Facilitating Biomarker Development: Strategies for Scientific Communication, Pathway Prioritization, Data-Sharing, and Stakeholder Collaboration
October 27, 2015

Discussion Guide

Introduction
The emerging field of precision medicine continues to offer hope for improving health outcomes while also controlling the overall cost of health care. Precision medicine strategies also have the potential to drive efficiencies across all stages of the medical product development process, providing timely information on promising therapeutic targets and optimal doses, facilitating the identification of patients likely to respond well to treatment, and offering novel ways to predict and measure safety and efficacy at an earlier stage of development. Together, these strategies may enable time and cost savings through smaller, more focused clinical trials with a higher overall probability of success. However, the impact of this approach has been limited in practice due to the absence of valid and reliable biomarkers for many disease areas, in many cases driven by the limited scientific understanding of a disease’s pathogenesis.

A biomarker is a defined characteristic (e.g., a molecular, histologic, radiographic, or physiologic characteristic) that is measured as an indicator of normal biological processes, pathogenic processes, or responses to an exposure or intervention, including therapeutic interventions. This definition broadly captures a variety of biomarkers that serve several functions in the nonclinical and clinical settings of medical product development. Thanks to advances in scientific research and their associated technologies, the discovery of novel biomarkers has accelerated rapidly. However, the pace of discovery has largely outpaced the broader scientific and biomedical community’s ability to validate and develop the evidence base for their clinical utility. In addition, the data supporting the ever-increasing number of putative biomarkers has been challenged by reproducibility when defining this clinical context. Biomarker development and regulatory acceptance can be a time and resource-intensive process, with challenges that range from scientific and regulatory to logistical and cultural.

Under a cooperative agreement between the U.S. Food and Drug Administration (FDA), the Center for Health Policy at the Brookings Institution is convening this workshop to advance the discussion on biomarker development and to identify strategies that can help to address some of these ongoing challenges.

Accelerating biomarker development
The need for regulatory-accepted biomarkers has given rise to numerous legislative and policy initiatives that seek to address challenges in their development and use. Improving regulatory review of biomarkers was a key commitment for FDA under the fifth reauthorization of the Prescription Drug User Fee Act, and the 21st Century Cures Act currently making its way through
Congress calls on FDA to take a number of additional steps to facilitate the development and regulatory acceptance process.¹,²

Over the last several years, FDA has issued guidances on the development and use of biomarkers in medical product development, engaged the stakeholder community through public meetings, established a voluntary submission process for pharmacogenomic data, and, most recently, developed a formal stand-alone qualification program for biomarkers (described in greater detail below). The Critical Path Initiative—FDA’s national strategy for accelerating innovation in the scientific processes through which medical products are developed, evaluated, and manufactured—has also identified biomarker development as a high-priority area for future research and collaboration among stakeholders.

More broadly, a diverse range of consortia are working to accelerate the discovery and development of publically-accessible medical product development tools and methods, including better predictive tools, methods for clinical trials, procedures for biospecimen handing, and collective research resources (molecular libraries and tissue repositories). Most are government-initiated, though others have been launched by third-party groups, industry, academia, and patient advocacy groups.³ As many as one in four are involved in biomarker research.⁴ The emergence of the consortia model indicates that stakeholders are increasingly becoming more open to sharing data, knowledge, resources, and capital to address pressing needs in biomedical research, especially in the area of biomarkers.

However, several barriers to collaboration and biomarker development remain, including the lack of a consistent set of definitions and taxonomy of biomarkers and their uses, uncertainty among sponsors regarding the appropriate pathway for achieving regulatory acceptance for a new biomarker, the lack of generally accepted evidentiary standards for qualifying novel biomarkers, and inadequate coordination of limited public and private resources available to support biomarker development and regulatory acceptance.

