shown in Exhibit E1. These phases were common between all four biomarker categories.

Exhibit E1: Biomarker Identification, Development and Validation Lifecycle

Different categories of biomarkers tend to share common development and validation process phases as well as process drivers (Exhibit E2), therefore the process and timeline to develop and validate biomarkers is similar across biomarker categories, with a few differences. One difference is the longer length of time needed for the Clinical Validation and Utility phase for prognostic biomarkers and surrogate endpoints, compared to predictive and safety biomarkers. This difference is attributable in part to the generally shorter length of follow-up time it takes to validate predictive and safety biomarkers with the outcomes they are measuring.

Exhibit E2: Timeframe for Biomarker Development

	Identification and Feasibility	Develop Assay and Intended Use	Analytical Validation	Clinical Validation and Utility
Predictive biomarkers	1 month to 3 years	3 months to 1.5 years	9 months to 1.5 years	10 months to 3 years
Prognostic biomarkers	3 months – 1 year	6 months – 2 years	1.5 years	1.5 – 10 years
Safety biomarkers	3 months – 1 year	3 months – 6 years	6 months – 1 year	6 months – 6 years
Surrogate endpoints	1-5 years	3 months to 3 years	1-2 years	1-10 years

Findings indicated that overall costs of biomarker development and validation varied considerably, both within and between each of the biomarker categories (Exhibit E3). Development and validation is most costly for surrogate endpoints, due in large part to the clinical trials and clinical validation needed.

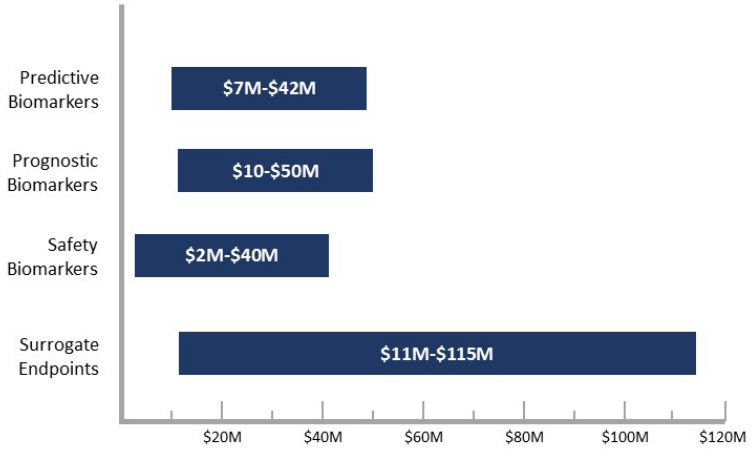
[3] Predictive biomarker: A biomarker used to identify individuals who are more likely than similar individuals without the biomarker to experience a favorable or unfavorable effect from a specific intervention or exposure.

[4] Prognostic biomarker: A biomarker used to identify the likelihood of a clinical event, disease recurrence, or progression in patients who have the disease or medical condition of interest.

[5] Safety biomarker: A biomarker used to indicate the presence or extent of toxicity related to an intervention or exposure.

[6] Surrogate endpoint: An endpoint that is used in clinical trials as a substitute for a direct measure of how the patient feels, functions, or survives; it does not measure the clinical benefit of primary interest in and of itself but rather is expected to predict the clinical benefit or harm based on epidemiological, therapeutic, pathophysiological, or other scientific evidence.

Exhibit E3: General Cost Range by Biomarker Category



This project identified and characterized seven cost drivers that can potentially impact overall biomarker development and validation costs: Scientific Knowledge Base, Sample Type, Sample Size, Novelty and Complexity of the Technology, Trial Recruitment, Follow-Up Time Needed, and Developmental Model (Private or Consortium).

Additionally, a cost framework analysis was developed for two biomarker categories: predictive biomarkers and surrogate endpoints. The cost framework identified variations in cost drivers between these two biomarker categories and phases and characterized interactions between cost drivers. The cost framework also provided cost estimates for each biomarker category, including costs associated with each development and validation phase (Exhibit E4).

Exhibit E4: Cost Ranges for Predictive Biomarkers and Surrogate Endpoints by Phase of Biomarker Development

	Predictive Biomarkers (Mean, range [min/max])		Surrogate Endpoints (Mean, range [min/max])	
Identification and Feasibility	Mean: $1,315,000		Mean: $1,983,333	
	Min: $550,000	Max: $3,000,000	Min: $500,000	Max: $5,000,000
Develop Assay and Intended Use	Mean: $2,757,000		Mean: $2,805,000	
	Min: $1,300,000	Max: $6,000,000	Min: $65,000	Max: $7,000,000
Analytical Validation	Mean: $3,668,750		Mean: $8,375,000	
	Min: $2,000,000	Max: $5,000,000	Min: $2,000,000	Max: $20,750,000
Clinical Validation and Utility	Mean: $8,875,000		Mean: $10,341,667	
	Min: $1,500,000	Max: $33,000,000	Min: $4,000,000	Max: $20,750,000
Total	Mean: $15,700,000		Mean: $23,505,000	

A discussion on biomarker failure drivers is also presented, and provides a method for calculating the cost of failure for biomarker development for application of cost data acquired in the future.

Due to project scope and characteristics of cost data, certain limitations of this project included data availability, data quality, and data categorization. Outstanding knowledge gaps include detailed information about prognostic and safety biomarkers, a lack of reliable information about cost of failure, uniformity of cost data, and a comprehensive understanding of cost allocation for biomarkers developed and validated by a consortium of organizations. Some of these gaps (e.g., regarding cost of failure and consortia-driven development) could be addressed by future efforts to convene subject matter experts with relevant knowledge in these areas.

1 INTRODUCTION

The current therapeutic development process takes 10 to 15 years from discovery to licensure and the investment required to bring a new drug to market has increased substantially in recent years. Since 2003, the cost of developing a new therapeutic product has increased at an annual rate well above the general price of inflation and some estimates now place the cost at between $1.4 and $2.6 billion per approved new compound (DiMasi et al., 2016). A recent study of 10,000 clinical trials over the past ten years found the overall likelihood of approval of drugs that enter phase I clinical trials is 9.6 percent. While the underlying reasons for the escalating timeline, costs and failure rates are complex, the high failure rate of late stage clinical trials is a major contributing factor (Thomas et al., 2016).

Biomarkers[7] are an important tool with the potential to decrease the time, costs, and failure rate of drug development. A recent study highlighted that inclusion of biomarker testing to select the clinical trial population was shown to increase the probability of success at every transition in drug development from phase I through phase III (Lopez et al., 2015). Importantly, the biggest increase in the probability of success occurs for drugs in phase III, a particularly significant transition point given the high costs of failure at this stage (Lopez et al., 2015).

BIOMARKER IMPACT ON DRUG APPROVAL

When biomarkers were used to select clinical trial populations, the probability of successful advancement to approval increased from 1:10 to 1:4 (Thomas, 2016).

Selection of individuals most likely to benefit from a drug or therapy is just one way that biomarkers can be used to improve drug development. Biomarkers can also help researchers monitor therapeutic response, assess safety, and evaluate clinical benefit sooner. Pharmacodynamic biomarkers can help monitor whether investigational drugs and therapeutics are hitting their target and can help reduce the overall costs of drug development by enabling researchers to identify ineffective drug candidates and halt development efforts more rapidly. Biomarkers that can identify safety issues earlier, ideally at a time when the unexpected harmful reaction to a therapeutic is still reversible, also have an important role to play in drug development. A recent portfolio review by AstraZeneca revealed that the top reason drug development projects failed or were closed was due to an unacceptable safety profile (Cook et al., 2014). Most of the safety failures occurred in the preclinical testing phase, highlighting the need for both preclinical and clinical safety biomarkers (Cook et al., 2014). Biomarkers that can serve as a substitute for a direct measure of how a patient feels, functions or survives, also known as surrogate endpoints, can dramatically reduce the time and investment needed to evaluate the therapeutic benefits of a new therapy. For example, estimated glomerular filtration rate (eGFR) is currently under development for use as a surrogate endpoint of kidney function and disease progression in chronic kidney disease. If successful, using eGFR as a surrogate endpoint to replace the existing measures would reduce the required clinical trial size by 20-35 percent, reduce the follow-up period in a clinical trial by two to three years and permit the inclusion of patients with near normal or mildly decreased eGFR (Badve et al., 2016).

Given the important role that biomarkers can play in drug development, it is important to understand how biomarkers are identified, developed, and validated.[8] Biomarkers have their own discovery and validation process that can vary in cost, duration and complexity depending upon the biomarker's purpose and the medical area(s) it serves. Adding to the complexity, biomarkers can be used individually or in sets or panels. Characteristics of the biomarker development and validation process, which vary by biomarker, determine

[7] Biomarker: a defined characteristic that is measured as an indicator of normal biological processes, pathogenic processes, or responses to an exposure or intervention, including therapeutic interventions.

[8] Validation: a process to establish that the performance of a test, tool, or instrument is acceptable for its intended purpose.

individual cost drivers[9] that impact overall project costs. Other factors that can potentially impact the cost and time required to develop and validate new biomarkers for drug development may include:

- Timing of biomarker development with relevant drug development efforts;
- Technology/platform used to measure the biomarker;
- Type and availability of biological samples needed to assess the biomarker; and
- Size and complexity of clinical studies needed to validate the biomarker.

There is limited published information and cost data describing the financial and other resources required for biomarker development that examines discrete phases of development and differences between biomarkers. The paucity of cost information is due to several factors, including: 1) the sensitive nature of proprietary information for commercially developed biomarkers; 2) long development times and multiple funding mechanisms for academic and consortia developed biomarkers; and 3) the complexity of assigning costs when biomarkers are developed using repurposed data and clinical trial materials originally collected for other purposes.

2 OBJECTIVES

The overall objective of this project was to develop a framework for estimating the cost of validating new biomarkers for use in drug development. This objective was further defined by specific project goals, scope and research questions described in the following sections.

2.1 Project Goals

Biomarker development is a complex, technical, resource-intensive process that occurs in a rapidly-changing environment of scientific and technical advances. Biomarker development plays a critical role in drug development, which is an equally complex, expensive and rapidly-evolving endeavor. The project sought to characterize the process for identifying, developing, and validating new biomarkers for use in drug development. Based on a greater understanding of these activities, the goal was to develop a framework for estimating the cost of validating new biomarkers for use in drug development.

2.2 Project Scope

This project focused primarily on four categories of biomarkers: predictive biomarkers, prognostic biomarkers, safety biomarkers and surrogate endpoints. Additionally, the project included more detailed findings for two of these biomarker categories: predictive biomarkers and surrogate endpoints. All biomarker categories were defined according to the Biomarkers, Endpoints and other Tools (BEST) Glossary (FDA-NIH Biomarker Working Group, 2016).

- Predictive biomarkers: A biomarker used to identify individuals who are more likely than similar individuals without the biomarker to experience a favorable or unfavorable effect from a specific intervention or exposure.
- Prognostic biomarkers: A biomarker used to identify the likelihood of a clinical event, disease recurrence, or progression in patients who have the disease or medical condition of interest.
- Safety biomarkers: A biomarker used to indicate the presence or extent of toxicity related to an intervention or exposure.

[9] Cost Driver: A factor that influences the costs of developing and validating new biomarkers. To be considered a cost driver, the factor had to affect the development and validation costs of at least two biomarkers explored in the case studies by at least 15 percent in one of the four developmental phases.

- Surrogate endpoints: An endpoint that is used in clinical trials as a substitute for a direct measure of how the patient feels, functions, or survives; it does not measure the clinical benefit of primary interest in and of itself but rather is expected to predict the clinical benefit or harm based on epidemiological, therapeutic, pathophysiological, or other scientific evidence.

While there are various ways to define the starting point of biomarker development, for the purposes of this project, the starting point for biomarker development was considered the point at which the intention is to develop a clinically useful biomarker that is indicative of biological processes, whether normal, pathogenic, or in response to an exposure/intervention. Biomarker candidates stemming directly from basic research are not in the scope of this project because they are not subjected to validation testing. For example, DNA hypomethylation has been explored in a research setting as a potential surrogate endpoint to monitor the efficacy of chemopreventive agents (Tao et al., 2004). However, considering that an assay capable of measuring this biomarker has yet to be developed for use in clinical trials, it is considered out of scope for the current effort. The scope of the project included, however, large screening assays performed to identify a potential biomarker for specific drug candidates, as well as any research conducted to define the biological and pathological relevance of the candidate biomarkers as they relate to drugs in development.

Similarly, there are various ways to define the completion of biomarker development. Depending upon the intended use, different biomarkers will require differing levels of validation. A greater burden of validation is generally required for biomarkers that directly inform treatment decisions compared to those that inform research decisions. For instance, biomarkers that are used to identify the population for which a drug is safe and effective or are included as companion diagnostics on the drug label require a much higher level of validation than biomarkers used to inform "go/no-go" decisions during drug research and development. Since this project was focused on the costs of validating a new biomarker for drug development, it considered the completion of biomarker development as achieving the level of biomarker validation needed for use in drug development. This endpoint can vary based on the category and biomarker's context of use. The additional costs of taking a validated biomarker through the Food and Drug Administration's (FDA) biomarker qualification process or demonstrating clinical utility that ultimately leads to use of a biomarker in clinical practice (in the post-market setting) are out of scope for the current effort. Refer to Exhibit 8 in Section 4.3.1. for an overview of the biomarker identification, development and validation lifecycle from beginning to end.

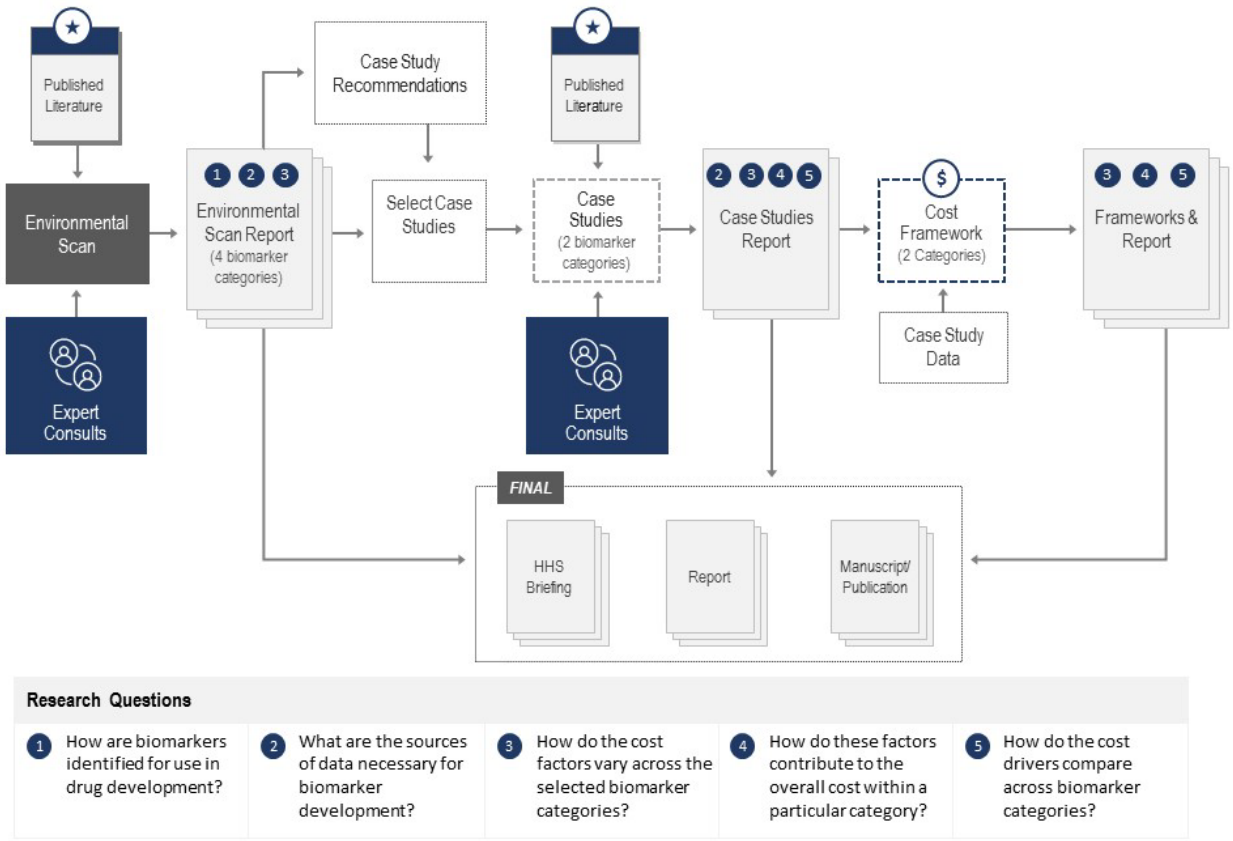
This project was designed to address the following five research questions:

1. How are biomarkers identified for use in drug development?
    a. How do these processes differ depending on the category of biomarker?
2. Given that regulatory acceptance of a biomarker is a data-driven process, what are the sources of data necessary for biomarker development?
    a. What is the range of costs associated with a given component of a biomarker development program (e.g., validation of a biomarker's analytical method and its associated platform)?
    b. What factors influence the cost of developing and validating biomarkers for use in drug development?
3. How do these factors vary across the selected biomarker categories?
4. How do these factors contribute to the overall cost of developing and validating biomarkers within a particular category?
5. How do the cost drivers compare across biomarker categories?

3 METHODS

The overall methodology of the project included an environmental scan, case studies and cost framework analysis. Exhibit 1, below, depicts how the project methodology addressed the five research questions. An environmental scan focused on the research questions 1 - 3. A series of case studies conducted following the environmental scan provided additional information to research questions 2 - 3 and served to inform research questions 4 - 5. Following the case studies, a cost framework analysis further informed research questions 3 – 5.

Exhibit 1: Project Methodological Overview (Text version of exhibit 1)



3.1 Environmental Scan

The environmental scan focused on addressing the first three of the five overarching research questions described in Section 2.2. The goal of the environmental scan was to explore how biomarkers are identified and validated for use in drug development, including how these processes differ for the four different categories of biomarkers covered in the environmental scan:

- Predictive biomarkers (including companion diagnostics)
- Prognostic biomarkers
- Safety biomarkers
- Surrogate endpoints

The environmental scan was designed to be the foundation of the entire project and provide a broad perspective. It encompassed general insights into the biomarker categories, processes for biomarker development and cost drivers. General validation process drivers and cost drivers are introduced in the environmental scan.

During the environmental scan, information was gathered using two methods: (1) a literature review of publicly available information on biomarkers; and (2) consultations with experts in the field of biomarker development (Appendix C). Publicly available information gathered during the literature review was used to identify processes and cost drivers common within and among the four categories of biomarkers included in the environmental scan. Using that information, a draft process model and list of cost factors were developed, presented during the expert interviews, and further refined during the consultations with experts as described in Section 3.1.2. The information gathered from these two sources was synthesized as described in Section 3.1.3, and is the basis for the findings presented in this report.

3.1.1 Literature Review

The purpose of the literature review was to gather publicly available information that would help address the project research questions. The literature review included peer-reviewed publications, guidelines, conference presentations, workshop proceedings, review articles and meta-analyses. The search engines used to identify sources included PubMed and Google Scholar. The initial list included approximately 100 sources and was prioritized for those publications that presented a synthesis on multiple studies in a given topic area, such as reviews, meta-analyses, guidelines and conference workshops. Literature was further down-selected for those containing data on specific biomarkers, information on the biomarker validation process and cost drivers for each of the four biomarker categories included in the environmental scan. Findings from the literature review were extracted using qualitative thematic analyses and synthesized into draft models for further refinement and validation during the expert consultations.

3.1.2 Expert Consultations

The purpose of the expert consultations was to refine and validate the draft models prepared with the findings from the literature review and gain deeper insights into the processes and cost drivers associated with biomarker development. The consultations focused on engaging biomarker experts with a background in industry who could speak to the budget needed to identify, develop, and validate new biomarkers. Biomarker experts with experience leading biomarker development consortia were also included in the consultations. A total of eleven experts were consulted for the environmental scan. Experts were chosen based on their knowledge of at least one of the biomarker categories. In total, each of the four biomarker categories had at least four experts with expertise in that category (Appendix C).

Each consultation lasted one hour and covered three main areas of inquiry: (1) Biomarker Insights; (2) Biomarker Discovery, Development and Validation Model; and (3) Cost Factors and Drivers for Biomarker Development and Validation. Due to the sensitive and proprietary nature of cost data, interviewees were not asked about the costs associated with a specific biomarker, but in general as a possible range for a given biomarker category or step in the development and validation process. All consultations were performed via teleconference and at a minimum two team members were present for each consultation, one to focus on guiding the consultation and a second to capture discussion notes.

3.1.3 Data Analysis and Synthesis

The findings from the literature review and expert consultations were reviewed to identify:

- Relevant disease areas;
- Biomarker uses in drug development;
- Processes to identify, develop, and validate biomarkers;
- Types of data needed for biomarker validation; and
- Factors that influence the cost of biomarker development.

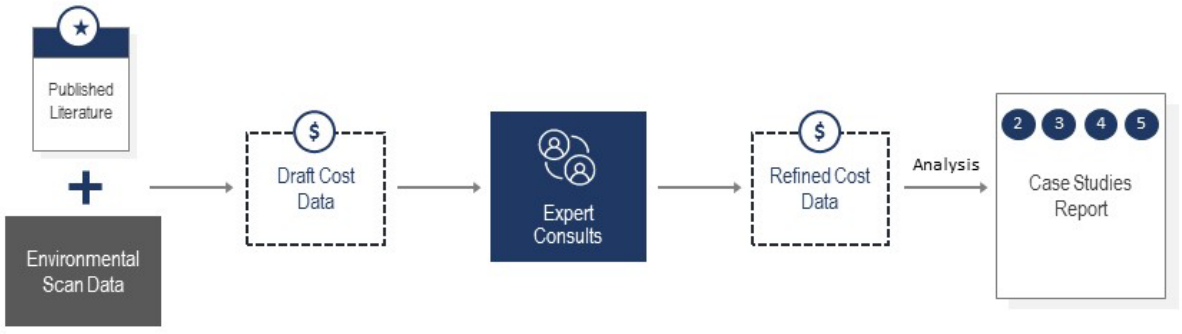
For each of these topics, the findings were extracted using qualitative thematic analysis. Particular attention was paid to the variability in these factors within and between biomarker categories.

3.2 Case Studies

The case studies examined the trajectories of six biomarkers (three predictive biomarkers and three surrogate endpoints) and focused on the factors that affected the process and investment needed to develop and validate the biomarkers. The case study methodology (Exhibit 2) consisted of four steps:

1. Gathering information from the published literature and using it to understand the context of use for a biomarker; the process through which the biomarker was identified, developed, and validated; and potential factors that influenced the amount of time and resources needed to validate the biomarker for use in drug development.
2. Combining data from the previously conducted environmental scan with the information gathered from the published literature evaluated in step 1 to develop draft cost data for each case study.
3. Discussing and refining the draft cost data during the expert consultations.
4. Synthesizing information gathered from published literature, the environmental scan, and expert consultations to form the basis for the findings presented in the case studies report.

Exhibit 2: Overview of Case Study Methodology



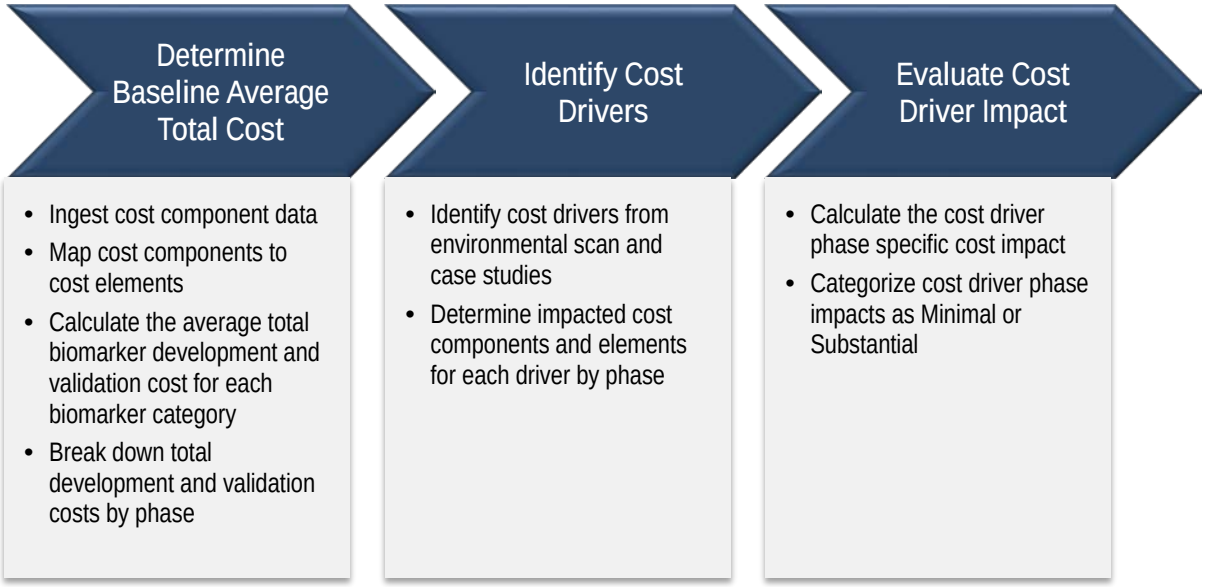
3.3 Cost Framework

The cost framework was designed to identify the key challenges and drivers influencing the cost of biomarker development, outline the variabilities associated with these challenges and drivers, and provide a methodology for estimating development cost across the development and validation lifecycle, including costs associated with failures. The cost framework report focused on predictive biomarkers and surrogate endpoints, the two biomarker categories addressed in the case studies.

3.3.1 Cost Framework Methodology

The cost framework was informed by findings from the environmental scan and the case studies. Exhibit 3 summarizes the methodology used to create the cost framework which included determining the baseline average total cost, identifying cost drivers and evaluating cost driver impact.

Exhibit 3: Cost Framework Methodology



3.3.1.1 Determine Baseline Average Total Cost

During the case studies, detailed information on the costs of developing and validating new biomarkers was gathered from subject matter experts (SMEs) for six biomarkers: three predictive biomarkers and three surrogate endpoints (for additional information, see Appendix D).

When the SMEs engaged during the case studies were unable to provide cost data for a specific cost component,[10] generalized cost information gathered from SMEs during the environmental scan was used. When no cost data was available from SMEs engaged during the environmental scan or case studies, an average annual cost estimate was generated based on industry standard data and validated with internal biomarker SMEs. These average annual cost estimates were assumed to be similar between predictive biomarkers and surrogate endpoints (e.g., salaries [for technician or radiologist] and data storage costs do not vary by biomarker category). Given that the expert input and industry average cost data used for the framework were contemporaneous, there was no additional normalization for inflation, and it can be assumed that cost estimates are given in 2016 dollars. Furthermore, because this report describes cost impact[11] as ranges rather than single point estimates, adjustments for inflation would likely have limited effects.

[10] Cost Component: The most granular representation of line item costs involved in the biomarker development and validation process. Cost components were identified from information gathered from interviews with industry subject matter experts (SMEs) or relevant industry data. Because cost component data was gathered from several sources, individual cost components may vary based on the data source due to differences in how the organization accounts for cost.

[11] Cost Impact: Describes the change in base expected biomarker development and validation costs due to cost driver effects. Cost impacts are incurred as an additional cost to base biomarker development and validation costs and are given as the

After assigning cost information gathered during case studies to cost components and filling in any missing cost data with industry averages, the average costs of developing and validating a new biomarker were calculated for predictive biomarkers and surrogate endpoints. The average costs were broken down by phase of biomarker development and cost component. These average costs were used to assess the impact of specific factors to assess whether they met the criteria for being considered a cost driver as described in Section 3.3.1.2 and evaluate the impact of cost drivers by phase as described in Section 3.3.1.3.

3.3.1.2 Identify Cost Drivers

A cost driver is defined as a factor that influences the costs of developing and validating a biomarker used in drug development (see Section 4.4). While many factors influence the cost of developing and validating new biomarkers, the purpose of the framework is to identify common factors that meaningfully influence the cost of developing and validating a new biomarker. During the environmental scan and case studies, experts highlighted several considerations that influenced the costs of developing and validating a new biomarker. The considerations noted by experts as key factors that influenced costs were used as a starting point to identify cost drivers. However, a factor affecting the cost of biomarker development and validation was only considered a cost driver if it influenced the development and validation costs of at least two biomarkers explored in the case studies by at least 15 percent in one of the four developmental phases.

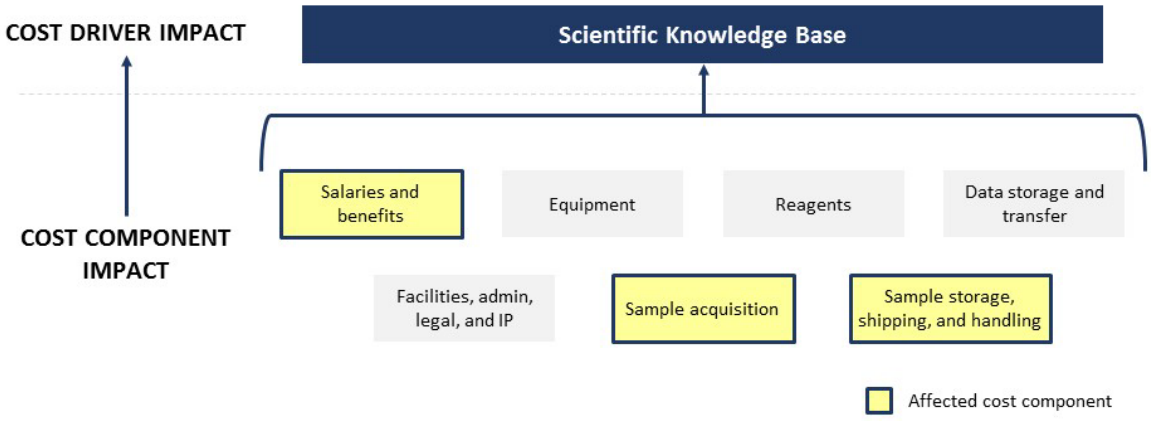
3.3.1.3 Evaluate Cost Driver Impact

For each cost driver, the associated cost components the cost driver affected within each phase of the development and validation process (see Section 4.3.1) were identified. Common ways cost drivers affect cost components are through extending the time spent in a phase, changing types of labor required (e.g., increasing or decreasing the amount of highly skilled labor required), changing the number or type of samples required, and requiring specialized equipment.

Once the affected cost components were identified, the overall change in costs was assessed for each driver by phase. For each cost driver, the incremental cost impacts were estimated as a percentage of baseline biomarker development and validation expenses. If a cost component was impacted, then an estimate was made to determine the magnitude of impact. To quantify the impact to cost components, data gathered from the case studies, environmental scan, SMEs, and industry averages on unit and annual costs were used. The individual impact to cost components was quantified, and then aggregated to the cost driver level. Exhibit 4 depicts how the impacts of a cost driver by phase were assessed. The cost components in yellow are those that substantially impact the Scientific Knowledge Base cost driver (see Section 4.4) in the Identification and Feasibility phase (see Section 4.3.1).

maximum of potential costs that could be incurred. The cost impact may be provided as a percentage change in base biomarker development and validation costs (e.g., a 10 percent cost impact means that the new expected cost for biomarker development and validation is 110 percent of base costs), or as a dollar amount (e.g., a $1,000,000 cost impact means the new expected cost for biomarker development and validation is increased by $1,000,000).

Exhibit 4: Cost Driver and Cost Component Impact Relationship Map



For example, if more labor and samples are required for the Identification and Feasibility phase (see Section 4.3.1.1) as a result of the Scientific Knowledge Base cost driver (see Section 4.4.1), the cost components impacted include salaries and benefits, sample acquisition, and sample storage, shipping, and handling. To evaluate the cost impact, the additional cost expected to be incurred for additional full-time equivalents and samples was calculated using unit cost[12] data. After the cost impact on each cost component was calculated, cost component impacts were summed to find the total cost impact on the cost driver level for the Identification and Feasibility phase (see Section 4.3.1).

To categorize the impact of cost drivers on the biomarker development and validation costs by phase, thresholds were established to categorize the change to the phase costs. Based on patterns observed in the cost impact data, cost driver impacts were categorized as "Minimal," where the cost impact was low, and "Substantial," where the cost impact was high. For consistency, a 15 percent increase in phase costs was used as a threshold to categorize cost driver impacts. Any cost driver impact that had less than a 15 percent increase in phase costs was categorized as "Minimal," and any cost driver impact that had a 15 percent or greater increase in phase costs was categorized as "Substantial." It is important to note that a Minimal cost impact categorization does not necessarily imply that no additional costs are incurred, only that the expected cost impacts are lower than cost drivers with Substantial cost impact.

The cost impact categorizations were based on the percent change in the biomarker development and validation costs by phase. Because cost impacts are calculated by phase, the Minimal and Substantial categorizations are best suited to comparisons across cost drivers within the same phase. It is possible that Minimal cost impacts are different between phases. For example, a 15 percent increase in costs in the Identification and Feasibility phase (see Section 4.3.1) for predictive biomarkers represents approximately $350,000 in additional costs, whereas a 15 percent increase in costs in the Clinical Validation and Utility phase represents approximately $1.48 million in additional costs. Thus, a smaller percentage increase in costs in the Clinical Validation and Utility phase can have the same or greater dollar impact to overall biomarker development and validation costs.

[12] Unit Cost: Describes the cost incurred by an organization for one unit of a particular product or service.

3.3.2 Cost Impact Scenarios

In addition to the cost framework, cost impact scenarios were developed to describe how a cost driver can impact the overall cost of biomarker development and validation (Exhibit 5). The cost impact scenario descriptions were intended to help provide an understanding of the conditions that would result in an increased cost for a cost driver. The cost scenario[13] descriptions were not intended to be exhaustive, but rather to describe some of the characteristics that a biomarker may have in each category. However, it is unlikely than any one biomarker would exhibit all the characteristics listed for any one scenario. Determining which scenario best encapsulates a given biomarker requires discretion and judgment.

As shown in Exhibit 5, the low impact level describes average or baseline scenarios. The average or baseline scenario was constructed using data from case studies described in Appendix D. The potential cost range associated with high-cost scenarios was also based on information drawn from interviews with experts during the environmental scan and case studies. The high-cost scenarios represent a range of potential additional cost based on the general experience of experts with biomarker development, and not limited to the six biomarkers explored in the case studies.

Exhibit 5: Description of the Cost Scenarios

Impact Level	Definition
Low	Describes an ideal scenario in which the cost driver minimally increases (≤ 15 percent) total overall costs above the average
High	Describes conditions that would increase the average overall costs by more than 15 percent

It is important to note the difference between Minimum/Substantial and Low/High designations. Minimal and Substantial designations are used when categorizing the *maximum cost driver impacts by phase*. For example, if Scientific Knowledge Base did not impact Identification and Feasibility costs by 15 percent even in the *maximum* case, it was labeled as "Minimal." Otherwise, the cost driver was labeled "Substantial" and listed as a cost driver for that phase (see Exhibit 11 and Exhibit 14 for a list of Substantial cost drivers by phase).

Once the cost driver is labeled as Substantial, the Low and High descriptions were applied to characterize potential variations in associated costs. For example, as discussed in Section 4.4.2, Sample Type is a Substantial cost driver in the Clinical Validation and Utility phase. Associated sample costs[14] can vary from $65 per sample (Low) to $10,000 (High). The Low and High descriptions are to provide additional context into the plethora of different real-world scenarios related to these cost drivers and what the cost impact would look like.

3.3.3 Estimating the Cost of Failure

Cost of failure (see Section 4.6.4.1) is an important factor to discuss when considering biomarker development and validation costs. The following section presents a discussion on biomarker failure drivers

[13] Cost Scenarios: Describes how a cost driver affects overall costs of biomarker development and validation for a given biomarker category; the scenarios describe the conditions or biomarker characteristics that would lead to a substantial increase in costs over average costs.

[14] Sample Costs: Cost component that includes sample acquisition costs and sample shipping, handling, and acquisition costs.

and a method for calculating the cost of failure[15] for biomarker development. In order to estimate the cost of failure, both out-of-pocket costs[16] and the cost of capital[17] must be considered. However, there is limited literature and published data on biomarker failure rates and costs. Experts interviewed for the environmental scan and case studies were hesitant to discuss the cost of failure citing proprietary information or a lack of specific information. Thus, this section presents a discussion on biomarker failure drivers and a method for calculating the cost of failure for biomarker development when more data become available.

3.3.3.1 Cost of Failure Calculation Methodology

Cost of failure can be discussed in terms of the *expected direct cost of failure,*[18] *realized direct cost of failure,*[19] and *opportunity cost*.[20]

The expected direct cost of failure is the planned cost of failure, based on the biomarker development and validation budget and probability of failure. The probability of failure is the probability the biomarker is unsuccessful (i.e., the biomarker does not complete all four stages of the development lifecycle, and ultimately does not translate into clinical practice). The probability of failure is inversely related to the probability of success (i.e., probability of failure = 1 – probability of success). It is important to note that the expected direct cost of failure is the theoretical out-of-pocket cost for failed biomarker efforts, projected based on historical data, rather than the actual out-of-pocket cost incurred if the biomarker effort fails. The expected direct cost of failure is typically assessed prior to the undertaking of a biomarker development project. The expected direct cost of failure can be calculated as:

$$Expected\ Direct\ Cost\ of\ Failure = Biomarker\ Development\ and\ Validation\ Budget \times Overall\ Probability\ of\ Failure$$

Realized direct cost of failure is the actual cost of failure incurred at the time at which a specific biomarker investment is declared a failure and activity stops. The realized direct cost of failure is the total amount spent on the biomarker investment. However, even if the biomarker initiative is deemed a failure, some residual benefit may be derived from the work completed that can aid in future biomarker development initiatives. Thus, the gross direct cost of failure is the funds spent on the failed biomarker investment, but the net direct cost of failure accounts for the residual benefit gained. Gross realized direct cost of failure[21] can be calculated as:

[15] Cost of Failure: Describes costs associated with a biomarker investment that did not successfully complete the biomarker development and validation process. Cost of failure includes both out-of-pocket and opportunity costs. Direct cost of failure can be discussed in terms of expected direct, realized direct, gross realized direct, and net realized direct cost of failure.

[16] Out-of-pocket Costs: Costs incurred for biomarker development and validation, not accounting for opportunity costs.

[17] Cost of Capital: Refers to the rate of return expected by investors to persuade them to make a given investment, and represents the opportunity cost of making a specific investment. ; from: https://hbr.org/2015/04/a-refresher-on-cost-of-capital.

[18] Expected Direct Cost of Failure: The theoretical or planned cost of failure based on historical biomarker investment performance. Expected direct cost of failure can be calculated from the biomarker development and validation budget and probability of failure.

[19] Realized Direct Cost of Failure: The actual cost of failure for a specific biomarker investment. The realized direct cost of failure is the amount spent on the biomarker investment until activity stops and failure is declared.

[20] Opportunity Cost: Refers to the loss of potential gain from other alternatives due to the decision to pursue the biomarker investment. Also known as time cost.

[21] Gross Realized Direct Cost of Failure: The actual cost of failure based on the specific biomarker investment, not accounting for any residual benefit derived.

$$Gross\ Realized\ Direct\ Cost\ of\ Failure = Biomarker\ Investment\ Capital\ Spent$$

Net realized direct cost of failure[22] can be calculated as:

$$\begin{aligned} &Net\ Realized\ Direct\ Cost\ of\ Failure \\ &\quad = Biomarker\ Investment\ Capital\ Spent \\ &\quad - Residual\ Benefit\ Derived\ from\ Biomarker\ Investment \end{aligned}$$

In addition to the out-of-pocket costs associated with biomarker development, the cost of failure also needs to account for opportunity costs. The opportunity cost captures the potential gains given up by pursuing the specific biomarker as opposed to an alternative investment. To evaluate opportunity costs, the cost of capital needs to be examined. The cost of capital represents the rate of return that could have been earned by investing funds into a different investment with equal risk (Gallo, 2015). While the biomarker is in development, investors lose out on any potential investment returns they could have made from alternative investments using the same capital. The opportunity cost may be calculated as:

$$\begin{aligned} &Opportunity\ Cost \\ &\quad = (Biomarker\ Investment\ Capital\ \times Expected\ Rate\ of\ Return) \\ &\quad - (Alternative\ Investment\ \times Alternative\ Expected\ Rate\ of\ Return) \end{aligned}$$

3.3.3.2 Cost of Capital Methodology

The cost of capital is based on the weighted average of the cost of equity and the cost of debt, since capital can be raised either from debt or equity. The cost of equity can be calculated using the capital asset pricing model (CAPM). CAPM can be calculated as:

$$Expected\ Cost\ of\ Equity\ Capital = Risk\ Free\ Rate + (\beta \times Risk\ Premium)$$

Beta is a relative measure of the riskiness of an individual investment. The higher the beta, the higher the cost of capital discount rate and the higher the risk. A beta of greater than one indicates the investment is riskier than the market. Conversely, a lower beta indicates a lower discount rate and lower risk, but also lower returns. The beta for a less risky investment will be less than one. The risk-free rate typically corresponds with the rate of return associated with a riskless investment. A U.S. Treasury bond yield, while not technically risk-free, is considered an example of a risk-free rate because it is generally expected that, based on the credit rating of the U.S. federal government, the note will be paid with 100 percent certainty. The risk premium is the difference between the expected market-wide return and the risk-free rate and represents the excess return an investment is expected to yield. The riskier the investment, the greater the risk premium.

The weighted average cost of capital (WACC) provides a company's average cost of raising capital based on its proportion of debt and equity. WACC can be calculated as:

$$WACC = Cost\ of\ Equity\ \times \frac{Equity}{Equity+Debt} + Required\ Return\ \times (1\ - corporate\ tax\ rate) \times \frac{Debt}{Equity+Debt}$$

To find the total capitalized cost and account for failed biomarker attempts, the out-of-pocket costs required for biomarker development is first multiplied by the cost of capital to calculate capitalized costs. This is calculated as:

[22] Net Realized Direct Cost of Failure: The actual cost of failure based on the specific biomarker investment, less any residual benefit gained.

$$Capitalized\ Biomarker\ Development\ and\ Validation\ Cost = Biomarker\ Investment_{Y1} + Biomarker\ Investment_{Y2}(1 + Cost\ of\ Capital) + \cdots + Biomarker\ Investment_{Yn}(1 + Cost\ of\ Capital)^{n-1}$$

Then, capitalized costs are divided by the average overall probability of failure. This figure provides an estimate of the full out-of-pocket and time costs incurred to develop and validate a biomarker to the point of marketing approval, accounting for failed attempts. This is calculated as:

$$Total\ Cost\ for\ One\ Successful\ Biomarker = \frac{Capitalized\ Biomarker\ Development\ and\ Validation\ Cost}{Overall\ Probabilty\ of\ Failure}$$

4 RESULTS

4.1 Defining Biomarker Categories

RESEARCH QUESTION 1

How are biomarkers identified for use in drug development?

Four categories of biomarkers are included in this report: predictive biomarkers; prognostic biomarkers; safety biomarkers; and surrogate endpoints.

4.1.1 Predictive Biomarkers

DEFINITION

Predictive biomarkers are used to identify individuals who are more likely than similar individuals without the biomarker to experience a favorable or unfavorable effect from a specific intervention or exposure.

The most common application for predictive biomarkers is in oncology, where significant heterogeneity in disease and response to therapy is frequently observed. Predictive biomarkers are most frequently used in drug development to identify subpopulations of patients with a given disease that are most likely to benefit from therapy. They are used in clinical trials either to enrich for individuals likely to respond to therapy or to exclude individuals unlikely to benefit from therapy. In general, biomarkers in this category are most useful in situations where there is reason to believe a drug will have good efficacy in a subgroup of the patient population and/or the potential for serious adverse effects for a patient subgroup, and that these subgroups can be identified by a distinct biomarker. Predictive biomarkers can also be used to tailor and optimize dosing strategy by identifying individuals that are likely to metabolize or respond to a therapy differently. For example, the companion diagnostic for the cancer drug crizotinib identifies those patients whose metastatic non-small cell lung cancer (NSCLC) is positive for the fusion protein, anaplastic lymphoma kinase (ALK), for these are the patients whose cancer will respond to crizotinib treatment (Pfizer Inc., 2011).

ALK AND CRIZOTINIB

Biomarkers can significantly change the trajectory of drug development. The cancer drug crizotinib was initially developed as a Mesenchymal-Epithelial Transition (MET) tyrosine kinase inhibitor. One of the crizotinib clinical sites, which was in the process of developing an anaplastic lymphoma kinase (ALK) rearrangement assay, identified an ALK mutation among one of the few trial participants who had a dramatic response to crizotinib. This serendipitous discovery, coupled with the understanding of how ALK rearrangement is involved in the pathogenesis of malignancies, enabled investigators to develop ALK rearrangement as a predictive biomarker. Crizotinib went on to receive conditional approval based on response rates of 50 percent and 61 percent in two single arm trials of ALK-rearranged non-small cell lung cancer patients. Without the discovery of ALK rearrangement as a biomarker that predicts response to crizotinib therapy, this drug would likely have never received approval (Ou, 2011).

Experts noted that identification of predictive biomarkers is largely dependent on basic research and the scientific knowledge of the molecular mechanisms underlying the disease state. Often, the development of predictive biomarkers at the early stages of drug development is limited by the lack of a clear understanding of the mechanism of action for a drug or a clear biomarker hypothesis. Predictive biomarkers are most useful and easiest to validate when there is a strong association between the biomarker and the response rate to a therapy. It is much more difficult to develop a biomarker when multiple factors contribute to the response rate, each with a small impact.

Despite the challenges in developing predictive biomarkers, they have shown tremendous benefit in terms of increasing the success rate of drug development. Between 2006 and 2015, clinical trials utilizing a predictive biomarker had a three-fold higher likelihood of a new drug approval as compared to those that did not (Thomas et al., 2016). Predictive biomarkers can also help to accelerate clinical trial enrollment, both by reducing the number of participants needed to demonstrate a statistically significant response and by providing individuals with a rationale for

enrolling in that particular trial, a factor that experts noted is especially important when there are multiple clinical trials for a potential participant to select from.

4.1.2 Prognostic Biomarkers

DEFINITION
Prognostic biomarkers are used to identify the likelihood of a clinical event, disease recurrence or progression in patients who have the disease or medical condition of interest.

Prognostic biomarkers provide information about the trajectory of a disease, for instance, aggressiveness or risk of recurrence. Prognostic biomarkers can be used as eligibility criteria in clinical trials to identify patients who are more likely to have clinical events or disease progression. Experts noted that prognostic biomarker technologies in development, such as liquid biopsies for circulating tumor cells, have the potential to identify patients who may experience a relapse of their disease at a much earlier timeframe. Prognostic biomarkers are particularly important for diseases that show great variability in the disease course across patient populations and/or when there is significant disease heterogeneity. Such biomarkers can be used to guide treatment decisions for patients, and in clinical trials they can segment patients to different treatment arms or enrich studies for patients with a common prognosis.