Developing a standard glossary of terms in biomarker development
The use of biomarkers across different settings and disciplines has naturally resulted in a great deal of variability in biomarker terminology, which has been recognized as a major barrier to their development and regulatory acceptance.⁵,⁶ In addition to hindering collaboration among stakeholders on important issues in biomarker development, inconsistent definitions also complicate downstream decision-making related to clinical use and reimbursement.⁷ Establishing a foundation of shared, well-accepted definitions for biomarkers and their various uses would go a long way towards improving collaboration among stakeholders who are working to address barriers to biomarker development. In particular, a consistent set of definitions could facilitate the process of defining specific contexts of use for biomarkers in medical product development and regulation, which could in turn help build consensus on the evidentiary standards that would need to be met.⁸

As a first step in driving this process, a joint working group of representatives from FDA and the National Institutes of Health (NIH) has begun developing a set of definitions that can serve as a
starting point for broader standardization. These terms are listed in Appendix 1 along with their proposed definitions. It is anticipated that this lexicon will be refined over time through consultation with various stakeholders. The joint working group welcomes feedback on these terms as well as proposals for additional terms that should be considered for inclusion in the future.

**Determining the appropriate pathway for biomarker development and regulatory acceptance**

There are generally three pathways through which a biomarker may receive regulatory acceptance for use in medical product development. The first is through general acceptance by the clinical, scientific, and regulatory communities, which usually occurs over an extended period of time as information accrues organically through scientific research. The second involves the traditional marketing authorization process for medical products (i.e., the Investigational New Drug/New Drug Application pathway for drugs and biologics, or the Investigational Device Exemption/Premarket Approval pathway for devices), in which a sponsor may engage with FDA to reach agreement on the use of a biomarker in a given development program. The third involves qualification through the recently launched Biomarker Qualification Program. Qualification is defined as “a conclusion that within the stated context of use (COU), a medical product development tool (MPDT) can be relied upon to have a specific interpretation and application in medical product development and regulatory review.” Once a biomarker has been fully qualified, it can be used in multiple medical product development programs without the need to collect additional data to support its use – as long as that use is consistent with the specific context and purpose for which it was first qualified, and that no new scientific information conflicts with its original intended use.

Emerging regulatory experience with these various pathways suggests that some types of biomarkers are more appropriate for one pathway over another. However, this choice is not always clear, particularly early on in the development stages. Developing criteria to assist in identifying the appropriate pathway for regulatory acceptance could facilitate biomarker development and help prioritize limited resources.

Towards that end, this workshop will include discussion of two biomarker case studies: Total Kidney Volume (TKV) and Epidermal Growth Factor Receptor (EGFR) in lung cancer. (See Appendix 2). TKV was recently qualified for use as a prognostic enrichment biomarker in studies for the treatment of autosomal-dominant polycystic kidney disease, while EGFR–related biomarkers (mutation status, protein expression) have been included as predictive biomarkers of drug response in a range of individual drug development programs and in the labels of several FDA-approved drugs that treat non-small cell lung cancer (NSCLC), including erlotinib, gefitinib, and afatinib. These two cases provide contrasting examples of the pathways towards regulatory acceptance and will serve to highlight essential elements that affect the feasibility and likely value of a biomarker development effort, including the context of use and the quality and availability of data.
Improving collaboration to advance biomarker development

Advancing the development of biomarkers will require coordinated partnerships involving a broad array of stakeholders including FDA, NIH, industry, academia, patients, and payers. There are several examples of successful public-private partnerships focused on biomarker qualification (indeed, four of the six qualified biomarker packages were submitted by consortium-led groups.) However, communication and coordination across these various efforts is limited, and there are several key barriers that further inhibit their progress.

Developing biomarker knowledge: improving data standardization and sharing

One of the greatest challenges in the development of biomarkers is the need for high-quality and robust packages of data that can support their use in drug development and clinical practice. This challenge is further complicated by the wide range of data types (e.g., preclinical, clinical, registry data) that need to be integrated to generate evidence for a biomarker and by the variety of platform technologies used at different institutions. These data often come in disparate formats and are typically proprietary. Consortia can play a critical role in promoting the harmonization of approaches to data collection, aggregation, and sharing to maximize the overall utility of data contributed by multiple organizations. Many data-sharing initiatives currently underway have worked towards this goal, although many have been organized to accomplish very different objectives. Strategies are needed to leverage existing data sharing initiatives towards biomarker development specifically.