Prognostic biomarkers are often identified through hypothesis-generating research, using microarrays or proteomic screening assays, and employ algorithms that can illuminate differential expression patterns of genes or proteins that predict the presence or severity of a disease or condition. Unfortunately, variability in disease outcomes is often due to multiple factors and complex interactions in disease pathology. This complexity can potentially be addressed by pursuing a genomics or proteomics approach. One example of a proteomics approach is a recent biomarker study that used serum proteomics via mass spectroscopy to try to define patterns that were prognostic of either benign or aggressive forms of multiple sclerosis. (Butterworth et al., 2016). Where a clear relationship has been defined between individual biomarkers and disease progression, simpler prognostic tests can be developed. Rheumatoid factor (RF) and anti-cyclic citrulinated proteins (anti-CCP) are prognostic biomarkers for rheumatoid arthritis that can be measured through simple blood tests (Taylor et al., 2011).

Unlike predictive biomarkers, prognostic biomarkers are not intended to identify individuals who are most likely to respond to a specific drug therapy or class of therapies with a similar mechanism of action; rather, they provide information about the likely progression or severity of a disease. Prognostic biomarkers typically play an informative role helping to guide treatment decisions but are less frequently used in drug development. Examples of prognostic biomarkers used in oncology drug development in particular are exceedingly rare. In fact, experts noted that one of the most well-known and validated prognostic biomarkers in the oncology space, Oncotype DX® Breast Cancer, has never been used in drug development for patient segmentation in a clinical trial. According to experts, a prognostic biomarker actually has the potential to limit the population that will receive a drug without the ability to identify patients more likely to benefit from a specific treatment, as is the case with companion diagnostic tests. As a result, prognostic tests do not reduce the uncertainty associated with a new drug treatment, making it more difficult for manufacturers to achieve the premium pricing for their drugs that would offset the reductions in the patient populations that can be treated with their drugs; this can be a significant disincentive for drug manufacturers to pursue the development of prognostic tests.

4.1.3 Safety Biomarkers

DEFINITION

Safety biomarkers are used to indicate the presence or extent of toxicity related to an intervention or exposure.

Safety biomarkers play an important role in drug development. A recent portfolio review by a major pharmaceutical company identified safety failures as the top reason that drug development programs fail, accounting for more than half of all failures (Cook et al., 2014). Safety biomarkers are used to indicate the presence or extent of toxicity related to an intervention. They can help improve the drug development process by:

(1) Enabling researchers to identify safety concerns earlier so that therapeutic candidates with an unacceptable safety profile can be terminated more rapidly and with less investment;

(2) Providing insight into the mechanisms of toxicity/pathology, enabling the design and development of safer therapeutics; or

(3) Enabling researchers to monitor the extent or severity of damage, better differentiate levels of risk, and discontinue therapeutic use at a stage when the damage is still reversible.

CROSS-OMICS ANALYSIS TO IDENTIFY NEW SAFETY BIOMARKERS

The European InnoMed-PredTox was a collaborative project aimed at identifying mechanism-linked safety biomarkers of liver and kidney toxicity. The project combined conventional toxicology parameters and -omics data, including microarray and proteomics data to identify novel biomarkers. The project confirmed three previously described biomarkers and identified several new pathways involved in organ toxicity. One of the key findings was that the identification of new and more specific biomarker candidates for organ toxicity was most successful when combining multiple types of -omics data with traditional toxicology data. The collaborative nature of this project, which included contributors from 15 pharmaceutical companies, two enterprises and three universities, was critical to bringing together the skills and expertise needed to perform the toxicology and -omics studies (Matheis et al., 2011).

The identification of new safety biomarkers is often very practical and needs-driven. The top organs or organ systems involved in safety failures, including cardiovascular, central nervous system, liver, kidneys and musculoskeletal (Cook et al., 2014), are the areas where safety biomarker identification and development efforts are most prolific. Typically, safety biomarkers are not tied to a specific drug or class of therapeutics but to an organ or organ system. Since safety biomarkers are not co-developed for a specific therapeutic, the experts consulted confirmed that consortia are needed to develop new safety biomarkers. Some safety biomarkers, such as atrial natriuretic peptide, are developed based on an understanding of the molecular pathways involved in toxicity and pathology while others, such as osteoactivin (Matheis et al., 2011), are identified using a screening method to identify differential expression of genes or proteins that are indicative of toxicity. Among the categories of biomarkers included in this effort, safety biomarkers are unique in that there is a need for both pre-clinical and clinical biomarkers. Typically, safety biomarker development and validation begins in animal models as a preclinical biomarker and in some cases progresses toward validation as a clinical biomarker. The animal model selected for safety biomarker development is often driven by the degree of homology between the biomarker in animals and humans as well as the extent of the similarities between the toxicity/pathological pathways in animals and human.

4.1.4 Surrogate Endpoints

DEFINITION
Surrogate endpoints are used in clinical trials as a substitute for a direct measure of how the patient feels, functions or survives; they do not measure the clinical benefit of primary interest in and of itself, but rather, are expected to predict the clinical benefit or harm based on epidemiological, therapeutic, pathophysiological or other scientific evidence.

Surrogate endpoints are biomarkers that are used as a substitute for clinical endpoints that predict the clinical benefit or harm from a therapeutic intervention. They are most commonly used in phase III clinical trials to support claims of efficacy of new therapeutics. Surrogate endpoints enable faster trials by reducing the amount of time needed for clinical follow-up. In general, surrogate endpoints are most useful when:

- The clinical endpoint takes years to develop;
- The surrogate endpoint seems to be obviously linked to the clinical endpoint of interest (e.g., tumor size in cancer);
- When other treatments exist, to alleviate difficulties of conducting trials when a new intervention must be proven as non-inferior to existing treatments; and/or
- When studying a population that is relatively at lower risk for a condition where using the primary clinical endpoint might be too burdensome (Institute of Medicine Committee on Qualification of Biomarkers and Surrogate Endpoints in Chronic Disease, 2010).

There is widespread consensus, both among experts and in the research literature, that surrogate endpoints are the most difficult category of biomarkers to identify, develop, and validate. Biomarker development and validation is guided by the principle of being linked to how they will be used (e.g., fit-for-purpose). In light of the serious consequences attached to an unreliable or inaccurate surrogate endpoint, the level of evidence required to validate a surrogate endpoint is much higher than for other biomarker categories such as prognostic biomarkers or predictive biomarkers. Understanding of the underlying disease mechanisms and pathology is critical for the development of surrogate endpoints. Unless the disease mechanisms are very well understood, it is not possible to be confident that a surrogate endpoint can predict all of an intervention's effects. In addition to understanding the underlying disease pathology, it is also critical to understand the role that a surrogate endpoint plays in the disease and therapeutic interventions. As experts noted, without this scientific knowledge base it is not possible to determine whether changes in a surrogate endpoint always correspond with the clinical endpoint or whether the surrogate is able to capture the entire effect of the intervention on the clinical endpoint.

Surrogate endpoints are commonly used in infectious disease where viral load is often used as a surrogate endpoint for treatment efficacy. HIV-1 RNA is one classic example of a surrogate endpoint that was used to accelerate the development and approval of HIV drugs during the AIDS epidemic in the 1990s (Institute of Medicine Committee on Qualification of Biomarkers and Surrogate Endpoints in Chronic Disease, 2010). Cardiovascular disease is another area where surrogate endpoints, such as low-density lipoprotein cholesterol (LDL-C) and blood pressure, are routinely used to assess the efficacy of new therapeutics. Surrogate endpoints are particularly useful in cardiovascular disease, because the clinical endpoints, such as myocardial infarction and stroke, are relatively rare and take a long time to develop. Other disease areas where there is a strong interest in using surrogate endpoints include cancer, degenerative joint diseases (e.g., osteoarthritis), metabolic disorders (e.g., diabetes), rare diseases such as muscular dystrophy and, to a lesser extent, neurodegenerative disorders.

4.2 Biomarker Applicability

RESEARCH QUESTION 1A

How do these processes differ depending upon the category of biomarker?

As summarized in Exhibit 6, biomarkers span eleven broad disease areas. Predictive biomarkers were the most prevalent type and cancer is the most common disease area for predictive biomarkers. As shown in Exhibit 6, in general, the disease states tended to cluster into one or two biomarker categories. Predictive biomarkers were most common in disease states with significant heterogeneity in response to therapeutics, such as cancer, to help identify sub-segments of the population most or least likely to benefit from treatment. Prognostic biomarkers that correlate with milder or more severe disease courses have been studied in several autoimmune indications as well as for a few cancer indications. Safety biomarkers tended to align with an organ or organ system with the potential to work across disease states while surrogate endpoints were most common in long-term, chronic, and infectious diseases.

Exhibit 6: Alignment of Biomarker Categories to Disease

Disease	Predictive Biomarkers	Prognostic Biomarkers	Safety Biomarkers	Surrogate Endpoints
Cancer	+++	++	–	+
Cardiovascular Disease	++	++	++	+
Degenerative Joint Diseases	–	++	–	++
Immune System Disorders	++	++	–	++
Infectious Diseases	–	–	–	+++
Kidney Disease	++	-	++	++
Liver Disease	–	–	+++	–
Metabolic/Endocrine Disorders	–	–	–	+++
Neurodegenerative Diseases	+	+++	–	+
Rare Diseases	++	++	–	++
Respiratory Diseases	–	+++	–	–
Skeletal Muscle Disorders	–	–	++	++

Legend: – few to no examples identified; + < roughly 20 percent of the biomarkers surveyed; ++ approximately 20-50 percent of the biomarkers surveyed; +++ >50 percent of the biomarkers surveyed for the disease area

Importantly, some biomarkers fit into more than one category or spanned more than one disease area. Most notably, some biomarkers initially developed as prognostic biomarkers were later shown to predict response to a particular therapy or intervention. For example, Oncotype DX® Breast Cancer was developed as a prognostic biomarker to assess the likelihood of disease recurrence in node negative, estrogen receptor positive breast cancers but was later shown to serve as a predictive biomarker that could identify which individuals were most likely to benefit from adjuvant chemotherapy.

The most common uses of the biomarker categories are listed in Exhibit 7. The common uses varied largely as a function of category with minimal overlap between categories. As with the classification of biomarkers, predictive biomarkers had the greatest overlap with prognostic biomarkers. While predictive biomarkers are generally used to predict which patients will have the greatest or least chance of responding to a specific therapy or class of therapies, prognostic biomarkers can also be used to enrich clinical trial design by identifying the group of individuals most likely to respond to therapeutic intervention based on their disease status or prognosis. For example, as stated above, the presence or absence of the serum biomarkers RF and/or anti-CCP have been shown to define four different rheumatoid arthritis patient

populations with prognoses of mild (negative for both biomarkers), intermediate (positive for one biomarker) and severe (positive for both biomarkers) (Taylor et al., 2011). These biomarkers could be used in clinical trials either to stratify patients according to prognosis of disease severity or to select/enrich for patients with a particular prognosis (e.g., only patients with a prognosis of severe disease).

Another example where the use of biomarkers is similar across categories is to reduce the number of trial participants needed to detect a statistically significant treatment effect. Predictive biomarkers can reduce the number of individuals that need to be included in the trial to observe a statistically significant treatment effect by either enriching for the individuals most likely to benefit from treatment or excluding those individuals least likely to experience therapeutic benefit. Similarly, surrogate endpoints can also reduce the number of individuals needed to demonstrate a statistically significant clinical benefit. Often, surrogate endpoints are used when the true clinical endpoint is relatively rare and/or requires a long follow-up period to observe. Surrogate endpoints can reduce trial size by assessing treatment response in a shorter timeframe, reducing the loss of participants to attrition. Surrogate endpoints can also monitor clinical events that are observed more frequently, reducing the sample size needed to observe a statistically significant event.

Exhibit 7: Common Uses of Biomarkers by Biomarker Category

	Predictive Biomarkers	Prognostic Biomarkers	Safety Biomarkers	Surrogate Endpoints
Identify patients most likely to benefit from therapy	✓			
Identify patients most likely to experience an adverse event	✓			
Identify patients most likely or at highest risk to develop a disease		✓		
Predict the speed or severity of disease progression		✓		
Detect treatment effects more rapidly than traditional clinical outcomes				✓
Reduce the number of trial participants needed to detect statistically significant treatment effects	✓	✓		✓
Detect toxicity or safety concerns earlier than existing clinical indicators			✓	
Monitor for safety concerns at a point at which it the damage is still reversible			✓	
Provide insight into the mechanism of action underlying toxicity	✓		✓	

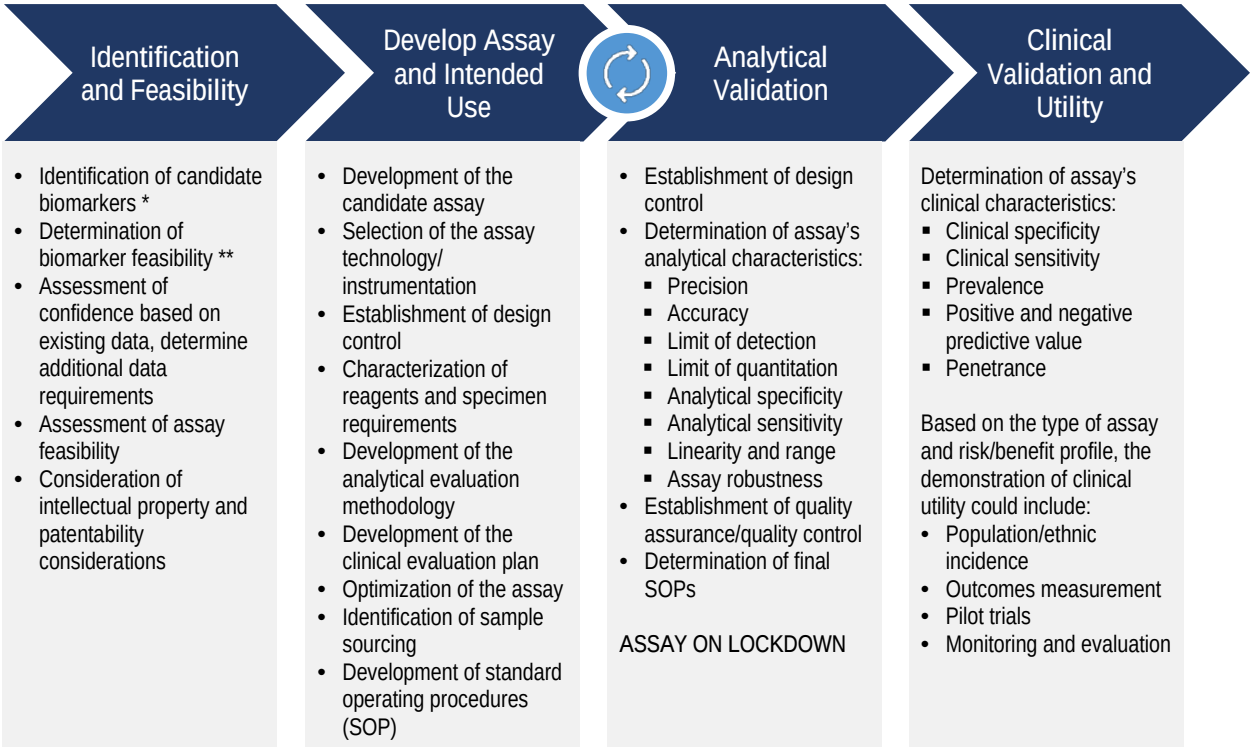
4.3 The Biomarker Development and Validation Process

Understanding the biomarker development and validation process is critical to developing an understanding of biomarker development and validation costs. Most biomarkers are identified by academia through basic biomedical research and -omics-based screening assays. Private or consortium-based organizations generally shepherd biomarkers through later development and validation phases. The sections below provide a summary of the development and validation process (Exhibit 8) and describe each phase of the process to inform and provide background for the ensuing discussion of factors influencing costs at each individual phase. As the biomarker development and validation process advances from phase to phase, "fit-for-purpose" must be accounted for, where the level of validation data associated with the biomarker is sufficient to support its context of use and allows researchers to tailor the validation process to the intended use.

4.3.1 Phases of the Biomarker Development and Validation Process

Exhibit 8, below, outlines the key steps in biomarker identification, development and validation. The process is largely the same for all biomarker categories.

Exhibit 8: Biomarker Identification, Development and Validation Lifecycle



4.3.1.1 Identification and Feasibility

The first step in the development of new biomarkers for drug development is to identify a potential biomarker and determine its feasibility of use. Ideally, it should be done as early as possible, during the preclinical stages or phase I of drug development when developing a new drug candidate. During this stage, candidate biomarkers are identified and the feasibility of using the biomarker in drug development is assessed. Assessing the feasibility of the biomarker involves getting a better sense of the degree to which the candidate biomarker is indicative of the processes or endpoints it is intended to measure, using preclinical animal models and clinical samples when available. Another important aspect of assessing the biomarker feasibility is establishing the extent to which there is confidence in a biomarker for its intended use. The process of building confidence that a biomarker is truly indicative of the process or endpoint it is intended to predict often involves surveying existing data and determining the additional data that are needed to build confidence in the biomarker prior to undertaking development and validation work. Once a biomarker has been selected, it is critical to determine whether an assay can be developed to measure the biomarker with the accuracy and precision needed for its intended use, a process that includes assessing the resources available for assay development, such as antibodies or well-characterized reference standards. Lastly, legal and intellectual property considerations need to be reviewed during this phase to ensure that there are no potential legal restrictions on the use of the biomarker or assay being developed.

4.3.1.2 Develop Assay and Intended Use

After the Identification and Feasibility phase is complete, the second step is to use the information gathered to develop the intended use and assay. Developing the intended use is the process of defining the clinical circumstances or purpose for which a biomarker is being developed. Based on the intended use, the analytical and clinical validation methodology are often defined during this stage, including the target values and acceptance limits. Using the information gathered during the Identification and Feasibility phase and the intended use, a candidate assay is developed using the most appropriate technology platform. Assays are developed using design control methodology which includes a design history file that provides evidence that the development of a biomarker meets the regulatory requirements of its intended use. The reagents used to perform the assay and specimen requirements for the assay are also defined at this stage. During this phase, the assay will often undergo a series of optimization experiments designed to evaluate and improve assay performance under a range of applicable matrices, samples, and assay conditions. After the assay is optimized, standard operating procedures and criteria for accepting or rejecting assay batches are developed to ensure the assay is reproducible. Lastly, planning for the analytical and clinical validation phases begins during this phase. The types of planning that are performed during this stage include developing a sample sourcing plan to ensure there are a sufficient number of the correct type of samples needed for the validation efforts and developing a clinical evaluation plan that demonstrates the clinical benefits of a biomarker assay outweigh the potential risks.

4.3.1.3 Analytical Validation

Analytical validation is the process of establishing that the performance characteristics of a biomarker assay are acceptable in terms of its sensitivity, specificity, accuracy, precision, and other relevant performance characteristics using a specified technical protocol (FDA-NIH Biomarker Working Group, 2016). In addition to establishing these performance characteristics, the analytical validation process includes assessment of assay robustness and development of the final standard operating procedure that will be used to train technicians performing the assay. In the case of omics-type assays where the interpretation of the analyte measurements is based on a software algorithm, both the assay measuring the analytes and the software algorithm interpreting the results need to be validated. Typically, the analytical validation process involves multiple iterations, with the assay being refined as additional information is gathered about the assay's performance. In parallel with the analytical validation, specimen stability and reagent integrity are assessed, and quality assurance and quality control steps are developed. After the analytical validation is complete, the assay is in 'lockdown;' that is, no changes can be made to the assay or the protocol. Any changes that need to be made to either the assay or the protocol after the analytical validation will require bridging studies that compare the performance of the original, analytically validated assay to the new assay to ensure that the changes do not adversely impact the assay's analytical performance.

4.3.1.4 Clinical Validation and Utility

Clinical validation is the process of connecting the biomarker response to the biological change that it is intended to measure. Clinical utility is the conclusion that use of a biomarker will lead to a net improvement in health outcome or provide useful information about diagnosis, treatment, management, or prevention of a disease (FDA-NIH Biomarker Working Group, 2016). The level of data required to demonstrate clinical validity and clinical utility varies depending upon the intended use of the biomarker. Prospective, randomized control trials (RCTs) are the gold standard for demonstrating clinical validity and utility. However, due to the time and expense involved in conducting RCTs, some types of biomarkers are validated with "retrospective-prospective" RCTs. While clinical validation and clinical utility are distinct

concepts, they are usually assessed concomitantly for biomarkers through the same trial or series of trials. Therefore, they are considered as a single step in the biomarker development and validation lifecycle, referred to in this report as the Clinical Validation and Utility phase.

4.3.1.5 Data Sources

To examine costs associated with biomarker development and validation, assessing data requirements necessary for regulatory acceptance is critical. Project findings indicated that overall regulatory data requirements include a well-defined intended use that clearly states the clinical circumstances or purpose for which a biomarker is being developed. Assay validation is the first process step where data is collected for regulatory acceptance. The type of data needed to validate the analytical performance of the assay depends on the type of assay selected. Examples of assay validation data needed for regulatory acceptance include:

RESEARCH QUESTION 2
Given that regulatory acceptance of a biomarker is a data-driven process, what are the sources of data necessary for biomarker development?

- Definitive quantitative biomarker assays require data that demonstrate the precision, accuracy, sensitivity, dynamic range, lower limit of quantification, and upper limit of quantification;
- Relative quantitative biomarker assays require data that demonstrate the precision and bias of the assay as well as data that assess the dilution linearity and parallelism between the endogenous biomarker and assay calibrator;
- Quasi-quantitative biomarker assays require data that demonstrate the clinical specificity and sensitivity using positive and negative controls;
- Qualitative biomarker assays require data demonstrating the specificity and consistency of the assay. If the assay relies upon scoring by an expert, data demonstrating the reliability of scoring between experts need to be present; and
- Data demonstrating sample stability under the collection and storage collections specified in the assay are important regardless of the assay type.

Clinical Validation and Utility data is also needed for regulatory acceptance. For prognostic biomarkers, predictive biomarkers and safety biomarkers, data from at least two well-designed retrospective-prospective RCTs demonstrating clinical validity and utility or one well-designed prospective RCT are generally required. For surrogate endpoints, data from multiple well-designed prospective RCTs demonstrating clinical validity and utility of the surrogate endpoint are generally required.

4.3.2 Timeframe and Drivers of the Development and Validation Process

Different categories of biomarkers tend to share common development and validation process phases as well as process drivers (Exhibit 9), therefore the process and timeline to develop and validate biomarkers is similar across biomarker categories, with a few differences.

Exhibit 9: Timeframe and Process Drivers for Biomarker Development

	Identification and Feasibility	Develop Assay and Intended Use	Analytical Validation	Clinical Validation and Utility
Predictive biomarkers	1 month to 3 years[23]	3 months to 1.5 years	9 months to 1.5 years	10 months to 3 years
Prognostic biomarkers	3 months – 1 year	6 months – 2 years	1.5 years	1.5 – 10 years
Safety biomarker	3 months – 1 year	3 months – 6 years	6 months – 1 year	6 months – 6 years
Surrogate endpoints	1-5 years	3 months to 3 years	1-2 years	1-10 years
Process drivers (defined as factors that impact the cost, duration, and likelihood of success by phase)	• Breadth, depth of basic research • Strength of association between the biomarker and outcome • Complexity of intellectual property and legal considerations	• Familiarity with technology • Availability of key reagents • Complexity of contracting if using a CRO	• Single analyte versus multiplex • Number of iterations or training sets needed	• Time needed to observe clinical outcome and link to biomarker • Availability of sample • Prevalence of biomarker in population of interest • Need for algorithm or expert to interpret assay • Planning for bridge studies, if needed

4.3.2.1 Timeframe

Two key differences were noted for development and validation timelines. The last phase of development, Clinical Validation and Utility, has the potential to be much longer than earlier phases for number of reasons, including time needed to observe clinical outcomes. Relative to other biomarker categories, the development timeline for prognostic biomarkers and surrogate endpoints is long, in part because outcome measures can have a long follow-up period.

4.3.2.2 Process Drivers

Process drivers are factors that impact the cost, duration, and likelihood of success across each development and validation phase.

4.3.2.2.1 Biomarker Identification and Feasibility Process Drivers

Depending upon the biomarker, the Identification and Feasibility phase can take anywhere from one month to five years to complete. The factors that can influence the amount of time it takes to complete the Identification and Feasibility phase include:

- The breadth and depth of basic science knowledge surrounding a biomarker;
- The strength of the association between the biomarker and the process or outcome it measures/predicts; and
- Level of complexity involved in the legal and intellectual considerations.

In general, experts indicated that the Identification and Feasibility phase is similar across the biomarker categories but on average may require less time in prognostic and safety biomarkers than in predictive biomarkers and surrogate endpoints. This can be explained at least in part by the fact that understanding the basic biology of the biomarker, one of the key driving factors influencing the amount of time needed to

[23] Time ranges consist of the minimum and maximum collected during both the environmental scan and the case studies. No one biomarker took the shortest or longest possible time across all four developmental stages.

identify and assess feasibility of a potential biomarker, is much more important for predictive biomarkers and surrogate endpoints, as discussed in Sections 4.1.1 and 4.1.4, respectively.

4.3.2.2.2 Develop Assay and Intended Use Process Drivers

There was consensus among experts that developing the assay and intended use typically takes between one month and six years. Variability in the amount of time needed to develop the assay and intended use is less driven by the biomarker category than other steps. Process drivers for this stage include the availability of key reagents, familiarity with the assay technology and complexity of the contracting process if using a Contract Research Organization (CRO). One expert noted that developing a new antibody can add six to nine months to the amount of time needed to develop the assay. Another expert reported that it took their organization two years to optimize a polymerase chain reaction (PCR) protocol for use in formaldehyde-fixed, paraffin-embedded tissues. However, now that the protocol has been developed and optimized, their organization can develop PCR assays for use in tissues stored under these conditions much more rapidly. CRO contracting issues that can increase negotiation time and delay project start dates predominantly involve intellectual property ownership/freedom to operate, ownership of developed assays, and liability/indemnification clauses in contracts. Budgetary items are a further cause of negotiation delay as a CRO may seek to have risk taken on by the partner. However, CRO organizations experienced in biomarker development can provide savings in both time and cost for drug manufacturers that lack such capabilities internally.

4.3.2.2.3 Analytical Validation Process Drivers

In general, the amount of time needed to perform the Analytical Validation phase can range from six months to two years. As with the prior phase, the amount of time needed for analytical validation is similar across all four biomarker categories. Variability in the amount of time needed for analytical validation as well as the associated cost is driven by two main factors:

(1) The number of iterations needed to optimize and validate the analytical performance of the assay; and

(2) The number of analytes involved in the biomarker assay (e.g., single analyte versus multiplexed).

Experts were quick to highlight the iterative nature of analytical validation. The assay initially developed often needs additional work to optimize the analytical performance such that it can provide the sensitivity, specificity, and reliability needed for the intended use. Often, the need for improvements to optimize the assay's performance does not become apparent until the assay begins to undergo analytical validation. For example, during assay validation it can become apparent that signal interference from other molecules in the sample, matrix effects, or the relative abundance of the analyte compared with other molecules in the sample are preventing the assay from achieving the needed performance profile, which would necessitate additional rounds of optimization. Any changes made to the assay or protocol necessitate repeating earlier steps in the validation process to ensure that changes made to impact one aspect of the assay's performance do not adversely impact other aspects of the assay's performance. The number of cycles needed for any given step will vary in large part based on the technology platform selected for the assay.

For example, experts indicated that an immunohistochemistry (IHC) assay may go through several iterations to optimize for factors such as the antibody concentration, antigen retrieval method, and incubation times. In contrast, PCR-based assays benefit from probe design software that yields better results early on and relatively standardized pre-validation steps. As a result, experts noted that PCR-based assays typically require fewer iterations to optimize than IHC assays. Regardless of the technology, though,

the analytical validation benefits from performing a series of experiments to optimize the assay conditions prior to the analytical validation experiments.

It is common, but highly undesirable, to have changes to the assay even after the clinical validation phase starts. If any changes are made to the assay, bridging studies need to be done comparing the performance of the original, analytically validated assay with the new assay to ensure that the changes do not adversely impact any aspects of the assay's analytical performance. The need to run bridging studies is costly in terms of additional time, expense and samples. Bridging studies can be problematic as the same clinical sample used in the original clinical validation needs to be used in the bridging study. Thus, plans to collect sufficient clinical samples for possible bridging studies need to be made early.

The number of analytes involved in the biomarker assay impact the amount of time needed for analytical validation. Experts noted that biomarker assays are increasingly moving from single analyte to multiplexed panels that rely on algorithms to analyze the cumulative effect of changes in multiple analytes. With these multiplexed panels, the analytical validation needs to be performed for each analyte to demonstrate that each one is being measured in an accurate and reproducible manner. Some multiplexed panels can include upwards of 50 analytes (Myers, 2016), each of which must undergo validation. Thus, the number of analytes included in the assay can have a major impact on the amount of time needed to validate the assay.

4.3.2.2.4 Clinical Validity and Utility Process Drivers

The amount of time needed for the demonstration of clinical validation and utility can vary considerably, from six months to 10 years. Experts noted that, in general, the timeline is longest for surrogate endpoints and prognostic biomarkers because of the length of time needed to assess the clinical outcome that the biomarker is intended to predict. For example, studies of prognostic biomarkers in Alzheimer's and other neurodegenerative diseases require durations of several years owing to the length of time needed to observe whether or not individuals with mild cognitive impairment would convert to fully developed Alzheimer's Disease (Adaszewski et al., 2013; Huang et al., 2017). Likewise, statin trials such as the West of Scotland Coronary Prevention Study Group, which were particularly helpful in providing support for the use of LDL-C as a biomarker for cardiovascular disease (Institute of Medicine Committee on Qualification of Biomarkers and Surrogate Endpoints in Chronic Disease, 2010), followed patients for nearly five years to track the incidence of myocardial infarction and death from cardiovascular causes (Shepherd et al., 1995). Such lengthy trials are quite expensive given the need for extensive patient follow-up; the need to power trials sufficiently to account for potential patient drop out during the course of the studies also adds to the expense of longer-duration studies.

In addition to the length of time needed to observe the clinical outcome of interest, there are a variety of other factors that can influence the amount of time needed to demonstrate the clinical validity and utility of a biomarker, such as the availability of samples for retrospective-prospective studies. For example, experts estimated that the clinical validation of Oncotype DX® Breast Cancer took approximately 1.5 years. The reason the clinical validation could be performed so quickly was that samples linked to clinical outcome data were available from previously conducted seminal clinical trials in early stage breast cancer. Had those samples not been available, the trials, which assessed the risk of recurrence over 10 years, would have taken more than 15 years. Likewise, there is significant variability in the amount of time needed to validate safety biomarkers because the trials need to be planned in conjunction with other trials where toxicity is a concern. If archived samples are available to validate safety biomarkers, the process can proceed fairly rapidly but if the validation has to wait for a clinical trial to launch the process will take much longer.

The relative abundance of the biomarker within the patient population is another important factor that influences the amount of time needed for the clinical validation of a biomarker. Patient recruitment is frequently the lengthiest step in a clinical trial (Emanuel et al., 2003). As discussed in Section 4.4.3, the prevalence of the biomarker in a population drives the number of patients that will need to be screened for inclusion in a potential trial examining the relationship between the biomarker and the clinical outcome of interest. As such, the prevalence of the biomarker can impact the amount of time needed to recruit a sufficient number of trial participants and the associated cost, and thus the overall amount of time and budget needed to clinically validate the biomarker.

There are also a number of factors related to the assay itself that will impact the amount of time needed to validate a biomarker. The need for algorithms to interpret the results of the biomarker or technical experts to run the assay or interpret the assay results (e.g., the need for a pathologist to score an IHC) can increase the amount of time needed to perform the clinical validation. Additionally, during the clinical validation stage it can become apparent that the threshold for making treatment decisions needs to be revised or the algorithm that interprets the assay results needs to be refined. Any changes in the interpretation of the biomarker assay results during the clinical validation process will necessitate another round of validation and significantly add to the timeline. In many instances, investigators will plan for this possibility by structuring their sample collection such that there are samples available, as needed, for additional bridging studies in the future. Planning sample collection for future bridging studies is time consuming. The informed consent must be written in a way that allows for samples to be reused at a later date for bridging studies. Significant resources have to be identified and acquired to facilitate the long-term storage of clinical specimen. Experts were in agreement that the additional planning and preparation that goes into gathering samples for future bridging studies is nontrivial and can significantly impact the amount of time needed and the associated cost for clinical validation.

4.4 Cost Drivers

RESEARCH QUESTION 2B

What factors influence the cost of developing and validating biomarkers for use in drug development?

Analysis of cost drivers for this project focused on two biomarker categories: predictive biomarkers and surrogate endpoints. Findings indicated that seven cost drivers that influence the costs of development and validating new biomarkers in these two categories. As shown in Exhibit 10, five cost drivers significantly affect both biomarker categories (as defined in Section 3.3.1.2), and two cost drivers affect surrogate endpoints only. None of the cost drivers affected only the predictive biomarker category.

The purpose of this section is to provide a general overview of each cost driver and a high-level description of how it affects costs. It is important to note that not all cost drivers affect the costs of each phase of the development process, and the manner in which the cost driver affects costs may vary by phase. The impact each cost driver has on the costs of each developmental phase is discussed in detail in the cost framework (Sections 4.5.1 and 4.5.2). Lastly, this section also describes interactions among the cost drivers at a high level.

Exhibit 10: Summary of Cost Drivers

No	Cost Driver	Description	Categories Significantly Impacted by the Driver
1	Scientific Knowledge Base	Refers to the breadth, depth, and quality of data surrounding the role that the biomarker plays in health and disease	Predictive Biomarkers Surrogate Endpoints
2	Sample Type	The biological material(s) needed to measure the biomarker (e.g., blood, plasma, urine, etc.)	Predictive Biomarkers Surrogate Endpoints
3	Sample Size	The number of samples or participants required	Predictive Biomarkers Surrogate Endpoints
4	Novelty or Complexity of the Technology	Refers to the technology used to measure the biomarker and the degree of specialized equipment or reagents, amount of specialized labor required, and how common the technology is within the context of clinical care	Predictive Biomarkers Surrogate Endpoints
5	Trial Recruitment	The process of enrolling participants to a clinical study or trial	Predictive Biomarkers Surrogate Endpoints
6	Follow-Up Time Needed	The amount of time needed to observe the clinical outcome the biomarker predicts	Surrogate Endpoints
7	Development Model, Private or Consortium	Indicates whether a biomarker is developed by a private company or companies where the intellectual property is retained by the organizations that developed the biomarker or as a consortium in which multiple organizations contribute financial or in-kind resources to develop and validate the biomarker with shared or open intellectual property.	Surrogate Endpoints

4.4.1 Scientific Knowledge Base

The Scientific Knowledge Base refers to the breadth, depth, and quality of data surrounding the role that the biomarker plays in health and disease. It includes how well characterized the biomarker is in humans and relevant animal models; the dynamic range of the biomarker in health and disease; the impact of common co-morbidities on the biomarker; and the knowledge base surrounding the role that the biomarker plays in the pathophysiology/mechanisms of disease.

The breadth, depth, and quality of the scientific knowledge base for a biomarker are important considerations that can affect the cost and timeline for developing new biomarkers. The scientific knowledge base is particularly important during the Identification and Feasibility phase because it directly affects the amount of additional research needed to assess biomarker feasibility and begin to build confidence in its usefulness. The quality of the knowledge base for a biomarker is also critically important. A lack of high-quality, reproducible preclinical research is one of the key reasons biomarker development and validation efforts fail (van Gool et al., 2017). Several experts have called for quality improvement efforts across the biomarker development lifecycle, including measures such as developing best practice guidelines or widely accessible, high-quality reference materials and biospecimens (Freedman et al, 2015; Piccoli & Sauer, 2017; van Gool et al., 2017).

4.4.2 Sample Type

Sample Type refers to the type of biological material (e.g., blood, urine, saliva, etc.) used to measure the biomarker. There was strong consensus among experts that the most significant cost driver of biomarker development and validation is the cost of acquiring clinical samples for clinical validation. Sample Type is an important cost driver that affects the costs of acquiring clinical samples, both during the Clinical Validation and Utility phase and earlier phases of development. The costs associated with collecting or acquiring samples for clinical validation can vary dramatically depending on the degree of difficulty

associated with obtaining the sample and the commonness of the disease or condition. Experts agreed that $65 to $200 per sample is a realistic cost to acquire samples that are easy to collect, such as serum, saliva, or urine, in common, chronic conditions like diabetes or cardiovascular disease. The most common price mentioned for an easy-to-collect sample was $100 per sample. For samples that are more difficult to collect, such as tumor biopsies or cerebrospinal fluid (CSF), or in less common diseases such as chronic myelogenous leukemia, experts noted that the costs are typically more like $1,000 to $3,000 per sample. Experts also noted that for exceedingly rare conditions, such as a blood sample in the case of septic shock or biopsies from individuals with inflammatory bowel disease, samples can cost between $5,000 and $10,000 each.

4.4.3 Sample Size

Sample Size is a major cost factor that influences the time and resources necessary to develop and validate a new predictive biomarker. As with Sample Type, Sample Size has the greatest impact on those phases that require the greatest number of samples — the Clinical Validation and Utility phase and, to a lesser extent, the Analytical Validation phase. The number of samples needed to develop and validate new biomarkers can vary substantially because it depends on a variety of factors, including the disease, application, and intended use. The cost components that Sample Size most strongly affects are material costs (i.e., equipment and reagents),[24] with labor and overhead[25] affected to a lesser extent.

There are several factors that influence the sample size needed to develop and validate a biomarker. One factor is the strength of the association between the biomarker and the clinical outcome, which affects the sample size needed to observe a statistically significant event during the Clinical Validation and Utility phase, with weaker associations requiring more samples. The strength of the association between the biomarker and the clinical outcome also affects the Identification and Feasibility phase because, in general, fewer samples and less work are needed to identify strong associations than weak ones. Sample size is also impacted by the prevalence, or proportion of a population of interest positive for the biomarker. Biomarker Prevalence mainly impacts cost by increasing the number of samples needed to be screened to identify enough “biomarker positive” samples for the development and validation effort.

4.4.4 Novelty or Complexity of the Technology

The Novelty or Complexity of the Technology refers to the technology used to measure the biomarker. Examples of straightforward or routine technology for assessing biomarkers with a low level of novelty or complexity include immunohistochemistry and enzyme-linked immunosorbent assay, while examples of more novel or complex technologies include next generation sequencing (NGS) or fluorodeoxyglucose positron emission tomography (FDG-PET).

Novel or complex technologies increase costs in several ways. Experts noted that novel technologies are less automated and require more labor for analysis and interpretation. The assay development process can be more challenging and take longer when the technology is new, especially if it is primarily used in a research setting with few to no applications in a clinical care setting. When a technology is not routinely used in clinical care, the infrastructure needed to run the assay, such as equipment and reagents, is likely not available at hospitals and clinical laboratories. Laboratory and clinical personnel may require additional

[24] Equipment and Reagents: Cost components that include equipment maintenance and operational costs, reagents, storage, shipping and handling costs.

[25] Labor and Overhead: Cost component that includes all salaries and benefits; facilities and administrative support; and overhead, including resources such as legal and intellectual property.

training to run the assay or handle and prepare specimens properly. In addition, the equipment and reagent costs tend to be costlier for novel technologies.

Unlike many of the other cost drivers, the Novelty and Complexity of the Technology predominantly affects the earlier phases of biomarker development, including the Identification and Feasibility phase and the Develop Assay and Intended Use phase, as this is when the assay methodology is being developed. The technology affects the Labor and Overhead costs and the Material costs. In some instances, the Novelty and Complexity of the Technology can increase outsourced services costs if the technology or expertise needed is not available within the organization or organizations developing the biomarker. Some organizations may elect to outsource part or all of the validation work even when the technology and expertise is available in house, but not to the scale required for the validation effort.

While novel and complex technologies primarily impact the first two stages of biomarker development, novel technologies in particular can increase the costs of the analytical validation phase, as they require more rigorous and extensive testing and validations. Such was the case with the plasma-based epidermal growth factor receptor (EGFR) assay, which was the first assay to receive U.S. Food and Drug Administration (FDA) approval using circulating free DNA as the analyte. In many cases, the number and type of controls needed for analytical validation is unclear with novel technologies or applications, as there is no precedent that can serve as guidance. As a result, the developer often has to engage in lengthy discussions with the FDA or other regulatory bodies to establish sufficient controls. One of the themes echoed throughout the project was the principle of using the simplest and most routine technology platform possible when developing a biomarker, both to control costs and to keep the regulatory approval process as straightforward as possible.

Complex technologies also increase the cost and amount of time needed for biomarker development and validation. Complex technologies typically require substantial amounts of highly skilled labor, which increases the costs. Assays that require highly skilled labor to run, analyze, or interpret the results will be costlier across the entire biomarker development and validation lifecycle, though particularly in the Analytical Validation and Clinical Validation and Utility phases. Examples of complex technology include imaging biomarkers, where the estimated labor costs associated with the data analysis alone can be $500 per scan. Complex assays typically require equipment that is expensive to procure and maintain. Complex technologies also tend to generate large amounts of data that are costlier to transfer and store. For example, with an imaging biomarker like FDG-PET, the cost to transfer each scan to a core facility is between $200 and $500. One expert noted that setting up a database to house the results from studies using polymerase chain reaction assays cost their project approximately $200,000 at that time, whereas they also noted building a repository to house data from a single study using NGS can cost $50,000. Thus, the data transfer and storage fees associated with more complex technologies can be substantial and increase the cost of the overall biomarker development and validation effort.

4.4.5 Trial Recruitment

Trial Recruitment refers to factors related to recruiting patients for the biomarker validation. The process of recruiting participants to a clinical trial is often the lengthiest stage of a clinical trial (Emanuel et al., 2003). As such, it affects both the time and costs associated with the Clinical Validation and Utility Phase with minimal impact on the earlier stages. Several factors influence the time and costs needed to recruit sufficient trial participants. For low prevalence diseases, it can be more difficult and time-consuming, and therefore more expensive, to recruit sufficient trial participants. The number of trials recruiting individuals concurrently for a given disease or condition can affect the amount of time and resources needed to recruit trial participants. When there is more than one trial recruiting for a disease or condition in the same

geographical region, the studies compete for potential participants, making it even more challenging to recruit sufficient participants to the trial. Other factors that affect the time and costs required to recruit participants include the potential for therapeutic benefit for trial participants and the burden placed on participants. Biomarkers that are tested in a trial alongside a therapeutic product may find it easier to recruit participants because there is a potential benefit of therapy to the participants compared with those trials that test the biomarker in isolation where participation does not benefit the participant as an individual. Likewise, trials that are cumbersome for participants in the amount of time or travel required or are uncomfortable or painful for participants may find it more challenging and time-consuming to recruit the necessary number of participants.

4.4.6 Follow-Up Time Needed

The amount of follow-up time needed to observe the clinical outcome of interest during the Clinical Validation and Utility phase is an important cost driver. The cost impact is primarily the result of the increased labor needed to run a longer trial, but there can also be modest increases in material costs if the trial requires sample collection at multiple time points. The length of the trial or trials needed to clinically validate a biomarker is primarily driven by the follow-up time needed to observe the clinical outcome. There are several reasons that longer trials are costlier than shorter trials. First, longer trials require more staff time to follow and track clinical outcomes in participants. Second, trial infrastructure, such as equipment or databases, must be maintained for a longer period, creating additional material costs. Moreover, the longer the trial, the greater the risk that the trial's equipment or infrastructure will become outdated or obsolete, adding to the overall cost of the trial by necessitating the purchase of new equipment. Third, longer follow-up periods tend to lead to higher dropout rates among the participants which in turn decreases the validity of the data collected. Thus, the development plan should be designed accordingly to mitigate these risks. While both predictive biomarkers and surrogate endpoints collect clinical information over a period of time, follow-up time only reached the cost driver threshold (defined in Section 3.3.1.2) for surrogate endpoints.

4.4.7 Developmental Model (Private or Consortium)

Developmental Model (Private or Consortium) refers to whether a biomarker is developed by a private company or companies where the intellectual property is retained by the organizations that developed the biomarker or as a consortium in which multiple organizations contribute financial or in-kind resources to develop and validate the biomarker with shared or open intellectual property.

Collaboration through a consortium model can improve efficiency by reducing duplication of effort or redundancies and building consensus around the data and evidence needed to support the validation process at the outset, reducing the need to redo studies. However, these gains are often offset by the inefficiencies inherent in collaborative projects. Developing consensus takes time and becomes more challenging with larger or more diverse stakeholder groups. The time needed to develop consensus often delays the development and validation effort, lengthening the overall process. The incentive structure for collaborative processes can also lead to inefficiencies. In private projects, there is a large reward for one organization, whereas in collaborative projects, there is a smaller reward for multiple organizations. As such, collaborative projects are typically not among an organization's top priorities and must often wait for resources to become available. Collaborative projects tend to use the resources available, which may not be optimal or ideal. Last, use of government funding, which is not uncommon for collaborative efforts, can create delays. Experts noted that government funding is subject to additional rules and regulations surrounding the acquisition process that can delay the process. Taken together, the challenges of working

through a collaborative model increase the time and costs associated with biomarker development and validation efforts.

4.4.8 Interdependencies and Relationships Among Drivers

Interactions between cost drivers influence the extent to which any one cost driver influences costs. This section describes some of the most common ways cost drivers influence one another at a high level. This section is not exhaustive but serves to give examples of frequent interdependencies between drivers. Interactions among drivers are shown in Sections 4.5.1 and 4.5.2 for predictive biomarkers and surrogate endpoints, respectively.

4.4.8.1 Scientific Knowledge Base and Sample Size

The Scientific Knowledge Base plays an important role in influencing the number of samples required for the biomarker development and validation effort, particularly during the Identification and Feasibility phase and the Clinical Validation phase. The goal of the Identification and Feasibility phase is to build confidence that the biomarker is reflective of the outcome it is intended to measure. The breadth and depth of the existing scientific knowledge plays an important role in determining the amount of additional work that will be needed to build confidence in the biomarker. The number of samples needed during the Identification and Feasibility phase is affected by the amount of research that needs to be performed to build confidence in the biomarker before moving to the Develop Assay and Intended Use phase.

For predictive biomarkers, the Scientific Knowledge Base also affects Sample Size in the Clinical Validation and Utility phase because it affects the trial strategy. A robust scientific knowledge base can help reduce the Sample Size needed for the pivotal clinical validation studies. For example, enrichment trial designs are an attractive choice for predictive biomarkers because of their ability to reduce trial size. However, for this strategy to work there needs to be compelling evidence that patients in the unselected biomarker group will not benefit from treatment (Mandrekar and Sargent, 2009). The evidence needed to support the rationale or justification for this trial strategy is heavily dependent on the scientific knowledge base surrounding the biomarker and the therapeutic target.