Developing strategies for the prioritization of limited resources: cross-sector and cross-consortia collaboration

Several public-private partnerships, consortia, and advocacy groups are active in the area of biomarker development. Additional efforts, however, are needed to identify potential areas for cross-consortia collaboration to maximize the use of limited resources and harness the unique capabilities of different organizations. The successful collaboration between the Critical Path Institute’s Predictive Safety Testing Consortium (PSTC), the Innovative Medicines Initiative’s Safer and Faster Evidence-based Translation Consortium (SAFE-T), and the Biomarkers Consortium has highlighted some of the strategic benefits of cross-consortia collaboration. The PSTC developed some of the data and evidence that supported the qualification of preclinical kidney injury biomarkers, which are also being evaluated in the clinical setting in trials conducted by SAFE-T and the Biomarkers Consortium.

However, this level of collaboration has been the exception rather than the norm. The idea of a broad-based umbrella consortium—first alluded to in the President’s Council of Advisors on Science and Technology report on improving innovation in drug development and review—could be a potential approach to encouraging high-level collaboration across consortia. More specifically, an overarching consortium could take on several key roles including: 1) coordinating existing partnerships and consortia so that they effectively direct efforts to the development and regulatory acceptance of high priority biomarkers identified by FDA and the scientific community; 2) developing and maintaining the infrastructure for biomarker data collection and curation; 3) conducting reviews and making recommendations to FDA on the adequacy of data packages submitted by sponsors; 4) supporting biomedical research needed
to advance the discovery and development of new biomarkers.\textsuperscript{16} In addition to biomarker development, such a consortium could help to promote the use of common biomarker terminology and serve as a forum for the broader community to reach consensus on an evidentiary framework for biomarkers and their diverse uses. However, further discussion is needed to elucidate the mission of such an over-arching consortium, its goals, governance and organizational structure, approaches to developing effective communication between member consortia, and mechanisms for ensuring sustainability.

**Meeting overview and objectives**

The objectives for this expert workshop are to: 1) discuss the common lexicon developed by FDA and NIH for the field of biomarker development; 2) use case studies to explore biomarker characteristics (including Context of Use) that can inform which biomarker development pathway is the most appropriate; and 3) develop an initial set of strategies that can help to ensure better cross-sector collaboration and communication in the area of biomarker development, including strategies for improving the standardization, aggregation, and dissemination of biomarker data.

**Session I: Developing a Standard Glossary of Terms in Biomarker Development**

*Objective:* Varied specification and irregular use of biomarker-related terminology have hindered progress in the field of biomarker development. In order to achieve greater clarity and facilitate more effective collaboration, FDA and NIH have partnered to develop a common lexicon. This lexicon will be circulated in advance to attendees, and a representative of the joint working group will present on the history and purpose of their development, as well as a select few of the terms that are most relevant to the day’s discussion. This session will include a brief discussion of the definitions and will focus on targeted questions related to the dissemination and acceptance of those definitions by the broader community. Participants will be able to submit written feedback on those definitions after the meeting. Discussion questions will include:

- What strategies can FDA and NIH pursue to encourage broad adoption that these definitions are: 1) acceptable to the community and 2) used widely in medical product development, clinical care, and research?
- Are there any major gaps in the lexicon? Specifically, are there any important terms that have not been included that should be?
Session II: Determining the Appropriate Pathway for Biomarker Development and Regulatory Acceptance

Objective: Uncertainty about when, why, and how a biomarker should be developed to support regulatory acceptance often contributes to confusion. This session will begin with a presentation from FDA on the three pathways for biomarker development. The session will then use case studies to highlight essential elements that affect the feasibility and likely value of a given biomarker development effort, including the Context of Use and the quality and availability of data. The first case study is TKV as a prognostic marker for polycystic kidney disease, which was recently qualified by FDA; the second case study is EGFR status as a predictive marker for EGFR-targeted therapy in NSCLC, which has been used in several individual drug development programs. Discussion questions will include:

- What are the implications of choosing one development pathway over another? To what extent are sponsors able to pursue both pathways?
- Are there examples of multiple stakeholders working together to address an identified biomarker need, but which exist outside of the qualification space?
- What are some of the benefits and challenges associated with each of the pathways?
- To what extent does qualification save time and effort for future medical product development programs, and to what extent have new individuals been able to more readily use biomarkers because of the evidence provided in someone else’s qualification package?
- What are the advantages and disadvantages of pursuing a relatively narrow scope for Context of Use as opposed to a broader scope?
- What broad types of Context of Use (for example, enrichment, patient stratification, surrogate endpoint, safety endpoint, etc.) would be considered high-value?

Session III: Strategies for Improving Data Standardization and Sharing

Objective: One of the key challenges in the field of biomarker development is the issue of data sharing (e.g., due to disaggregated data, differing data standards, and proprietary concerns). This session will provide an opportunity to: 1) discuss the main barriers to biomarker data sharing, including issues related to standardization, aggregation, and dissemination, and 2) identify possible strategies to address those barriers. Discussion questions will include:

- What are the main barriers to broader biomarker data aggregation and dissemination?
- How and to what extent can biomarker data from individual medical product development programs be shared more broadly?
- How and to what extent can data from ongoing qualification programs be shared more broadly?
- What incentives can FDA and other stakeholders (e.g., existing biomarker consortia) use to encourage data sharing?
Session IV: Facilitating Collaboration and Cross-Sector Communication

Objective: There are a number of organizations, partnerships, and consortia that are working to develop biomarkers in specific areas. However, their efforts are not well-coordinated, and there is a need to identify approaches to collaboration and communication that can: 1) help to identify and prioritize areas of highest unmet need in the field, and 2) ensure that consortia and other key stakeholders are working collaboratively and sharing critical information. This session will serve to identify and explore possible approaches to achieving these goals in the short and long-term. Discussion questions will include:

- How best to identify and communicate priorities for biomarker development.
- How best to move forward with establishing an evidentiary framework for the regulatory acceptance of biomarkers for various Contexts of Use.
- Lessons learned from the successful collaboration between the PSTC/SAFE-T/Biomarkers Consortia that may be applied to future efforts.
- Need for and feasibility of developing an “uber consortium” that could help to address these major issues.
- Possible scope and role for such an “uber consortium.”
References

   Goals and procedures fiscal years 2013 through 2017. Retrieved from –
   congress/house-bill/6/text
   and Development. Washington, DC. Retrieved from http://consortiapedia.fastercures.org/assets/PDF/45700-
   ConsortiaReport.pdf
   242cm6.
   Multiple Stakeholder Framework for Biomarker Development, Regulatory Acceptance, and Utilization. Clinical
   Pharmacology & Therapeutics, 98(1), 34-46.
   Medicine—Payer and Manufacturer Perspectives and Implications for Health Economics and Outcomes Research:
   Multiple Stakeholder Framework for Biomarker Development, Regulatory Acceptance, and Utilization. Clinical
   Pharmacology & Therapeutics, 98(1), 34-46
   Development Tools. Retrieved from
   Volume in Studies for Treatment of Autosomal Dominant Polycystic Kidney Disease. Retrieved from
11. Table of Pharmacogenomic Biomarkers in Drug Labeling. (05/20/2015). Retrieved from
12. Table of Qualified Biomarkers and Supporting Information. (08/20/2015 ). Retrieved from
13. C-PATH and IMI 2nd Annual Meeting. (2014). Session 3: Maximizing the Value of Data Shared by Multiple
   Organization. Retrieved from
   Multiple Stakeholder Framework for Biomarker Development, Regulatory Acceptance, and Utilization. Clinical
   Pharmacology & Therapeutics, 98(1), 34-46
15. Executive Office of the President. President’s Council of Advisors on Science and Technology. (2012). Report to
   the President on Propelling Innovation in Drug Discovery, Development, and Evaluation. Retrieved from
   https://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-fda-final.pdf
   Multiple Stakeholder Framework for Biomarker Development, Regulatory Acceptance, and Utilization. Clinical
   Pharmacology & Therapeutics, 98(1), 34-46
Appendix A: Draft Biomarker Lexicon under development by joint FDA-NIH working group