4.4.8.2 Sample Size and Sample Type

Sample Size influences the costs associated with Sample Type at multiple phases throughout the biomarker development and validation lifecycle for both predictive biomarkers and surrogate endpoints. Large sample sizes can drive up Sample Type costs because Sample Size acts as a multiplier of costs associated with Sample Type costs. As a result, the extent to which the Sample Type affects the overall costs of biomarker development and validation is affected by the Sample Size.

4.4.8.3 Sample Size and Trial Recruitment

Sample Size can also affect the costs associated with Trial Recruitment. Recruiting sufficient participants to larger trials can be difficult and increase the per participant recruitment costs as compared to smaller trials. With predictive biomarkers, low biomarker prevalence can make it more difficult, time-consuming, and costly to recruit participants. The lower the biomarker prevalence, the greater the number of samples that will need to be screened to identify sufficient positive samples for the trial. The impact of biomarker prevalence on Trial Recruitment is exacerbated in diseases with a low prevalence, because identifying trial participants becomes increasingly difficult.

4.4.8.4 Sample Type and Novelty and Complexity of the Technology

The Sample Type selected to measure the biomarker will influence the Novelty or Complexity of the Technology needed to measure the biomarker. Different types of biospecimens used as a sample (e.g., blood plasma vs. tissue) require use of different types of technology for assays with inherently different levels of cost. For example, magnetic resonance imaging (MRI) samples require imaging equipment and infrastructure, while viral DNA samples require sequencing technology. Additionally, Sample Type may impact additional technology needs during assay development such as collection, storage and extraction of samples suitable for surrogate endpoint evaluation.

4.4.8.5 Follow-Up Time Needed and Trial Recruitment

For surrogate endpoints, the Follow-Up Time Needed can affect the costs associated with Trial Recruitment. Longer trials tend to lose a greater number of participants. The longer the trial, the greater the risk that participants will move or lose interest in the trial and be lost to follow-up. Lengthier trials need to account for this risk during the recruitment phase by recruiting a greater number of participants at the outset to ensure that the trial is adequately powered at the conclusion. Recruiting additional participants to account for the increased participant attrition associated with lengthier trials increases both the time and costs of the Clinical Validation and Utility Phase.

4.4.9 Cost Driver Summary

In total, there are seven cost drivers that substantially affect the costs of biomarker development and validation at one or more phases of development. For predictive biomarkers, only five of these cost drivers substantially impact one or more developmental phases (Exhibit 11), while all seven cost drivers affect the costs of one or more phases of biomarker development and validation for surrogate endpoints (Exhibit 14). Additionally, even among those cost drivers that affect the costs of both predictive biomarkers and surrogate endpoints, the phases impacted by each cost driver differ between the two biomarker categories. As a result, two cost frameworks have been created, one describing how the cost drivers affect the costs of predictive biomarkers (Section 4.5.1) and one describing how the cost drivers affect the costs of surrogate endpoints (Section 4.5.2).

4.5 Cost Frameworks

Cost frameworks developed for predictive biomarkers and surrogate endpoints provide a structure for identifying cost drivers, estimating how cost drivers affect costs at individual development and validation phases, and estimating overall costs over the course over the biomarker development and validation lifecycle. Seven cost drivers can potentially have a substantial effect on the costs of biomarker development and validation at one or more phases of development.

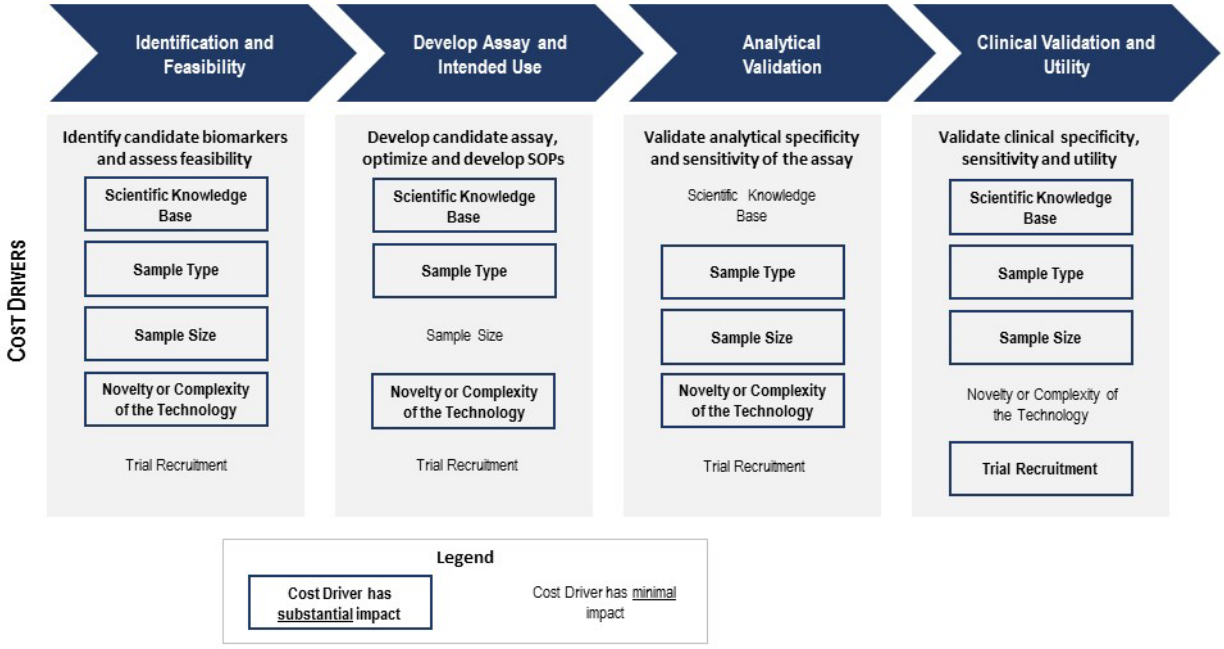
RESEARCH QUESTION 3
How do the factors that influence the cost of developing and validating biomarkers for use in drug development vary across the selected categories?
RESEARCH QUESTION 4
How do these factors contribute to the overall cost of developing and validating a biomarker within a particular category?
RESEARCH QUESTION 5
How do the cost drivers compare across biomarker categories?

4.5.1 Predictive Biomarker Cost Framework

For predictive biomarkers, only five of these cost drivers substantially impact one or more developmental phases (Exhibit 11). During the Identification and Feasibility phase, four cost drivers have a significant impact: Scientific Knowledge Base, Sample Type, Sample Size, and Novelty and Complexity of the Technology. During the Develop Assay and Intended Use Phase, three cost drivers have a significant impact: Scientific Knowledge Base, Sample Type, and Novelty

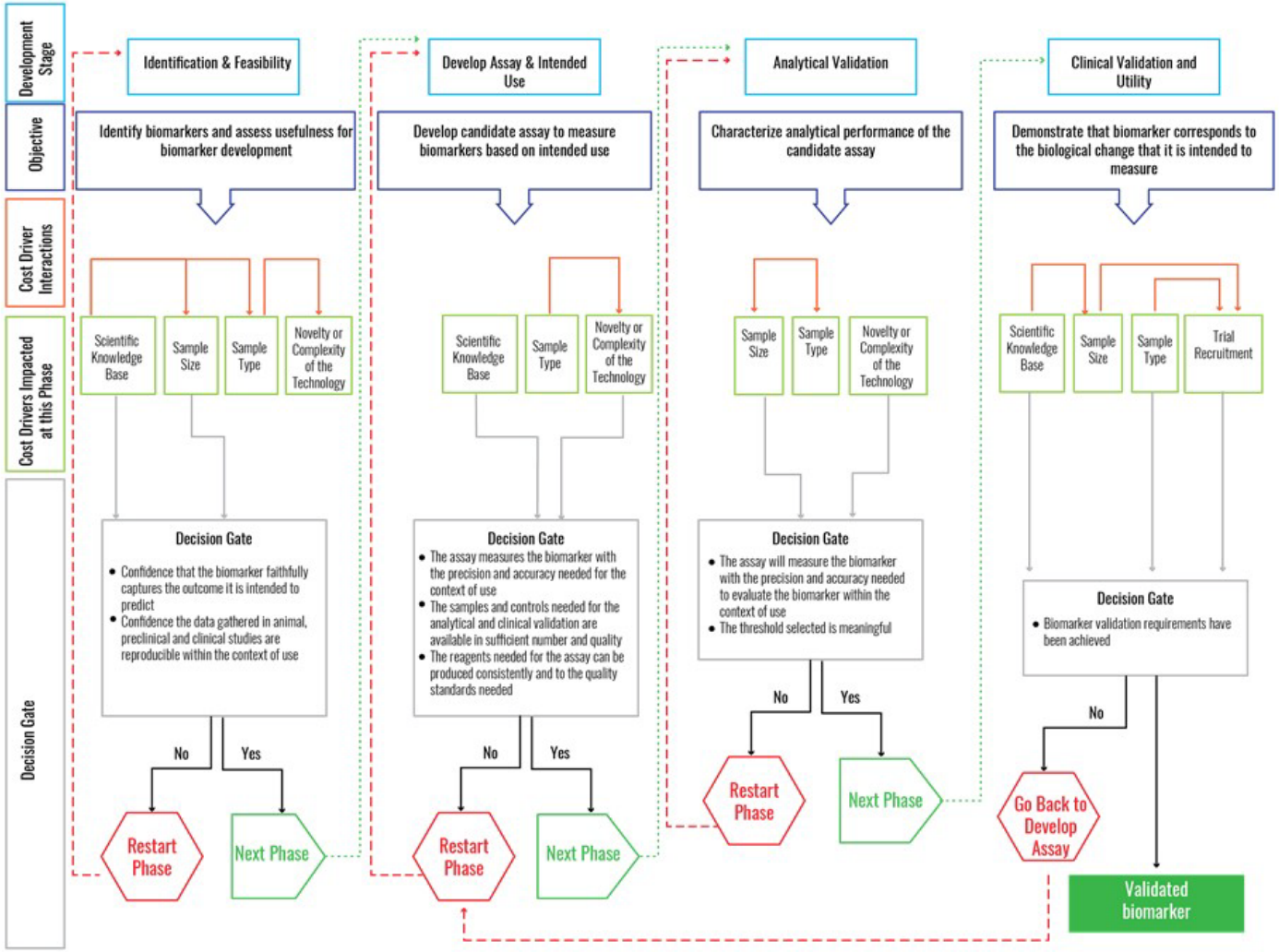
and Complexity of the Technology. During the Analytical Validation phase, three cost drivers have a significant impact: Sample Type, Sample Size, and Novelty and Complexity of the Technology. During the Clinical Validation and Utility phase, four cost drivers have a significant impact: Scientific Knowledge Base, Sample Type, Sample Size, and Novelty and Trial Recruitment.

Exhibit 11: Cost Driver Summary for Developing and Validating Predictive Biomarkers



The predictive biomarker cost framework (Exhibit 12) maps cost drivers, interactions between cost drivers, and the influence of cost drivers on decision gates that determine progressions across biomarker development and validations phases. Grey arrows indicate cost drivers with the most substantial impact on the decision gates.

Exhibit 12: Predictive Biomarker Cost Framework



In the Identification and Feasibility phase, during which time the objective is to identify and assess usefulness for biomarker development, four cost drivers have a substantial impact: Scientific Knowledge Base, Sample Type, Sample Size, and Novelty and Complexity of the Technology. Interactions between these cost drivers were discussed in section 4.4.8. Scientific Knowledge Base and Sample Size affect the decision gate, which takes two primary criteria into consideration: 1) confidence that the biomarker faithfully captures the outcome it is intended to predict; and 2) confidence that the data gathered in animal, preclinical, and clinical studies are reproducible within the context of use. If these decision gate criteria are met, then development and validation efforts can move on to the next phase, Develop Assay and Intended Use. If these decision gate criteria are not met, then more time needs to be spent in the Identification and Feasibility phase, potentially starting from the beginning of the phase.

In the Develop Assay and Intended Use phase, during which time the objective is to develop a candidate assay to measure biomarkers based on intended use, three cost drivers have a substantial impact: Scientific Knowledge Base, Sample Type, and Novelty and Complexity of the Technology. Interactions between these cost drivers were discussed in section 4.4.8. Scientific Knowledge Base and Novelty or Complexity of the Technology affect the decision gate, which takes three primary criteria into consideration: 1) that the assay measures the biomarker with the precision and accuracy needed for the context of use; 2) that the sample and controls needed for the analytical and clinical validation are available in sufficient number and quality; and 3) that the reagents needed for the assay can be produced consistently and to the quality of standards needed. If these decision gate criteria are met, then development and validation efforts can move on to the next phase, Analytical Validation. If these decision gate criteria are not met, then more time needs to be spent in the Develop Assay and Intended Use phase, potentially starting from the beginning of the phase.

In the Analytical Validation phase, during which time the objective is to characterize analytical performance of the candidate assay, three cost drivers have a substantial impact: Sample Size, Sample Type, and Novelty and Complexity of the Technology. Interactions between these cost drivers were discussed in section 4.4.8. Sample Size and Novelty or Complexity of the Technology affect the decision gate, which takes two primary criteria into consideration: 1) that the assay will measure the biomarker with the precision and accuracy needed to evaluate the biomarker for the context of use; and 2) that the threshold selected is meaningful. If these decision gate criteria are met, then development and validation efforts can move on to the next phase, Clinical Validation and Utility. If these decision gate criteria are not met, then more time needs to be spent in the Analytical Validation phase, potentially starting from the beginning of the phase.

In the Clinical Validation and Utility phase, during which time the objective is demonstrate that the biomarker corresponds to the biological change that it is intended to measure, four cost drivers have a substantial impact: Scientific Knowledge Base, Sample Size, Sample Type, and Trial Recruitment. Interactions between these cost drivers were discussed in section 4.4.8. Scientific Knowledge Base, Sample Type, and Trial Recruitment affect the decision gate, which takes one primary criterion into consideration: 1) that the biomarker validation requirements have been achieved. If this decision gate criterion is met, then the biomarker has been validated. If this decision gate criterion is not met, then additional bridging studies need to be conducted, and the development process returns back to the Develop Assay and Intended Use phase.

The summary of cost activities (Exhibit 13) identifies research and development activities associated with cost drivers across the phases of biomarker development for predictive biomarkers.

Exhibit 13: Impact of Cost Drivers on the Predictive Biomarker Development and Validation Process

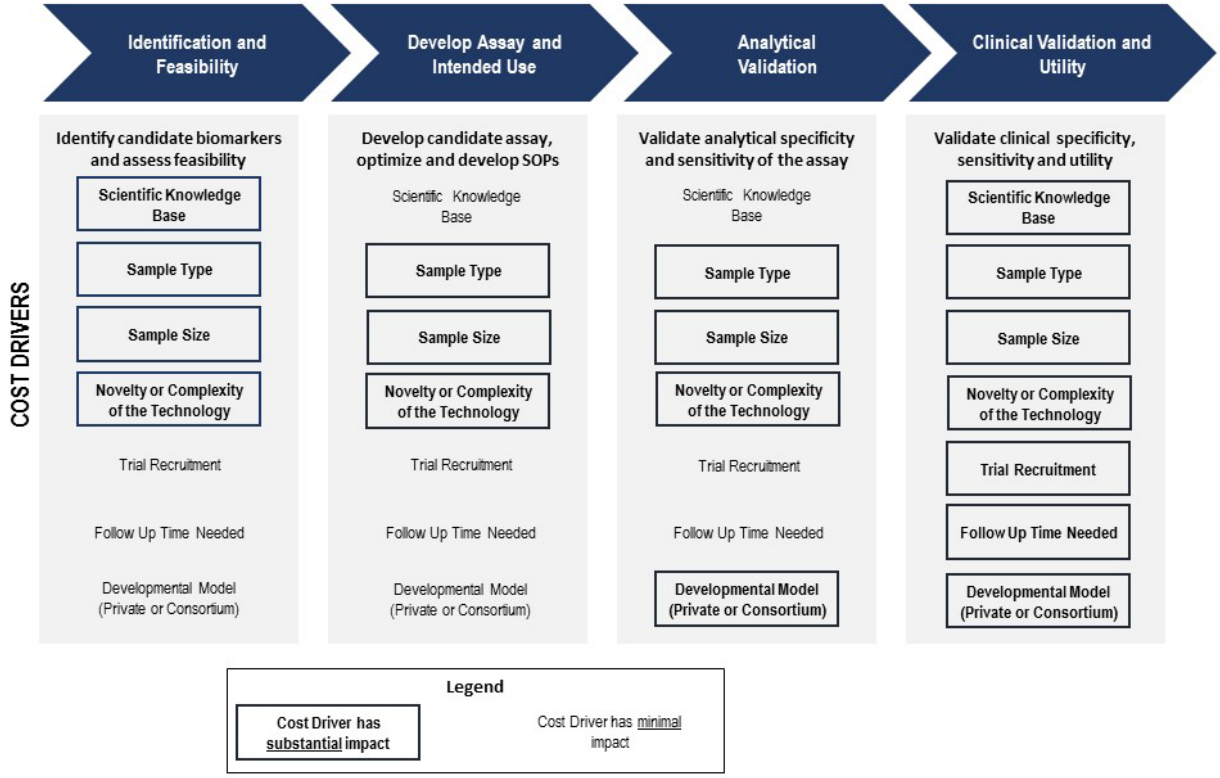
Cost Driver	Identification and Feasibility	Develop Assay and Intended Use	Analytical Validation	Clinical Validation and Utility
Scientific Knowledge Base	• Performing the experiments and reviewing the literature to identify biomarker candidates • Performing experiments to assess the relationship between the biomarker and the clinical outcome of interest • Obtaining samples for animal preclinical and clinical studies • Use of equipment and reagents • Use of laboratory infrastructure and overhead • Determining prevalence of the biomarker	• Assessing the most appropriate assay method to measure the biomarker • Performing experiments to assess the dynamic range of the biomarker and understand the potential of comorbidities on the biomarker • Obtaining clinical samples • Use of equipment and reagents • Use of laboratory infrastructure and overhead • Developing protocols	[minimal cost impact]	• Driving clinical trial design and strategy • Determining if the biomarker can be used as part of the inclusion criteria for an enrichment trial design, or whether an unselected design will be used in which participants are included in the trial regardless of their biomarker status
Sample Type	• Obtaining specimens o Obtaining the sample (e.g., blood draw, cheek swab, etc.) o Sample storage	• Obtaining specimens o Obtaining the sample (e.g., blood draw, cheek swab, etc.) o Sample storage	• Obtaining specimens o Obtaining the sample (e.g., blood draw, cheek swab, etc.) o Sample storage	• Obtaining specimens o Obtaining the sample (e.g., blood draw, cheek swab, etc.) o Sample storage
Sample Size	• Quantity of samples of samples needed • Shipping costs	[minimal cost impact]	• Quantity of samples of samples needed • Shipping costs	• Quantity of samples of samples needed • Shipping costs
Novelty or Complexity of the Technology	• Assessing and choosing technology • Building any algorithms needed to measure the biomarker • Data storage[26] and transfer costs	• Developing the infrastructure – e.g., building databases • Developing the protocol • Developing any training materials needed • Data storage and transfer costs	• Skilled labor to run, analyze or interpret the results • Expensive equipment • Data storage and transfer costs	[minimal cost impact]
Trial Recruitment	[minimal cost impact]	[minimal cost impact]	[minimal cost impact]	• Recruiting patients • Retaining patients • Onboarding the sites and maintaining equipment (site overhead) • Training • Shipping and processing samples

[26] Data Storage and Transfer: Cost component that pertains to data storage and transfer of data. Biomarkers that collect genetic or other storage-intensive data have significant data storage and transfer needs.

4.5.2 Surrogate Endpoint Cost Framework

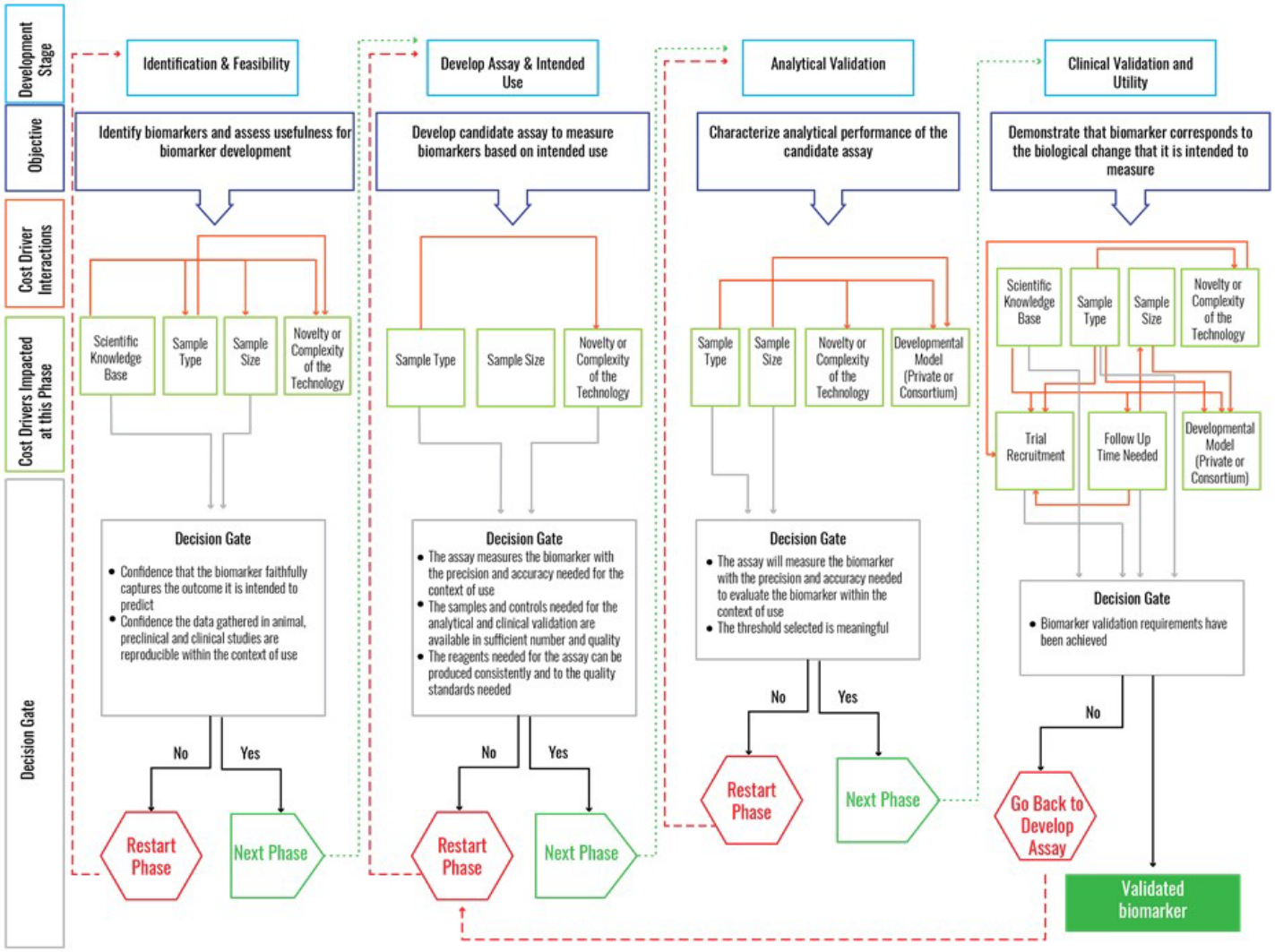
For surrogate endpoints, seven cost drivers substantially impact the costs of one or more phases of biomarker development and validation (Exhibit 14). During the Identification and Feasibility phase, four cost drivers have a substantial impact: Scientific Knowledge Base, Sample Type, Sample Size, and Novelty and Complexity of the Technology. During the Develop Assay and Intended Use phase, three cost drivers have a substantial impact: Sample Type, Sample Size, and Novelty and Complexity of the Technology. During the Analytical Validation phase, four cost drivers have a substantial impact: Sample Type, Sample Size, Novelty and Complexity of the Technology, and Developmental Model (Private or Consortium). During the Clinical Validation and Utility, all seven cost drivers have a substantial impact: Scientific Knowledge Base, Sample Type, Sample Size, Novelty and Complexity of the Technology, Trial Recruitment, Follow-up Time Needed, and Developmental Model (Private or Consortium).

Exhibit 14: Cost Driver Summary for Developing and Validating Surrogate Endpoints



The surrogate endpoint cost framework (Exhibit 15) maps cost drivers, interactions between cost drivers, and the influence of cost drivers on decision gates that determine progressions across biomarker development and validations phases.

Exhibit 15: Surrogate Endpoint Cost Framework



In the Identification and Feasibility phase, during which time the objective is to identify and assess usefulness for biomarker development, four cost drivers have a substantial impact: Scientific Knowledge Base, Sample Type, Sample Size, and Novelty and Complexity of the Technology. Interactions between these cost drivers were discussed in Section 4.4.8. Scientific Knowledge Base and Sample Size affect the decision gate, which takes two primary criteria into consideration: 1) confidence that the biomarker faithfully captures the outcome it is intended to predict; and 2) confidence that the data gathered in animal, preclinical, and clinical studies are reproducible within the context of use. If these decision gate criteria are met, then development and validation efforts can move on to the next phase, Develop Assay and Intended Use. If these decision gate criteria are not met, then more time needs to be spent in the Identification and Feasibility phase, potentially starting from the beginning of the phase.

In the Develop Assay and Intended Use phase, during which time the objective is to develop a candidate assay to measure biomarkers based on intended use, three cost drivers have a substantial impact: Sample Type, Sample Size, and Novelty and Complexity of the Technology. Interactions between these cost drivers were discussed in section 4.4.8. Sample Type and Novelty or Complexity of the Technology affect the decision gate, which takes three primary criteria into consideration: 1) that the assay measures the biomarker with the precision and accuracy needed for the context of use; 2) that the sample and controls needed for the analytical and clinical validation are available in sufficient number and quality; and 3) that the reagents needed for the assay can be produced consistently and to the quality of standards needed. If these decision gate criteria are met, then development and validation efforts can move on to the next phase, Analytical Validation. If these decision gate criteria are not met, then more time needs to be spent in the Develop Assay and Intended Use phase, potentially starting from the beginning of the phase.

In the Analytical Validation phase, during which time the objective is to characterize analytical performance of the candidate assay, for cost drivers have a substantial impact: Sample Type, Sample Size, Novelty and Complexity of the Technology, and Developmental Model (Private or Consortium). Interactions between these cost drivers were discussed in section 4.4.8. Sample Type and Sample Size affect the decision gate, which takes two primary criteria into consideration: 1) that the assay will measure the biomarker with the precision and accuracy needed to evaluate the biomarker for the context of use; and 2) that the threshold selected is meaningful. If these decision gate criteria are met, then development and validation efforts can move on to the next phase, Clinical Validation and Utility. If these decision gate criteria are not met, then more time needs to be spent in the Analytical Validation phase, potentially starting from the beginning of the phase.

During the Clinical Validation and Utility phase, during which time the objective is demonstrate that the biomarker corresponds to the biological change that it is intended to measure, seven cost drivers have a substantial impact: Scientific Knowledge Base, Sample Size, Sample Type, Novelty and Complexity of the Technology, Trial Recruitment, Follow-Up Time Needed and Developmental Model (Private or Consortium). Interactions between these cost drivers were discussed in section 4.4.8. Scientific Knowledge Base, Sample Type, Trial Recruitment, and Follow-Up Time Needed affect the decision gate, which takes one primary criterion into consideration: that the biomarker validation requirements have been achieved. If this decision gate criterion is met, then the biomarker has been validated. If this decision gate criterion is not met, then additional bridging studies need to be conducted, and the development process returns back to the Develop Assay and Intended Use phase.

Exhibit 16 summarizes activities across cost drivers and phases of surrogate endpoint development.

Exhibit 16: Impact of Cost Drivers on the Surrogate Endpoint Development and Validation Process

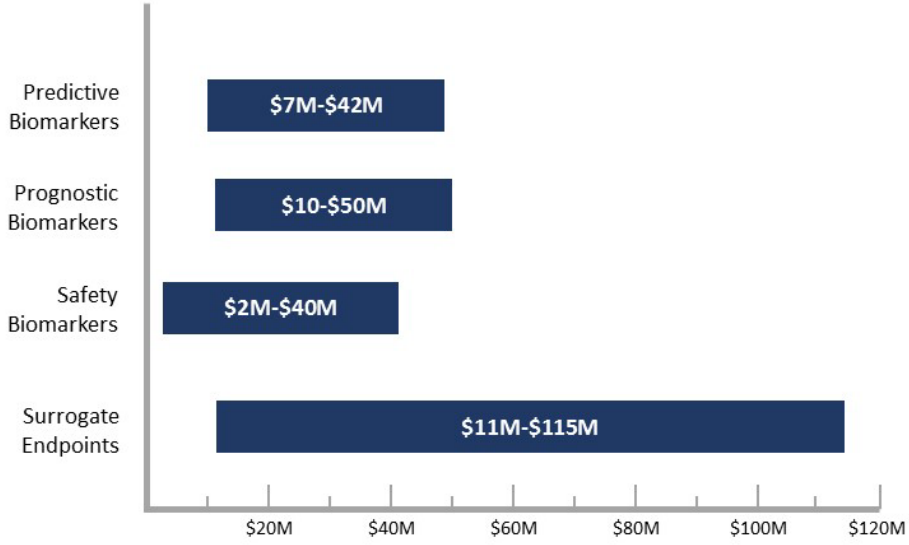
Cost Driver	Identification and Feasibility	Develop Assay and Intended Use	Analytical Validation	Clinical Validation and Utility
Scientific Knowledge Base	• Drives the number, type and size of studies needed for the assessment of feasibility	[minimal cost impact]	[minimal cost impact]	• Drives the number and size of additional studies needed for the clinical validation
Sample Type	• Obtaining specimens o Obtaining the sample (e.g., blood draw, cheek swab, etc.) o Sample storage	• Obtaining specimens o Obtaining the sample (e.g., blood draw, cheek swab, etc.) o Sample storage	• Obtaining specimens o Obtaining the sample (e.g., blood draw, cheek swab, etc.) o Sample storage	• Obtaining specimens o Obtaining the sample (e.g., blood draw, cheek swab, etc.) o Sample storage
Sample Size	• Quantity of samples of samples needed • Shipping costs	• Quantity of samples of samples needed • Shipping costs	• Quantity of samples of samples needed • Shipping costs	• Quantity of samples of samples needed • Shipping costs
Novelty or Complexity of the Technology	• Assessing and choosing technology • Building any algorithms needed to measure the biomarker • Data storage and transfer costs	• Developing the infrastructure – e.g., building your databases • Developing the protocol • Developing any training materials needed • Data storage and transfer costs	• Developing validation thresholds • Validating multiple components of assay protocol • Data storage and transfer costs	• Training specialized staff • Validating equipment • Running assays • Data storage and transfer costs
Trial Recruitment	[minimal cost impact]	[minimal cost impact]	[minimal cost impact]	• Recruiting patients • Retaining patients • Onboarding the sites and maintaining equipment (site overhead) • Training • Shipping and processing samples
Follow-Up Time Needed	[minimal cost impact]	[minimal cost impact]	[minimal cost impact]	• Maintaining clinical trial infrastructure and databases • Procuring equipment if existing equipment becomes outdated • Recruiting additional participants to offset larger attrition for longer trials
Developmental Model (Private or Consortium)	[minimal cost impact]	[minimal cost impact]	• Developing consensus about optimal protocol • Coordination between sites • Training and standardization among sites	• Developing consensus about optimal protocol • Coordination between sites • Training and standardization among sites

4.6 Biomarker Development Costs

4.6.1 Cost Ranges by Biomarker Category

Overall costs of development and validation vary by biomarker category. Exhibit 17, below, shows an estimated range of what it costs to develop a new biomarker for each of the four biomarker categories. There is considerable variability, both within and between each of the biomarker categories. There was a high degree of expert input and cost validation for cost range estimates for predictive biomarkers. There was also general consensus among biomarker experts that prognostic biomarkers are more expensive to develop than predictive or safety biomarkers which may be due in part to the amount of follow-up time needed to observe the clinical endpoint/outcome of interest. Safety biomarkers are slightly less expensive and have somewhat less variability in terms of the actual dollar amount, which may be due in part to the consortium approach to safety biomarker development where expenses such as labor may not be wholly accounted for in the costs. Safety biomarkers also include pre-clinical biomarkers that have reduced clinical demands. Experts agreed that surrogate endpoints are the most costly and time-intensive to develop. This is due in large part to the clinical trials and clinical validation that are needed for surrogate endpoints. The findings highlighted that surrogate endpoints for common diseases can be from three to 50 times as expensive to develop as predictive biomarkers.

Exhibit 17: General Cost Range by Biomarker Category

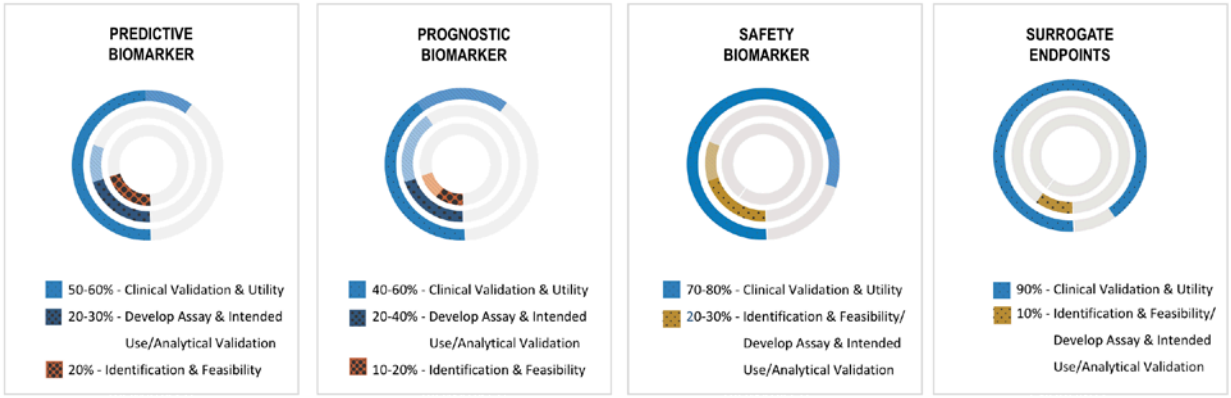


There was a paucity of publicly available data in the research literature regarding the costs associated with the identification, development or validation of new biomarkers. Most biomarker costs are borne by commercial entities such as pharmaceutical companies and this information is considered proprietary. The cost data represented here is almost entirely based on the consultations with 29 industry and academic biomarker experts interviewed for the project. The data collected were validated with multiple experts when possible. The data are most robust for predictive biomarkers, which have many examples of validated biomarkers; however, consulted experts were able to provide some cost information for each biomarker category.

Exhibit 18 depicts each biomarker category and how the development costs break down as a percentage of the total cost. Experts were not always able to give dollar amount figures for each of the four steps in the

process as in many cases budgets are not aligned to process steps. The cost breakdown can vary even within a biomarker category depending on the cost drivers. The biggest source of variability in costs among the categories comes during the final stage, the demonstration of clinical validation and utility which ranges from 40-90 percent of the total cost of biomarker development. Findings resulting from discussions with experts indicated that Clinical Validation and Utility was generally the most expensive development and validation phase. For predictive biomarkers, 50-60 percent of costs were attributed to this phase, with 20-30 percent going toward Develop Assay and Intended Use and Analytical Validation, and 20 percent of costs attributed to the Identification and Feasibility phase. For prognostic biomarkers 40-60 percent of costs were attributed to the Clinical Validation and Utility phase, with 20-40 percent going toward Develop Assay and Intended Use and Analytical Validation, and 10-20 percent of costs attributed to the Identification and Feasibility phase. For safety biomarkers, experts attributed 70-80 percent of costs toward Clinical Validation and Utility, with the remaining 20-30 percent of costs going toward the other three phases. For surrogate endpoints, experts attributed 90 percent of overall costs to Clinical Validation and Utility, with the remaining 10 percent going toward the remaining three phases.

Exhibit 18: General Cost Percentages by Biomarker Category and Phase[27] *(text version)*



Note: values shown as a range represent the minimum and maximum possible percentage of total costs for each development step.

[27] Different experts classified costs slightly differently; accordingly, some of the bars in the graphic represent more than one stage. For prognostic biomarkers and predictive biomarkers, experts broke the costs down into three stages: Identification and Feasibility (inner orange bar); Develop Assay and Intended Use and Analytical Validation (middle purple bar); and the Clinical Validation and Utility (outer darker blue bar). The lighter shade portion of the bar shows the variability in the estimates obtained for each stage (e.g., 40-60 percent). For the last two categories, safety biomarkers and surrogate endpoints, experts classified the costs into two steps: the first step included everything up to clinical validation & utility (Identification and Feasibility; Develop Assay and Intended Use; and Analytical Validation; inner gold); while the second stage included the costs of clinical validation and demonstration of clinical utility (outer blue bar). As with the first two pictures, the lighter shade portion of the bar indicates the variability in the estimates for each stage. The innermost, light grey circle in the safety and surrogate endpoint charts is a placeholder included for consistency.

4.6.2 Cost Ranges by Phase of Development

RESEARCH QUESTION 2A
What are the ranges of costs associated with a given component of a biomarker development program (e.g., validation of a biomarker's analytical method and its associated platform)?

Exhibit 19 presents the mean and range (minimum and maximum) of development validation costs associated with predictive and surrogate endpoints. Cost ranges were based on general information and experience discussed during interviews with experts, and the data displayed is necessarily limited to the six biomarkers explored in the case studies (Appendix D). As shown below, there can be a significant range in costs within a phase of development.

Exhibit 19: Cost Ranges for Predictive Biomarkers and Surrogate Endpoints by Phase

	Predictive Biomarkers (Mean, range [min/max])		Surrogate Endpoints (Mean, range [min/max])	
Identification and Feasibility	Mean: $1,315,000		Mean: $1,983,333	
	Min: $550,000	Max: $3,000,000	Min: $500,000	Max: $5,000,000
Develop Assay and Intended Use	Mean: $2,757,000		Mean: $2,805,000	
	Min: $1,300,000	Max: $6,000,000	Min: $65,000	Max: $7,000,000
Analytical Validation	Mean: $3,668,750		Mean: $8,375,000	
	Min: $2,000,000	Max: $5,000,000	Min: $2,000,000	Max: $20,750,000
Clinical Validation and Utility	Mean: $8,875,000		Mean: $10,341,667	
	Min: $1,500,000	Max: $33,000,000	Min: $4,000,000	Max: $20,750,000
Total	Mean: $15,700,000		Mean: $23,505,000	

4.6.3 Cost Impact Scenarios

Cost impact scenarios were developed to describe how a cost driver can impact the overall cost of biomarker development and validation. The cost impact scenario descriptions are intended to help provide an understanding of the conditions that would result in an increased cost for a cost driver. The cost scenario descriptions were not intended to be exhaustive, but rather to describe some of the characteristics that a biomarker may have in each phase. It is unlikely than any one biomarker exhibits all the characteristics listed for any one scenario. Estimated cost impacts were based on discussions with experts about the six biomarker case studies (Appendix D).

Exhibit 20 presents a description of example scenarios and cost impact estimates broken down into low and high cost impact for the Scientific Knowledge Base cost driver. Included are example biomarkers that would fall into the category.

Exhibit 20: Cost Impact Scenarios for Scientific Knowledge Base

Impact Level	Description	Examples	Estimated Cost Impact
Low	• Biomarker is well characterized in animal models and humans • The biomarker's role in health and disease is well understood • Minimal additional research is necessary to assess feasibility • The dynamic range of the biomarker is well understood in healthy/normal controls and individuals with the condition of interest • Impact of co-morbidities on the biomarker is well understood and documented • The underlying pathophysiology/mechanism is well understood • Confidence in the predictive power of the biomarker is high • There is sufficient evidence to consider using the biomarker as part of the inclusion criteria	EGFR, second version assay (predictive biomarker) Hepatitis C virus (HCV) RNA viral load (surrogate endpoint)	Little to no cost impact over baseline in the low scenario for predictive biomarkers or surrogate endpoints
High	• Biomarker is only characterized in animal models or poorly understood • Studies of the biomarker have been exploratory/small in both animal models and humans, if applicable • There is a limited understanding of the biomarker's role in health and disease • Significant additional research is needed to assess feasibility • Few studies are available looking at the dynamic range of the biomarker • Additional work is needed to determine the full dynamic range and potential impact of co-morbid conditions • Confidence in the predictive power of the biomarker is low • There is likely insufficient evidence to consider using the biomarker as an inclusion criterion for a trial	EGFR, first version (v1) of the assay (predictive biomarker) FDG-PET (surrogate endpoint)	$740,000 to $4.9 million for predictive biomarkers $587,000 to $3.9 million for surrogate endpoints

Exhibit 21 presents a description of example scenarios and cost impact estimates broken down into low and high cost impact for the Sample Type cost driver. Also included are examples of what would fall in each category. The impact that Sample Type has on the overall costs to develop and validate a biomarker will also depend upon the Sample Size and, to a lesser extent, Amount of Follow-up Time Needed.

Exhibit 21: Cost Impact Scenarios for Sample Type

Impact Level	Description	Examples	Estimated Cost Impact
Low	• Easily obtained samples (e.g., saliva, blood, urine) for common diseases	• Blood sample for diabetes or hypertension	Little to no cost impact over baseline in the low scenario for predictive biomarkers or surrogate endpoints.
High	• Easily obtained samples for rare or less common diseases • Samples that are more difficult to obtain but for relatively common diseases • Difficult-to-obtain samples for rare diseases • Extremely rare diseases • Biomarkers with a tight and unpredictable temporal window (e.g., during septic shock)	• Blood samples for abacavir hypersensitivity (predictive biomarker) • Lung biopsies from individuals with NSCLC • CSF for rare neurological disorders • Blood samples from individuals with progeria • Blood samples during septic shock	$1.1 million to $7.4 million for predictive biomarkers $2.0 million to $13.1 million for surrogate endpoints

Exhibit 22 presents a description of example scenarios and cost impact estimates broken down into low and high cost impact for the Sample Size cost driver. Also included are examples of what would fall in each category. In general, experts asked about the sample size needed to clinically validate a new biomarker described studies with 500 or fewer participants as small. The cost estimates presented for these categories in Exhibit 22 reflect the range of fluctuation possible with sample sizes.

Exhibit 22: Cost Impact Scenarios for Sample Size

Impact Level	Description	Examples	Estimated Cost Impact
Low	• Fewer than 500 samples	• HCV RNA viral load (surrogate endpoint)	Little to no cost impact over baseline in the low scenario for predictive biomarkers or surrogate endpoints
High	• More than 500 samples	• EGFR, Programmed death-ligand 1 (PD-L1) (predictive biomarkers) • Estimated glomerular filtration rate (eGFR) (surrogate endpoint)	$3.4 million to $22.9 million for predictive biomarkers $2.3 million to $15.6 million for surrogate endpoints

Exhibit 23 presents a description of example scenarios and cost impact estimates broken down into low and high cost impact for the Novelty and Complexity of the Technology cost driver. Also included are examples of what would fall in each category. The potential range of additional costs is high for surrogate endpoints because they typically take a long time to develop and require a greater number of trials, and therefore the complexity of the technology to assess the samples can have a larger impact.

Exhibit 23: Cost Impact Scenarios for Novelty or Complexity of the Technology

Impact Level	Description	Examples	Estimated Cost Impact
Low	• No new or specialized equipment or reagents needed • Technology platform selected is routinely used in clinical care throughout the health care system • Cost of equipment or reagents is modest or low • No need to develop additional training materials • Protocols, types of controls, and experiments needed for the Analytical Validation phase are well established • Assay run, analysis, and interpretation is mainly automated • Little to no additional training needed for personnel	• HLA-B*5701 (predictive biomarker)	Little to no cost impact over baseline in the low scenario for predictive biomarkers or surrogate endpoints
High	• Some specialized equipment or reagents needed • Technology platform is used in clinical care, though perhaps not as ubiquitous as in the low impact scenario • Cost of purchasing and maintaining equipment is moderate or high, as is the cost of reagents • Running, analyzing, and interpreting the assay requires specialized labor • Protocols, types of controls, and experiments needed for the Analytical Validation phase are not well established and will require discussion with the appropriate regulatory agency or agencies • Additional training for personnel that prepare the specimen, run the assay, or interpret the results is required	• PD-L1 (predictive biomarker) • FDG-PET (surrogate endpoint)	$396,000 to $2.6 million for predictive biomarkers $861,000 to $5.7 million for surrogate endpoints

Exhibit 24 presents a description of example scenarios and cost impact estimates broken down into low and high cost impact for the Trial Recruitment cost driver. Sample Size, Sample Type, and the Amount of Follow-up Time Needed can influence the extent to which Trial Recruitment impacts the cost of developing and validating new biomarkers.

Exhibit 24: Cost Impact Scenarios for Trial Recruitment

Impact Level	Description	Examples	Estimated Cost Impact
Low	• Smaller trial size • Higher biomarker prevalence • Common disease • Potential benefit for trial participants	• HCV RNA viral load (surrogate endpoint)	Little to no cost impact over baseline in the low scenario for predictive biomarkers or surrogate endpoints
High	• Large trial size • Low biomarker prevalence (predictive biomarkers only) • Low disease prevalence • No potential benefit for trial participants • Competition for patient populations from other studies actively recruiting for the same disease or condition	• PD-L1, EGFR (predictive biomarker) • FDG-PET (surrogate endpoint) • eGFR (surrogate endpoint)	$456,000 to $3 million for predictive biomarkers $732,000 to $4.9 million for surrogate endpoints

Exhibit 25 presents a description of example scenarios and cost impact estimates broken down into low and high cost impact for the Follow-up Time Needed cost driver. Note that this cost driver meets the substantial cost impact threshold described in Section 3.3.1.3 only for surrogate endpoints and not predictive biomarkers because of the long follow-up time needed for surrogate endpoint clinical validation.

Exhibit 25: Cost Impact Scenarios for Follow-Up Time Needed

Impact Level	Description	Examples (Surrogate Endpoint)	Estimated Cost Impact
Low	• The amount of time needed to observe the clinical outcome the surrogate endpoint is intended to predict is relatively short, typically less than three years • Studies performed during the Identification and Feasibility phase and Clinical Validation phase do not need to follow participants for a long period of time and can be completed more rapidly	• None	Little to no cost impact over baseline in the low scenario for surrogate endpoints
High	• The amount of time needed to observe the clinical outcome the surrogate endpoint is intended to predict is relatively long, typically three to 10 years • As described in the low impact scenario, the amount of time needed to observe the clinical outcome will impact the length of both the Identification and Feasibility studies as well as the Clinical Validation and Utility phase studies	• FDG-PET • HCV RNA viral load • eGFR	$1.3 million to $8.9 million for surrogate endpoints

Exhibit 26 presents a description of example scenarios and cost impact estimates broken down into low and high cost impact for the Developmental Model (Private or Consortium) cost driver. Note that the cost driver is specific to surrogate endpoints, not predictive biomarkers.