Harmonization of terms used in translational science and medical product development, particularly those related to the study of biomarkers and endpoints, was identified as a priority need in spring 2015 by the FDA-NIH Joint Leadership Council. Lack of clear definitions and inconsistent usage of key terms can interfere with effective communication about medical product development programs and interpretation of evidence and therefore hinder efficient translation of promising medical discoveries to approved medical products. With the goal of improving communication, aligning expectations, and improving scientific understanding, the two agencies are developing a consensus glossary showing important terms, definitions, hierarchy, and inter-relationships.

The initial focus of the glossary aims to capture the distinction between biomarkers (BM) and clinical outcome assessments (COA) and to delineate their various roles in biomedical research and medical product development. Given the goal of broad applicability to multiple stakeholder communities, any set of definitions must account for the meaning of terms that have been variably used to date. The intent of the glossary is to capture important concepts so they can be used consistently recognizing that some of the terms have acquired nuanced and situation-specific interpretations. However, as can be expected, this is a challenging process. Therefore, during session I of the meeting, we plan to discuss the intersection of three important terms: susceptibility/risk biomarker, prognostic biomarker, and predictive biomarker, to illustrate the difficulty in adequately defining these terms that are used broadly in many settings and to get feedback on their appropriate application. In addition, we hope to also discuss the challenge of capturing unique concepts of the biomarker subtypes.

This glossary is meant to be a “living” resource that will be updated over time with additional terms and clarifying information. In the near future, an updated version will include additional terms, as well as specific examples.

Initial defined terms relevant to today’s discussion include:

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. Molecular, histologic, radiographic, or physiologic characteristics are examples of biomarkers. A biomarker is not an assessment of how a patient feels, functions, or survives. Types of biomarkers include:
   a. **Susceptibility/Risk biomarker** - A biomarker that indicates the risk for developing a disease or sensitivity to an exposure in an individual without clinically apparent disease.
   b. **Diagnostic biomarker** - A biomarker used to identify individuals with the disease or condition of interest or to define a subset of the disease.
   c. **Monitoring biomarker** - A biomarker used to detect a change, over time, in the degree or extent of disease, safety indicator, or exposure.
d. **Prognostic biomarker** - A biomarker used to identify likelihood of a clinical event, disease recurrence or progression.

e. **Predictive biomarker** - A biomarker used to identify individuals who are likely to experience a favorable or unfavorable effect from a specific intervention or exposure.

f. **Pharmacodynamic biomarker** - A biomarker used to show that a biological response has occurred in an individual who has received an intervention or exposure.

g. **Safety biomarker** - A biomarker used to monitor toxicity.

2. **Surrogate endpoint** - An endpoint that is used in clinical trials as a substitute for a direct measure of how a patient feels, functions, or survives. A surrogate endpoint does not measure the clinical benefit of primary interest in and of itself, but rather is expected to predict that clinical benefit or harm based on epidemiologic, therapeutic, pathophysiologic, or other scientific evidence.

From a U.S. regulatory standpoint, surrogate endpoints can be characterized by the level of clinical validation: validated, reasonably likely, and candidate.

a. A **validated surrogate endpoint** is supported by a clear mechanistic rationale and clinical data providing strong evidence that an effect on the surrogate predicts a clinical benefit. A validated surrogate endpoint is a predictor of clinical benefit and therefore can be used to support traditional approval without the need for additional efficacy information.

b. A **reasonably likely surrogate endpoint** is supported by clear mechanistic and/or epidemiologic rationale but insufficient clinical data to show that it is a validated surrogate. Such endpoints can be used for accelerated approval. In this case, additional trial data, assessing the effect of the intervention on the clinical benefit endpoint of interest will be collected in the post-marketing setting to verify whether an effect on the reasonably likely surrogate actually predicts clinical benefit in the specific context under study.

c. A **candidate surrogate endpoint** is still under evaluation for its ability to predict clinical benefit.

3. **Endpoint** - A precisely defined variable intended to reflect an outcome of interest that is statistically analyzed to address a particular research question. A precise definition of an endpoint typically specifies the type of assessments made, the timing of those assessments, the assessment tools used, and possibly other details, as applicable, such as how multiple assessments within an individual are to be combined.

4. **Validation** – Establishing that the performance of a test, tool, or instrument is acceptable for its intended purpose. Elements of validation include but are not limited to the following:

   a. **Analytical validation** - Establishing the performance characteristics of a test, tool, or instrument in terms of its sensitivity, specificity, accuracy, precision, and other relevant performance characteristics using a specified technical protocol (which may include specimen collection, handling and storage procedures) to demonstrate that the item
delivers reliable and reproducible measurements under stated conditions. This is validation of the test’s, tool’s, or instrument’s technical performance, but is not validation of the item’s usefulness.

b. **Clinical validation** - Establishing the extent to which a test, tool, or instrument identifies, measures, or predicts the presence or severity of a clinical characteristic, condition or predisposition in an individual. What is to be validated depends on its purpose.

5. **Context of use (COU)** - A statement that fully and clearly describes the way the medical product development tool (MPDT) is to be used and the medical product development-related purpose of the use. The COU defines the boundaries within which the available data adequately justify use of the MPDT and describes all important criteria regarding the circumstances under which the MPDT is qualified.

6. **Intended use** - The specific clinical circumstance or purpose for which a medical product or test is being developed. The description should include what is being measured and in what clinical context. In the regulatory context, “intended use” refers to the objective intent of the persons legally responsible for the labeling of medical products.¹

7. **Fit-for-Purpose** - A conclusion that the level of validation associated with a medical product development tool (MPDT) is sufficient to support its context of use.

8. **Qualification** – In a regulatory context, a conclusion that within the stated context of use (COU), a medical product development tool (MPDT) can be relied upon to have a specific interpretation and application in medical product development and regulatory review.

9. **Accelerated approval** - Regulatory mechanism by which new drugs² meant to treat serious, life-threatening diseases that provide meaningful therapeutic benefit to patients over existing treatments can be approved on the basis of adequate and well-controlled clinical trials establishing that the drug has an effect on a reasonably likely surrogate endpoint or on the basis of an effect on a clinical endpoint other than survival or irreversible morbidity (intermediate clinical endpoint). Postmarketing confirmatory trials have been required to verify and describe the anticipated effect on irreversible morbidity or mortality (IMM) or other clinical benefit.

---

¹ 21 CFR 201.128
² References to drugs or drug products include both human drugs and biological drug products regulated by CDER and CBER unless otherwise specified.
Terms currently in the process of being defined:

- Assay
- Assessment/Measurement
- Clinical benefit
- Clinical outcome
- Clinical outcome assessments (COAs)
  - Clinician-reported outcome (ClinRO)
  - Observer-reported outcome (ObsRO)
  - Patient-reported outcome (PRO)
  - Performance outcome (PerfO)
- Clinical utility
- Concept
- Construct validation
- Content validation
- Criterion validation
- Intermediate clinical endpoint
- Medical product development tool (MPDT)
- Outcome
- Outcome assessment
- Target entity
- Test, Tool, or Instrument
Appendix B: Case Study: Total Kidney Volume as a Prognostic Biomarker for Polycystic Kidney Disease

Introduction
Autosomal Dominant Polycystic Kidney Disease (ADPKD) is the most common form of hereditary kidney disease and the fourth most common cause of kidney failure in adults affecting more than 12 million individuals worldwide. The disease is characterized by progressive enlargement of the kidneys due to the formation and growth of cysts progressing to end-stage renal disease (ESRD) in many cases. There are no approved therapies in the United States for ADPKD although improved understanding of the molecular biology of the disease has resulted in a variety of experimental therapies aimed at slowing or stopping disease progression. In patients with ADPKD, the decline in renal function is preceded by an increase in total kidney volume (TKV), and, according to the published literature, TKV strongly predicts future loss of kidney function. These observations led to interest in the use of TKV as a prognostic enrichment biomarker in clinical trials of ADPKD.