Exhibit 26: Cost Impact Scenarios for Developmental Model (Private or Consortium)

Impact Level	Description	Examples (Surrogate Endpoint)	Estimated Cost Impact
Low	The biomarker is developed through a collaborative effort of a few organizations or entities, typically two or three	HCV RNA viral load	$783,000 to $1.8 million for surrogate endpoints
High	The biomarker is developed through a collaborative effort of several organizations or entities, typically more than five	FDG-PET	Up to $5.2 million for surrogate endpoints

4.6.4 Overarching Considerations

Due to the lack of robust biomarker-specific data available, cost of failure and cost of capital are not directly accounted for in the cost frameworks. Cost of failure and cost of capital formulas are provided in Section 3.3.3 to be used once data is available. However, they are two important factors to consider when assessing biomarker development and validation costs. Total costs for the development and validation of one successful biomarker must consider costs associated with failure of earlier biomarker efforts. Similarly, the cost of capital accounts for development time and can be a significant driver of total costs. Because data are not public or robust for biomarkers, these factors will be discussed qualitatively in the following sections.

4.6.4.1 Cost of Failure

Failure of the biomarker development and validation effort may be declared for many reasons, and there is likely to be a large variation in probabilities of success by indication and biomarker category. Fluctuations in failure rates can have significant impacts to total costs associated with one successful biomarker. As there is limited published information about biomarker development and validation costs, there is a lack of published information about the cost of failure for biomarkers. Both internal factors (i.e., factors tied to innate biomarker properties and performance) and external factors (i.e., factors not tied to innate biomarker properties and performance) can lead to biomarker failure. The distinction between internal and external failure factors is important to account for the development of an appropriate method of determining the rate of biomarker failure due to inadequacies of the biomarker itself or of the associated biomarker program.

A biomarker may fail at any phase of the biomarker development and validation lifecycle. However, experts noted that most biomarker candidates fail early in the development process. Researchers aim to fail early rather than late as less time and resources have been invested.

Internal factors that can lead to biomarker failure in the Identification and Feasibility phase include biomarker implausibility or lack of feasibility, technical inadequacy, or poor performance. If the biomarker does not provide confidence that it captures the outcome it is intended to predict, or if studies are unreproducible, researchers may decide to keep the biomarker candidate in the phase for longer to conduct additional characterization studies or elect to abandon the biomarker development effort.

In the Develop Assay and Intended Use phase, the biomarker may fail due to suboptimal performance characteristics. The required assay may not show consistent precision and accuracy needed for the context of use, be too expensive to produce, or a single assay result may be insufficient to characterize a heterogeneous sample.

As the biomarker moves to the Analytical Validation phase, the biomarker may fail if results are unable to be reproduced. If the biomarker has too many false positives or false negatives, or doesn't meet sensitivity or precision requirements, the biomarker effort may be abandoned.

In the Clinical Validation phase, the biomarker should be well characterized, but can still fail for a variety of reasons. The Clinical Validation phase is the costliest of the phases, so the overall rate of failure is typically

lower compared to the previous three phases. Reasons for failure include less robustness than initially thought or lack of clinical utility. The biomarker may fail to change clinical decision making or standards of care, or the use of the biomarker may bring harm from over-diagnosis. Even if the biomarker is robust and shows clinical utility, the marker may not be implemented due to cost, inconvenience, or other disadvantages. Additionally, the biomarker may be disproved or abandoned after initial use. As new data becomes available, or from re-analysis of the data, the original claims may fail to be validated.

Throughout the biomarker development and validation process, methodological shortcomings may lead to biomarker failure (e.g., hidden structures in the original data that result in biased data outcomes. Normal variation may dominate the observations, resulting in false positives or overfitting of the data. In addition, application of inappropriate statistical approaches or sampling methods my result in the publication of poorly substantiated findings.

Other than discontinuation of biomarker development due to failures from the biomarker itself, the biomarker may be abandoned due to external factors. These include the failure of the partner drug being developed, discontinuation of the drug program or discontinuation of the biomarker program due to a change in business priorities or shifts in the competitive landscape or funding. If the companion drug being developed fails, then there may be no need to develop the biomarker any further. If the drug or biomarker development program shifts priorities, or is shut down entirely, it may not be a priority to continue to fund further research into the specific biomarker project. All of these external factors may also result in failure of the biomarker in development.

4.6.4.2 Capital Costs

Cost of capital is not directly accounted for in the cost frameworks but is an important factor when considering the total cost for biomarker development and validation. Cost of capital increases out-of-pocket costs by the cost of capital for every year from the initial investment to approval (Paul et al, 2010). Because the biomarker development and validation lifecycle can be a relatively lengthy process (i.e., months or even years), the cost of capital can potentially have a major impact on the final cost of the biomarker. The cost of capital can vary from company to company, and varies based on profitability, creditworthiness, and operating history, among other factors. For pharmaceutical research and development costs, the average cost of capital is around 11 percent (Mestre-Ferrandiz, 2012) and capitalized costs can represent as much as 33 percent of the total cost to bring one successful drug to market.

A number of factors can affect development times throughout the biomarker development and validation process. If the biomarker is not well characterized, more time will be spent in early phases to address key knowledge gaps and further develop understanding. This will primarily affect the Identification and Feasibility phase and Develop Assay and Intended Use phase, where additional time is spent validating the relationship between the biomarker and what it is intended to measure, assessing biomarker feasibility, and building confidence in its usefulness. If critical questions about biomarker performance cannot be confidently answered, more time is spent in the phase to conduct additional studies or the biomarker effort is abandoned. In the later stages of biomarker development and validation, development times can be extended due to difficulties in patient recruitment, or length of the clinical trial. Surrogate endpoints are particularly susceptible to extended development and validation times because of the long follow-up times required to observe a clinical outcome. Similar to earlier in the development process, if the biomarker's performance or usefulness does not meet the requisite standards, more time can be spent in the Analytical Validation, Clinical Validation and Utility, and Intended Use phases. Extending the time in development not only increases out-of-pocket costs, but also the potential returns that could have been earned from an alternative investment.

5 DISCUSSION

5.1 Summary of Findings

Overall findings from this project identified and characterized four phases of the biomarker development and validation process: 1) Identification and Feasibility; 2) Develop Assay and Intended Use; 3) Analytical Validation; and 4) Clinical Validation and Utility. These phases were common between all four biomarker categories. This project also identified and characterized seven cost drivers that can potentially impact overall biomarker development and validation costs: Scientific Knowledge Base, Sample Type, Sample Size, Novelty and Complexity of the Technology, Trial Recruitment, Follow-Up Time Needed, and Developmental Model (Private or Consortium). This project developed a cost framework analysis for two biomarker categories – predictive biomarkers and surrogate endpoints. The cost framework identified variations in cost drivers between the two biomarker categories and between the four development and validation phases and characterized interactions between cost drivers. The cost framework also provided cost estimates for each biomarker category, including costs associated with each development and validation phase. To promote development and validation of biomarker candidates, a better understanding of associated costs can encourage dedication of resources and investments by supporting research organizations.

5.2 Limitations

Due to project scope and characteristics of cost data, certain limitations of this project should be considered when applying findings and conclusion to inform development and validation of predictive biomarkers and surrogate endpoints, as well as other type of biomarkers. Biomarkers used in the drug development space are diverse, complex and may use a variety of technologies to build a useful tool for application to drug development. It is important to acknowledge that this cost framework is limited due to 1) use of a limited set of biomarkers that does not represent the full range of biomarker diversity by exclusion of imaging and genomic biomarkers to name a few, 2) lack and limited quality of available data, and 3) large variability in the maturity of the science as relates to a proposed biomarker. The number of observations for this project (i.e., sample size) provided data to address the research questions; however, results may not be broadly generalizable. Case studies of prognostic and safety biomarkers were not carried out for this report, so less cost information is available for these two biomarker categories. Additional project limitations include availability, quality, and categorization of cost data.

5.2.1 Data Availability

There is currently extremely limited public, open-source cost data on biomarker development and validation. The sensitive nature of proprietary information for commercially developed biomarkers in addition to the long development cycle and multiple funding mechanisms for consortia-developed biomarkers contribute to the difficulty of collecting direct cost information. While there are several private biomarker development data sources available, the cost to procure that information was prohibitive for this project. Additionally, in interviews conducted with industry biomarker development teams, scientists did not elect to provide data on the cost of failure. Because of the lack of biomarker-specific data, a discussion is provided on the role cost of failure and cost of capital play when considering total costs rather than directly incorporating the cost impact into the framework.

5.2.2 Data Quality

The cost data used in this report was gathered through an environmental scan, case study interviews, industry average data, and SME input. Cases studies, which generally provided more granular cost data, were conducted on three predictive biomarkers and three surrogate endpoints (Appendix D). Because of this relatively small sample size and greater amount of data collected for predictive biomarkers compared to surrogate endpoints, the calculated baseline biomarker development and validation costs, as well as assessed cost impacts, may not be broadly representative of the biomarker industry. Average annual cost estimates for some cost components (e.g., salaries [for technicians or radiologists] and data storage) were assumed to be similar between biomarker categories, which may not be a broadly generalizable assumption. Additionally, there were significant in-kind labor costs, particularly from consortia-based development models, for which conservative estimates were necessary. Total costs reported in the case studies likely underestimate the holistic cost of developing a biomarker. In general, the estimated total cost reported here may be less than what it would cost for *de novo* biomarker development.

5.2.3 Data Categorization

In addition, experts noted that a major confounder in biomarker cost analysis is accurately attributing activities that served many purposes to the cost of biomarker development. For example, overhead costs and expensive equipment (e.g., MRI technology) that could be funded by and used for multiple purposes were challenging to estimate for biomarker-specific activities. Therefore, different experts categorized cost data differently across case studies. Assigning costs to specific phases or activities in the cost framework was particularly challenging for biomarker candidates developed in a consortium model, which were predominantly surrogate endpoints and safety biomarkers. Additionally, experts participating in consortia development model did not always have visibility into certain phases of development because their participation was limited to specific phases.

5.3 Policy Relevance

Throughout the course of the project, a number of issues were identified that impact the timeline and complexity of biomarker development without necessarily impacting the costs or process to validate directly. Nonetheless, these issues, if they remain unresolved, could potentially limit companies' or organizations' willingness to undertake new biomarker development and validation efforts. As such they have been included in this report. The three major issues that were identified, primarily through the expert consultations, include:

- Lack of harmony on informed consent, particularly globally;
- Lack of clarity, consensus, or guidelines around the types and amount of data needed to validate certain categories of biomarkers, specifically surrogate endpoints; and
- Unfair competition for companion diagnostic tests between laboratories that provide tests as a service in accordance with Clinical Laboratory Improvement Amendments (CLIA) regulations and diagnostic kit manufacturers who must comply with FDA regulations.

One of the challenges of doing multisite trials is site-to-site variation. Historically, part of this variation has arisen from the fact that different clinical sites fall under the direction of local institutional review boards (IRBs), even when those sites are part of the same study. Often, local IRBs will have similar but slightly different requirements for the informed consent document that could potentially prevent samples collected at one or more locations from being used in the future for biomarker bridging studies. These differences in the informed consent document are only exacerbated when the trial sites span more than one country. As

one expert pointed out, if an organization is conducting a trial at sites in more than one country, it is almost a guarantee that there will be multiple IRBs and multiple versions of the informed consent form, some of which will not allow for the retesting of clinical samples at a later date. Experts felt that even being able to achieve a situation in which there was one informed consent form per country would be a huge step forward.

Many of the challenges associated with having to work with multiple IRBs in the United States are being addressed through recent legislation and policy changes, including the passage of the 21st Century Cures Act and the HHS Common Rule (Common Rule, 45 CFR 46), both of which stipulate that multisite trials should move to using a single IRB.[28] However, the challenges remain for trials that are performed beyond U.S. borders.

Second, experts also mentioned that the lack of guidelines or consensus around the amount and type of data needed to validate a surrogate endpoint is an obstacle that makes it more difficult to develop new surrogate endpoints. Experts were quick to praise the FDA guidelines on companion diagnostics (FDA, 2014) and the clarity they provide around the types of data needed to bring new predictive biomarkers to market. The transparency and predictability of the requirements for developing a new companion diagnostic make it a much more attractive and viable proposition for pharmaceutical companies. Experts indicated that developing a similar set of guidelines or convening additional working groups to add clarity to the requirements in an application-specific manner would greatly improve the process for surrogate endpoints. Several experts described situations in which they had submitted data to support the use of a biomarker as a surrogate endpoint and were told that while the evidence was good, it was insufficient for the validation of the biomarker as a surrogate endpoint. While the experts specifically mentioned surrogate endpoints, the feedback can be broadly applied to other biomarker categories as well. The safety biomarker experts consulted highlighted that one of the key functions of safety biomarker consortia is to build consensus within the pharmaceutical and biotechnology community as well as among regulators around the intended use and evidence needed to validate new safety biomarkers. An expert noted that this is a model that could potentially be applied on a case-by-case basis to help build clarity and consensus around the evidence needed to validate other biomarkers.

Lastly, several experts noted the challenges posed by CLIA regulations for the development of companion diagnostic tests. Diagnostic developers and pharmaceutical companies will invest millions to develop and validate an assay that reliably measures a biomarker for a given purpose. The FDA holds these developers to a set of standards for assay reliability and performance which ultimately increase the development costs because they often necessitate several rounds of assay optimization as well as evidence of analytical and clinical validation and a formal application for market approval from the FDA. In contrast, CLIA laboratories are free to develop their own version of the assay without the same requirement to demonstrate clinical validity and no requirement to apply for market approval. Experts noted that this creates an unfair competitive advantage for CLIA laboratories, who advertise a similar but non-FDA approved assay at a significantly lower cost. Most experts indicated that CLIA laboratories are typically able to perform a similar assay at approximately half the cost of the FDA-approved assay. Experts were quick to note that the profit margin for diagnostics is already very small and that the unfair competitive advantage provided to CLIA laboratories further weakens the business case for developing new biomarkers. As Dracopoli and Boguski (2017) point out in their review of oncology therapeutic products brought to market since 1998, “the cost-to-benefit ratio for the great majority (90.4 percent) of the approved drugs was not in favor of developing a

[28] Section 3023 of the 21st Century Cures Act streamlines the IRB process for multisite clinical trials and Section 3056 strikes the requirement that a medical device trial must always use a local IRB, allowing for the use of centralized IRBs with medical device trials.

[companion diagnostic] test." Addressing the inequities inadvertently created by CLIA could help improve the business case for investing in new biomarkers.

5.4 Key Knowledge Gaps

While this report includes robust information drawn from literature reviews, case studies and expert consultations addressing development and validation costs for predictive biomarkers and surrogate endpoints, less is known about other biomarker categories including prognostic and safety biomarkers. Given significant variations in costs between predictive biomarkers and surrogate endpoints, extension of findings from this report to other biomarker categories is limited. Additionally, the lack of reliable information about the cost of failure for biomarker candidates is an impediment to modeling overall biomarker development resources. Cost of failure represents a significant financial outlay and is persistent factor in decision-making across the development and validation lifecycle. When considering the relatively small sample size of case studies in this report along with the inherent variability across biomarker categories and development phases, uniform cost data applicable across the development and validation landscape is still lacking. In addition, a comprehensive understanding of development and validation costs for biomarkers supported by consortia-driven initiatives is lacking. Given these large-scale collaborative initiatives include multiple organizations contributing to different phases of development and validation, individual participating experts generally have a limited perspective on the overall end-to-end process.

5.5 Areas for Further Inquiry

This project focused on the cost of developing biomarkers to a specific entity - whether the entity was a company, team or consortium. Costs should be evaluated in comparison to the benefits of the efforts. Additional studies are needed to understand the entire scope of benefits the sponsor (e.g. biopharmaceutical companies) receives for their efforts. Also, more research is needed to understand the social benefits of biomarker development, including the benefits realized by potential competitors that can use the validated biomarker to support their drug development programs.

Further analyses of costs could yield valuable information for organizations seeking to develop and validate biomarker candidates. Development of a financial cost model for putative biomarker candidates could be explored based upon the cost framework and cost data generated by this project; these models could be constructed to incorporate additional cost data and understanding of biomarker cost drivers generated by future research projects. The cost framework analysis could also be extended to include consideration of indirect costs.

Finally, future research and analysis efforts could seek to address key knowledge gaps identified above. Research methodologies described in this report could be extended to explore costs associated with prognostic and safety biomarker categories. Future convening initiatives could include short workshops of subject matter experts to address stated knowledge gaps for biomarker development and validation. For example, a panel of both economic and scientific experts could focus on developing more advanced and detailed methodologies for estimating costs of failure that stand upon models that take various assumptions into account. Additionally, stakeholders from across a broad scope of organizations that participate in consortia-based biomarker development could construct a comprehensive assessment of processes and costs across this relatively diffuse and fragmented collaborative development model. Analysis of which organizations bear which development costs would be beneficial, which could also be extended to non-consortia-based biomarker development and validation pathways that include multiple participants, but are not coordinated centrally by an organization or consortia model.

REFERENCES

Adaszewski, S., Dukart, J., Kherif, F., Frackowiak, R., Draganski, B., & Alzheimer's Disease Neuroimaging Initiative. (2013). How early can we predict Alzheimer's disease using computational anatomy? *Neurobiology of Aging*, *34*(12), 2815–2826. https://doi.org/10.1016/j.neurobiolaging.2013.06.015

Badve, S. V., Palmer, S. C., Strippoli, G. F. M., Roberts, M. A., Teixeira-Pinto, A., Boudville, N., … Johnson, D. W. (2016). The Validity of Left Ventricular Mass as a Surrogate End Point for All-Cause and Cardiovascular Mortality Outcomes in People With CKD: A Systematic Review and Meta-analysis. *American Journal of Kidney Diseases: The Official Journal of the National Kidney Foundation*, *68*(4), 554–563. https://doi.org/10.1053/j.ajkd.2016.03.418

Butterworth, S. E., Ingram, G., & Robertson, N. P. (2016). Advances in biomarker research in multiple sclerosis. *Journal of Neurology*, *263*(3), 621–623. https://doi.org/10.1007/s00415-016-8062-y

Cook, D., Brown, D., Alexander, R., March, R., Morgan, P., Satterthwaite, G., & Pangalos, M. N. (2014). Lessons learned from the fate of AstraZeneca's drug pipeline: a five-dimensional framework. *Nature Reviews. Drug Discovery*, *13*(6), 419–431. https://doi.org/10.1038/nrd4309

DiMasi, J. A., Grabowski, H. G., & Hansen, R. W. (2016). Innovation in the pharmaceutical industry: New estimates of R&D costs. *Journal of Health Economics*, *47*, 20–33. https://doi.org:10.1016/j.jhealeco.2016.01.012

Dracopoli, N. C., & Boguski, M. S. (2017). The Evolution of Oncology Companion Diagnostics from Signal Transduction to Immuno-Oncology. *Trends in Pharmacological Sciences*, *38*(1), 41–54. https://doi.org/10.1016/j.tips.2016.09.007

Emanuel, E. J., Schnipper, L. E., Kamin, D. Y., Levinson, J., & Lichter, A. S. (2003). The costs of conducting clinical research. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, *21*(22), 4145–4150. doi:10.1200/JCO.2003.08.156

FDA-NIH Biomarker Working Group. (2016). *BEST (Biomarkers, EndpointS & Other Tools) Resource*. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK326791/ (Accessed February 14, 2017)

Freedman, L. P., Cockburn, I. M., & Simcoe, T. S. (2015). The economics of reproducibility in preclinical research. *PLoS Biology*, *13*(6), e1002165. doi:10.1371/journal.pbio.1002165

Gallo, A. (2015, April 30). A refresher on cost of capital. *Harvard Business Review*. Retrieved from https://hbr.org/2015/04/a-refresher-on-cost-of-capital

Huang, M., Yang, W., Feng, Q., Chen, W., & Alzheimer's Disease Neuroimaging Initiative. (2017). Longitudinal measurement and hierarchical classification framework for the prediction of Alzheimer's disease. *Scientific Reports*, *7*, 39880. https://doi.org/10.1038/srep39880

Institute of Medicine Committee on Qualification of Biomarkers and Surrogate Endpoints in Chronic Disease. (2010). *Report on the evaluation of Biomarkers and Surrogate Endpoints in Chronic Disease.* Washington, DC: National Academies Press.

Lopez, G., Parker, J. L., Willan, A., Shah, S., Weerasinghe, A., Rubinger, D., & Falconi, A. The role of biomarkers in improving clinical trial success: A study of 1,079 oncology drugs [abstract]. In: 2015 American Society of Clinical Oncology Annual Meeting; May 29 – Jun 2, 2015; Chicago, Illinois. Alexandria (VA): ASCO; 2015. Abstract nr e17804.

Matheis, K. A., Com, E., Gautier, J.-C., Guerreiro, N., Brandenburg, A., Gmuender, H., … Staedtler, F. (2011). Cross-study and cross-omics comparisons of three nephrotoxic compounds reveal mechanistic insights and new candidate biomarkers. *Toxicology and Applied Pharmacology*, *252*(2), 112–122. https://doi.org/10.1016/j.taap.2010.11.006

Mandrekar, S. J., & Sargent, D. J. (2009). Clinical trial designs for predictive biomarker validation: One size does not fit all. *Journal of Biopharmaceutical Statistics*, *19*(3), 530–542. doi:10.1080/10543400902802458

Mestre-Ferrandiz, J., Sussex, J., & Towse, A. (2012). *The R&D cost of a new medicine.* Retrieved from https://www.ohe.org/publications/rd-cost-new-medicine#

Myers, M. B. (2016). Targeted therapies with companion diagnostics in the management of breast cancer: current perspectives. *Pharmacogenomics and Personalized Medicine*, *9*, 7–16. https://doi.org/10.2147/PGPM.S56055

Piccoli, S. P., & Sauer, J. M. for the Biomarker Assay Collaborative Evidentiary Considerations Writing Group, Critical Path Institute (C-Path). (2017). *Points to consider document: Scientific and regulatory considerations for the analytical validation of assays used in the qualification of biomarkers in biological matrices.* Retrieved from https://healthpolicy.duke.edu/sites/default/files/atoms/files/white_paper_6_14_17_0.pdf

Pfizer Inc. (2011). *XALKORI® (Critzotinib) Capsules, Oral: Highlights of prescribing information.* Retrieved from http://www.accessdata.fda.gov/drugsatfda_docs/label/2011/202570s000lbl.pdf

Shepherd, J. (1995). The West of Scotland Coronary Prevention Study: a trial of cholesterol reduction in Scottish men. *The American Journal of Cardiology*, *76*(9), 113C–117C.