Brief development history
A longitudinal study published in 2002 demonstrated that in patients with ADPKD, a decline in renal function as measured by glomerular filtration rate (GFR), correlated with an increase in renal volume assessed using ultrasound. Another study was launched in 2001 by the Consortium for Radiologic Imaging Studies of Polycystic Kidney Disease (CRISP) to develop standardized imaging techniques that could accurately and reliably measure kidney and cyst volume in patients with ADPKD; and to follow patients longitudinally for a 3-year period (subsequently extended for another 5 years). CRISP demonstrated: that total kidney and cyst volume increased in patients with ADPKD; that baseline TKV predicted the subsequent rate of increased kidney volume; that increased TKV was associated with reduced GFR; and that there was a substantial lag between the increase in TKV and decline in GFR. CRISP also showed that kidney and cyst volumes are the strongest predictors of declining renal function.

In 2007, the Polycystic Kidney Disease Outcomes Consortium (PKDOC) was established under the leadership of the Critical Path Institute (C-Path) bringing together the PKD foundation, the Clinical Data Interchange Standards Consortium (CDISC), four academic institutions, and three pharmaceutical companies with the aim of developing tools to accelerate the development of treatments for PKD. The Consortium set as its initial goal to improve the efficiency of ADPKD clinical trials by qualifying TKV as a prognostic biomarker for enriching ADPKD clinical studies. The consortium’s ultimate goal is to qualify TKV as a surrogate endpoint to establish the effectiveness of therapies for ADPKD.

Given that developing the evidence required to support qualification would require integrating data from multiple studies, the PKDOC worked with CDISC to develop PKD-specific clinical data standards. This enabled aggregation of data from two longitudinal studies and three patient registries from the University of Colorado-Denver, the Mayo Clinic, and Emory University, which were used to develop a predictive model linking baseline TKV, in combination with age and baseline estimated GFR (eGFR) with declining renal function. These cohorts included a combined total of 2355 patients ranging in age from 0-84 years with at least one TKV measurement by at least one of three modalities. After exclusion of patients with missing data, 1140 patients remained.
<table>
<thead>
<tr>
<th><strong>Biomarker Name</strong></th>
<th>Total Kidney Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Biomarker Classification</strong></td>
<td>Prognostic biomarker for Polycystic Kidney Disease</td>
</tr>
<tr>
<td><strong>Intended/defined context(s) of use</strong></td>
<td>Total Kidney Volume, measured at baseline, will be used as a prognostic enrichment biomarker in conjunction with patient age and baseline estimated glomerular filtration rate (eGFR) to identify individuals with Autosomal Dominant Polycystic Kidney Disease (ADPKD) who are at greater risk for a substantial decline in renal function.</td>
</tr>
<tr>
<td><strong>Pathway to Regulatory Acceptance</strong></td>
<td>A draft guidance was issued in August 2015 for the qualification of TKV as a prognostic biomarker for enriching clinical trials with patients with ADPKD at greater risk for a substantial decline in renal function.</td>
</tr>
<tr>
<td><strong>Sponsor(s)</strong></td>
<td>Polycystic Kidney Disease Outcomes Consortium (PKDOC), one of 10 consortia of the Critical Path Institute.</td>
</tr>
<tr>
<td><strong>Description of tests</strong></td>
<td>TKV should be calculated from the left and right kidneys measured with a validated and standardized image acquisition and analysis protocol within the trial. TKV can be measured by MRI, CT, or ultrasound imaging.</td>
</tr>
</tbody>
</table>
In March 2014, the PKDOC submitted a final briefing book to support qualification of TKV as a prognostic biomarker in clinical trials of ADPKD. Based on its analyses, the consortium proposed the following context of use for TKV for clinical trial enrichment in patients with ADPKD:

Baseline TKV can be applied as a prognostic enrichment biomarker that, in combination with patient age and baseline eGFR, can be used to help identify those ADPKD patients who are at the greatest risk for a substantial decline in renal function defined as (1) 30% worsening of eGFR, (2) 57% worsening of eGFR (equivalent to doubling of serum creatinine), or (3) End-Stage Renal Disease (ESRD, defined as dialysis or transplant).