Tao, L., Wang, W., Kramer, P. M., Lubet, R. A., Steele, V. E., & Pereira, M. A. (2004). Modulation of DNA hypomethylation as a surrogate endpoint biomarker for chemoprevention of colon cancer. *Molecular Carcinogenesis*, *39*(2), 79–84. https://doi.org/10.1002/mc.20003

Thomas D. W., Burns J., Audette J., Carroll C., Dow-Hygelund C., Hay C. (2016). Clinical development success rates (2006–2015). Retrieved from www.bio.org

U.S. Food and Drug Administration, Center for Drug Evaluation and Research. (2014). *Guidance for industry and FDA staff: Qualification process for drug development tools*. Retrieved from https://www.fda.gov/downloads/drugs/guidances/ucm230597.pdf

van Gool, A. J., Bietrix, F., Caldenhoven, E., Zatloukal, K., Scherer, A., Litton, J.-E., … Ussi, A. (2017). Bridging the translational innovation gap through good biomarker practice. Nature Reviews. Drug Discovery, 16(9), 587–588. doi:10.1038/nrd.2017.72

APPENDIX A: GLOSSARY

Term	Definition
Analytical Validation Phase	The third step in the development of a new biomarker that involves establishing that the performance characteristics of a biomarker are acceptable in terms of its sensitivity, accuracy, precision, and other relevant performance characteristics using a specified technical protocol.
Biomarker Prevalence	The proportion of the population of interest that is positive for the biomarker.
Clinical Utility	The conclusion that a given use of a medical product will lead to a net improvement in health outcome or provide useful information about diagnosis, treatment, management, or prevention of a disease. Clinical utility includes the range of possible benefits or risks to individuals and populations.
Clinical Validation and Utility Phase	The fourth and final stage of the biomarker development process that establishes that the assay acceptably identifies, measures, or predicts the clinical outcome of interest. The clinical utility of the biomarker is also established during this phase.
Companion Diagnostic	Companion diagnostics are FDA-approved diagnostic tests that are required for the safe and effective use of a corresponding therapeutic product; this requirement is stated within the therapeutic products' FDA-approved labels/package inserts. They are an important subset of predictive biomarkers. The term predictive biomarker will be used to signify the predictive biomarker/companion diagnostic category.
Context of Use	A statement that fully and clearly describes the way the medical product development tool is to be used and the medical product development-related purpose of the use.
Cost Component	The most granular representation of line item costs involved in the biomarker development and validation process. Cost components were identified from data gathered from interviews with industry SMEs or relevant industry data. Because cost component data was gathered from several sources, individual cost components may vary based on the data source due to differences in how the organization accounts for cost.
Cost Driver	A factor that influences the costs of developing and validating new biomarkers. To be considered a cost driver, the factor had to affect the development and validation costs of at least two biomarkers explored in the case studies by at least 15 percent in one of the four developmental phases.
Cost Element	Describes a common group of cost components. Because individual cost component data may vary by data source, cost components are mapped to standardized cost elements (e.g., Labor and Overhead, Materials, and Outsourced Services) to provide a common structure and taxonomy for costs and to facilitate direct cost comparisons and analysis.
Cost Impact	Describes the change in base expected biomarker development and validation costs due to cost driver effects. Cost impacts are incurred as an additional cost to base biomarker development and validation costs, and are given as the maximum of potential costs that could be incurred. The cost impact may be provided as a percentage change in base biomarker development and validation costs (e.g., a 10 percent cost impact means that the new expected cost for biomarker development and validation is 110 percent of base costs), or as a dollar amount (e.g., a $1 million cost impact means the new expected cost for biomarker development and validation is increased by $1 million).
Cost of Capital	Refers to the opportunity cost of making a specific investment. Cost of capital is the rate of return that could have been earned by using the same investment capital into a different investment with equal risk. The cost of capital is the rate of return required to persuade the investor to make a given investment.
Cost of Failure	Describes costs associated with a biomarker investment that did not successfully complete the biomarker development and validation process. Cost of failure includes both out-of-pocket and opportunity costs. Direct cost of failure can be discussed in terms of expected direct, realized direct, gross realized direct, and net realized direct cost of failure.
Cost Scenarios	Describes how a cost driver affects overall costs of biomarker development and validation for a given biomarker category; the scenarios describe the conditions or biomarker characteristics that would lead to a substantial increase in costs over average costs
Data Storage and Transfer	Includes costs for data storage and transfer. Data storage and transfer costs typically increase when novel or complex technologies that generate massive amount of data are required, such as next generation sequencing or imaging studies. On average, annual data storage and sample cost were assumed to be between $10,000 to $50,000, but can increase to be as much as $300,000.

Term	Definition
Equipment and Reagents	Includes the annual equipment and reagent costs required for biomarker development and validation. On average, annual equipment and reagent costs were assumed to be between $50,000 to $200,000, though in cases where specialized equipment or reagents are necessary, additional costs are likely to be incurred.
Expected Direct Cost of Failure	The theoretical or planned cost of failure based on historical biomarker investment performance. Expected direct cost of failure can be calculated from the biomarker development and validation budget and probability of failure.
Facilities and Administrative	Includes overhead costs such as facilities, administrative support, and other miscellaneous costs associated with the biomarker development and validation effort.
Gross Realized Direct Cost of Failure	The actual cost of failure based on the specific biomarker investment, not accounting for any residual benefit derived.
Identification and Feasibility Phase	The first step in the development of new biomarkers for drug development that involves identifying candidate biomarkers and assessing the feasibility of using the biomarker in drug development. Assessing the feasibility of the biomarker involves getting a better sense of the degree to which the candidate biomarker is indicative of the processes or endpoints it is intended to measure using preclinical animal models and clinical samples, when available.
Labor and Overhead	Cost element that includes all salaries and benefits; facilities and administrative support; and overhead, including resources such as legal and intellectual property.
Legal/IP	Includes costs associated with assessing legal and intellectual property considerations to ensure there are no potential legal restrictions on the use of the biomarker being developed. Licensing fees and restrictions on use may be involved in instances when the entity or group developing the biomarker does not hold all rights to a critical component of the biomarker assay, whether it be rights to the application of the biomarker itself or to a critical technology being used in the assay. Average legal costs were assessed to be between $200,000 to $400,000, though additional legal costs can be incurred if the patent is challenged.
Materials	Cost element that includes equipment maintenance and operational costs, reagents, sample acquisition, storage, shipping and handling costs, as well as data storage and transfer costs.
Net Realized Direct Cost of Failure	The actual cost of failure based on the specific biomarker investment, less any residual benefit gained.
Opportunity Cost	Refers to the loss of potential gain from other alternatives due to the decision to pursue the biomarker investment. Also known as time cost.
Outsourced Services	Cost element that includes any cost spent engaging a contract research organization (CRO) during the development and validation process.
Out-of-Pocket Cost	Describes the costs incurred for biomarker development and validation, not accounting for opportunity costs.
Phase	A distinct stage in the process of developing and validating a new biomarker for use in drug development. The Biomarker development and validation lifecycle referenced in this report is composed of four phases: (1) Identification and Feasibility; (2) Develop Assay and Intended Use; (3) Analytical Validation; and (4) Clinical Validation and Utility. The phases are described more fully in Exhibit 8.
Predictive Biomarker	Per the Biomarkers, Endpoints and other Tools Glossary that the U.S. Food and Drug Administration – National Institutes of Health Biomarker Working Group created, the definition of a “predictive biomarker” is “a biomarker used to identify individuals who are more likely than similar individuals without the biomarker to experience a favorable or unfavorable effect from a specific intervention or exposure.
Realized Direct Cost of Failure	The actual cost of failure for a specific biomarker investment. The realized direct cost of failure is the amount spent on the biomarker investment until activity stops and failure is declared
Retrospective-prospective RCT	A trial in which samples from a previously conducted RCT are used to validate a biomarker; while it can reduce the time and costs associated with clinical validation, it is dependent on the availability of an appropriate number of high-quality samples from a well-conducted clinical trial and thus not always possible.

Term	Definition
Salaries and Benefits	Includes costs for employee salaries and benefits, taxes, and other payroll costs. Industry average salary rates were converted to burdened labor rates by using two times annual salary as a basis for estimating the indirect costs associated with labor. Three labor categories were used – low, medium, and high. The low burdened rate (<$150,000 per year) accounts for positions such as laboratory technicians and research assistants. The medium burdened rate ($150,000-$300,000) accounts for positions such as statisticians, data analysts, project managers, regulatory affairs lead, principal investigators, and medical writers. The high burdened rate (>$300,000) accounts for highly specialized clinical staff, such as radiologists, oncologists, and pathologists. Different labor mixes were developed for the different stages of the biomarker development and validation process. In earlier stages, the labor mix was skewed toward low and medium category personnel, while in later stages the labor mix required more medium and high category personnel.
Surrogate Endpoint	An endpoint that is used in clinical trials as a substitute for a direct measure of how the patient feels, functions or survives; it does not measure the clinical benefit of primary interest in and of itself but rather is expected to predict the clinical benefit or harm based on epidemiological, therapeutic, pathophysiological, or other scientific evidence.
Sample Acquisition	Includes the amount to acquire samples. Sample acquisition costs are highly variable depending on the type of sample being acquired and commonness of the disease or condition. Sample costs can range from under $100 (for easily obtainable samples such as blood, saliva, and urine) to $10,000 per sample (for difficult to obtain samples for rare diseases). Average sample costs were assumed to be $1,000.
Sample Storage, Shipping and Handling	Includes the cost to transport and store samples. On average, specimen shipping and handling costs were estimated to be between $35,000 to $100,000 annually.
Targeted/Enrichment Design	A trial in which biomarker status is used as part of the selection criteria for participant inclusion in the trial.
Trial Recruitment	Trial Recruitment includes the costs for retaining a CRO. CRO costs play a major role in the Clinical Validation and Utility phase. CRO recruit patients and gather specimens and data for clinical trials. Seventy percent to 80 percent of clinical trial costs go to CROs to patient recruitment and sample acquisition, while 20 percent to 30 percent go to sponsor costs. This breakdown in costs was assumed to hold for all three trial designs, the standalone prospective randomized control trial (RCT), the prospective RCT alongside a therapeutic product, and the prospective-retrospective RCT.
Unit Cost	Describes the cost incurred by an organization for one unit of a particular product or service.

APPENDIX B: ACRONYMS

ALK	Anaplastic lymphoma kinase
CAPM	Capital asset pricing model
CLIA	Clinical Laboratory Improvement Amendments
CRO	Contract research organization
CSF	Cerebrospinal fluid
eGFR	Estimated glomerular filtration rate
EGFR	Epidermal growth factor receptor
FDA	U.S. Food and Drug Administration
FDG-PET	Fluorodeoxyglucose positron emission tomography
HCV	Hepatitis C virus
HLA-B*5701	Human leukocyte antigen-allele B*5701
IHC	Immunohistochemistry
IP	Intellectual property
LDL-C	Low-density lipoprotein cholesterol
MRI	Magnetic resonance imaging
NGS	Next-generation sequencing
NSCLC	Non-small cell lung cancer
PCR	Polymerase chain reaction
PD-L1	Programmed death-ligand 1
RCT	Randomized controlled trial
SME	Subject matter expert
WACC	Weighted average cost of capital

APPENDIX C: EXPERT CONSULTATIONS

Over the course of the project 29 experts were interviewed. The table below lists the organizations with which the experts consulted are currently affiliated as well as an overview of their areas of expertise. Each organization is only listed once regardless of how many experts were interviewed. In many cases, the input given was based on the expert's experience with many biomarkers and biomarker categories at different organizations. It is important to note that the information and opinions shared by experts during the course of the consultations were based on their own experiences and opinions, not those of the companies or organizations that employ them.

	Wide Spectrum Biomarker Development Perspective: Leader/Vision	Industry Experience	Process: Discovery → Validation, Regulatory	Technology Impact	Costs and Cost Drivers	Safety Biomarker	Companion Diagnostic/Predictive	Surrogate Endpoint	Prognostic Biomarker
INTERVIEWED EXPERTS									
Agilent Technologies	♦	♦	♦	♦	♦		♦		
Critical Path Institute, Predictive Safety Testing Consortium	♦	♦	♦	♦	♦	♦			
FDA Center for Drug Evaluation and Research			♦	♦	♦			♦	
GenePeeks Inc.	♦	♦		♦	♦		♦		
Genomic Health	♦	♦	♦	♦	♦				♦
Gilead Sciences	♦	♦	♦	♦	♦	♦	♦	♦	♦
GlaxoSmithKline	♦	♦	♦	♦	♦		♦		
Janssen Research & Development	♦	♦	♦	♦	♦	♦	♦	♦	♦
Memorial Sloan Kettering			♦	♦	♦			♦	
Merck and Co.		♦	♦	♦	♦	♦		♦	
National Cancer Institute, Cancer Imaging Program			♦	♦	♦			♦	
Opus Three LLC (Biomarker Consulting CRO)	♦		♦	♦		♦	♦		♦
Roche Molecular Systems	♦	♦	♦	♦	♦		♦		
Takeda	♦	♦	♦	♦	♦		♦	♦	
Toronto Western Hospital			♦	♦	♦			♦	
Tufts University			♦	♦	♦			♦	
University of Michigan	♦		♦	♦			♦	♦	♦
Yale University			♦	♦	♦		♦		

APPENDIX D: CASE STUDY SUMMARY

Biomarkers:	EGFR	PD-1/ PD-L1	HLA-B 57:01	eGFR	FDG-PET	HCV Viral Load
Knowledge Base						
Body of published literature sufficient	✓	✓	✓	✓	✓	✓
Prior case studies available	✓		✓			
Appropriate experts identified	✓	✓	✓	✓	✓	✓
Mechanism of action known	✓	✓	✓	✓	✓	✓
Stage in the validation process	FV	FV	FV	FV	FV	FV
Assay Characteristics						
Gene- or genomic-based	✓		✓			✓
Protein- or proteomic-based		✓				
Radiographic/imaging					✓	
Physiological (e.g., eGFR, blood pressure)				✓		
Single analyte	✓	✓	✓		✓	✓
Multiplexed/multi-analyte				✓		
Assay type (DQ/RQ/QQ/QL)	QQ	QL	QQ	RQ	SQ	QQ
Clinical Trial Characteristics						
Trial design (R/P, P-RCT, SA-P-RCT, COH, Multi)	R/P, P-RCT	R/P, P-RCT	SA-P-RCT	R/P	P-RCT	P-RCT
Disease Areas or Conditions						
Cancer	✓	✓			✓	
Cardiovascular						
Kidney				✓		
Rare diseases						
Infectious disease			✓			✓
Chronic				✓		✓
Acute	✓	✓	✓		✓	
Biomarker Uses						
Identify patients most likely to benefit from therapy	✓	✓	✓			
Identify patients most likely to experience an adverse event			✓			
Reduce the number of trial participants needed to detect statistically significant treatment effects				✓	✓	✓

AV = Analytical Validation phase; COH = longitudinal cohort study; CV = clinical validation phase; DA = Develop Assay and Intended Use phase; DQ = definitive quantitative; FV = fully validated; I&F = Identification and Feasibility phase; Multi = multiple types of trials used to support validation; P-RCT = prospective RCT; QL = qualitative; QQ = quasi-quantitative; R/P = retrospective-prospective RCT; RCT = randomized controlled trial; RQ = relative quantitative; SA-P-RCT = stand-alone prospective RCT. FV = Fully validated biomarker that has completed the four step validation process: 1) Identification and Feasibility, 2) Develop Assay and Intended Use, 3) Analytical Validation, 4) Clinical Validation and Utility.

APPENDIX E: PREDICTIVE BIOMARKER CASE STUDY INFORMATION

Exhibit 27 summarizes the estimated overall development and validation costs associated with the EGFR, HLA-B*5701, and PD-L1 biomarker assays as well as several key characteristics identified as important cost drivers. The findings from the case studies confirmed that these factors are important cost drivers that can significantly affect the overall costs of developing and validating predictive biomarkers.

Exhibit 27: Case Study Information for Predictive Biomarkers

Category	EGFR	HLA-B*5701	PD-L1
Unique biomarker characteristics	• PCR-based assay, with challenges arising from rapid technological evolution • First liquid biopsy (plasma-based) validated biomarker • Multiple iterations and expanded intended use for a biomarker	• Predictor of safety rather than of efficacy • No assay validation needed - used a preexisting assay • Biomarker developed as a post-marketing commitment rather than alongside a therapeutic agent	• Rapid development of a biomarker to support the rapidly evolving field of immune-oncology • IHC-based assay, with challenges standardizing the threshold and interpretation of results
Overall estimated development and validation costs[29]	$7 million to $10.5 million	$30 million to $42 million	$6.5 million to $14.25 million
Sample costs	$600 to $1,500 per sample, depending on specimen type	$500 to $3,000 per sample	$1,000 per sample
Trial design	Prospective RCTs, retrospective-prospective RCTs, and bridging studies	Stand-alone, prospective RCT	Prospective RCTs, retrospective-prospective RCTs, and bridging studies
Sample size needed for clinical validation	1,000 participants	1,800 participants	1,000-2,500 participants
Length of follow-up required to observe clinical outcome	10-12 months	6 weeks	20-27 months
Developmental Model (Private or Consortium)	Private	Private	Private
Key pain points	• Novel technology and the need to work through appropriate controls	• Lack of archived patient samples from clinical trials that could have been used for retrospective studies	• Firewall between assay development efforts led to lack of consistency among threshold and scoring criteria • Knowledge base for the role of the biomarker in health and disease not well characterized at the outset of biomarker development
Cost driver examples	• Need for new equipment and software because of technological advancements • Need for large number of samples because of low frequency of EGFR mutation • Need for plasma samples from cancer patients for controls and dilution matrix for plasma assay	• Stand-alone trial needed for clinical validation • Extensive planning and collaboration needed to design first-of-its-kind trial • Clinical validation study larger than strictly necessary because it was the first-of-its-kind	• Extensive background research required because the role of the biomarker was not well understood at the outset • Sample sizes for clinical validation were large, particularly when the biomarker was used as part of inclusion criteria

[29] Estimates determined using the cost methodology described in Section 3.3. These estimates are limited to the information gathered during the project and may not encompass all the costs of developing the biomarker.

APPENDIX F: SURROGATE ENDPOINTS CASE STUDY INFORMATION

Exhibit 28 summarizes the estimated overall development and validation costs associated with eGFR, FDG-PET, and HCV RNA viral load as well as several factors identified as important cost drivers. The findings confirmed that these factors are important cost drivers that can significantly affect the overall costs of developing and validating surrogate endpoints. The direct cost to develop two surrogate endpoints was significantly less than expected based on the category average findings. Experts pointed out that surrogate endpoint development often has substantial indirect development costs because the process relies on data from prior clinical studies or ongoing routine medical care. Development could not have occurred without these expensive clinical studies. In response to the experts' strong recommendation to provide a full understanding of the effort necessary to develop and validate a surrogate endpoint, cost ranges for biomarker development with and without 100 percent of clinical trial costs attributed to biomarker development are provided in Exhibit 28.

Exhibit 28: Case Study Information for Surrogate Endpoints

Category	eGFR	FDG-PET	HCV RNA Viral Load
Unique biomarker characteristics	• Equation incorporating demographic and analytic variables • Used mostly retrospective data	• Imaging biomarker • Required multisite trials for validation • Required a relatively large amount of highly skilled labor to develop and perform the assay	• Infectious disease viral load • Long-term follow-up to determine impact of virus elimination on clinical outcome
Overall estimated development and validation costs[30]	• $12 million to $15 million (direct) • $72 million to $85 million (with all clinical studies)	• $42 million to $51 million (direct)	• $11 million to $15 million (direct) • $31 million to $68 million (with all clinical studies)
Sample costs	• Majority used existing sample data, a small number of banked samples, or prospectively gathered samples for specific populations	• $2,000 to $3,000 per scan	• $200 per whole blood sample
Trial design	• Retrospective	• Prospective	• Multiple cohort studies, mainly retrospective, some prospective
Sample size needed for clinical validation	• 3,896 participants	• 180 participants	• 530 participants
Length of follow-up required to observe clinical outcome	• 10 years	• 1-5 years	• 2-10 years
Developmental Model (Private or Consortium)	• Consortium	• Consortium	• Private/consortium
Key pain points	• Researchers had to standardize the measurement of analytes • Data from many different population demographics were needed	• Recruitment was a challenge because of the time commitment and potential lack of therapeutic benefit • Ensuring adherence to a standardized protocol across sites during multisite trials required significant effort	• Follow-up to observe clinically meaningful endpoints was lengthy

[30] Estimates determined using the cost methodology described in section 3.3. These estimates are limited to the information gathered during the project and may not encompass all the costs of developing the biomarker.

Category	eGFR	FDG-PET	HCV RNA Viral Load
Cost driver examples	• Large sample sizes were required to determine the appropriate coefficients for demographic variables • Lengthy follow-up times of up to 10 years were required to link estimates of GFR with long-term clinical outcomes	• Recruitment challenges increased both the cost and the timeline for validating FDG-PET as a surrogate endpoint for new indications • A significant amount of highly skilled labor was required • The necessary technology was costly to purchase and maintain	• Trial design to follow cohorts of patients from multiple hospitals increased the time and cost required to validate HCV RNA viral load • Lengthy follow-up times of up to 10 years increased costs • The consortium of hospitals each maintaining its own database of patients increased the timeline and cost to allow for data aggregation and analysis

Text Only Version of Graphics

Exhibit 1: Project Methodological Overview

The overall methodology of the project included an environmental scan, case studies and cost framework analysis addressing the five research questions:

1. How are biomarkers identified for use in drug development?
2. What are the sources of data necessary for biomarker development?
3. How do the cost factors vary across the selected biomarker categories?
4. How do these factors contribute to the overall cost within a particular category?
5. How do the cost drivers compare across biomarker categories?

Published literature and expert consults were conducted to inform the Environmental Scan activity, producing the Environmental Scan Report which addressed Research Questions 1 – 3 across 4 biomarker categories: Predictive Biomarkers, Prognostic Biomarkers, Safety Biomarkers, and Surrogate Endpoints.

The Environmental Scan Report informed case study recommendations and the selection of case studies for the case study activity, which included 2 biomarker categories: Predictive Biomarkers and Surrogate Endpoints. The selected case studies, published literature, and expert consults informed the Case Studies activity, producing the Case Studies Report which addressed Research Questions 2 - 5.

The Case Studies Report and case study data informed the Cost Framework analysis activity, which included 2 biomarker categories: Predictive Biomarkers and Surrogate Endpoints, resulting in the production of the Cost Frameworks and Cost Framework Report which addressed Research Questions 3 - 5.

The Environmental Scan Report, the Case Studies Report, and the Cost Frameworks Report all informed the production of the Final HHS Briefing, The Final Biomarkers Report, and a manuscript for publication.

(return to page 7)

Exhibit 18: General Cost Percentages by Biomarker Category and Phase

For prognostic biomarkers and predictive biomarkers, experts broke the costs down into three steps:

1) Clinical Validation and Utility;
2) Develop Assay and Intended Use/Analytical Validation; and
3) Identification and Feasibility.

For Predictive Biomarkers, Clinical Validation and Utility was 50-60% of costs; Develop Assay and Intended Use/Analytical Validation was 20-30% of costs, and Identification and Feasibility was 20% of costs.

(Note: values shown as a range represent the minimum and maximum possible percentage of total costs for each development step.)

For Prognostic Biomarkers, Clinical Validation and Utility was 40-60% of costs; Develop Assay and Intended Use and Analytical Validation was 20-40% of costs, and Identification and Feasibility was 10-20% of costs.

For Safety Biomarkers and Surrogate Endpoints, experts classified the costs into two steps:

1) Clinical Validation and Utility; and
2) Identification and Feasibility/Develop Assay and Intended Use/ Analytical Validation.

For Safety Biomarkers, Clinical Validation and Utility was 70-80% of total costs, with 20-30% of costs going to the final three phases (Identification and Feasibility, Develop Assay and Intended Use, and Analytical Validation).

For Surrogate Endpoints, Clinical Validation and Utility was 90% of total costs, with 10% of costs going to the final three phases (Identification and Feasibility, Develop Assay and Intended Use, and Analytical Validation).

(return to page 45)