According to the PKDOC’s analyses, the probability of reaching a 30% worsening of eGFR at three years of follow up was greater in patients with larger TKV (≥ 1L) compared to those with smaller TKV. Their analyses also showed that age, baseline eGFR, and log-transformed baseline TKV were independently associated with decline in eGFR. The consortium developed a multivariate model based on predicted probabilities of a 30% decline in eGFR to select patients for an early stage clinical trial of a preventive therapy. They performed a similar analysis for the probability of reaching a 57% worsening of eGFR as a means of identifying patients for a disease progression trial of a treatment for reducing complications of the disease as well as the probability of reaching the endpoint of end-stage renal disease (ESRD), i.e., start of dialysis or kidney transplant, in order to select patients for a treatment designed to reduce progression to ESRD.

As part of its review, the FDA performed its own analyses of the submitted data. FDA’s conclusion was that, relative to a model that did not include log (TKV), a fitted survival model including log (TKV) improved the predictive performance of event risk for a confirmed 30% decline in eGFR. The improvement was observed using: (1) the consortium’s data alone for model development and cross validation and (2) using clinical trial data that were available internally to FDA for independent validation. According to the FDA review, for the endpoints of 57% decline in eGFR and ESRD, there were too few events over the timeframe of a feasible clinical trial to perform meaningful analyses. The FDA also performed analyses to determine the impact of using the best fit model with TKV on the number of patients needed to produce one event and the number that would need to be screened. These analyses supported the utility of using the model with TKV to enrich the trial population. The analyses also suggested that using a multivariate risk score to enrich the trial population was more efficient than specifying independent entry criteria for the parameters of interest.

Based on its analyses, FDA issued a draft guidance containing qualification recommendations for the use of TKV, measured at baseline, as a prognostic enrichment biomarker to select patients with ADPKD at high risk for a progressive decline in renal function (defined as a confirmed 30% decline in the patient’s estimated glomerular filtration rate (eGFR)) for inclusion in interventional clinical trials. The draft guidance, issued in August 2015, states that baseline TKV can be used in combination with a patient’s age and baseline eGFR as an enrichment factor in clinical trials to select ADPKD patients at high risk for a progressive decline in renal function. The draft guidance states that TKV should be calculated from the left and right kidneys measured with a validated and standardized image acquisition and analysis protocol within the trial.
Lessons Learned from the development process

- The qualification effort utilized a fair amount of resources (both on the submitter end and on the FDA side); perhaps the greatest benefit of the exercise is that it quantified the amount of information that “was added” by using TKV to enrich the trial population.
- Registry data can be critical for establishing the value of a biomarker as a tool in drug development, but there are challenges associated with using and interpreting registry data.
- Biomarker qualification packages are based on the totality of data available to the submitter; however, sometimes FDA has access to other large datasets (i.e., data from drug development programs) that speak to the utility of a biomarker. It is unclear when and how we should use these sources of information to confirm the utility of a biomarker for a proposed context of use.
- Data standardization is essential to enable aggregation of data from multiple datasets.
- We need to consider whether the outcome of biomarker qualification should be a quantitative model describing how the biomarker and other factors influence the outcome of interest. Such a model would incorporate uncertainties in the parameters of the model and in the resulting predictions. The model output should have direct applicability to the intended use. For example, for a biomarker prognostic for clinical events, outputs should inform the numbers needed to screen and to enroll in order to achieve a single outcome event.
References


Appendix C: Case Study: EGFR mutation status as a predictive marker for EGFR-targeted therapy

Introduction
Lung cancer is the leading cause of cancer deaths in the U.S. with an estimated number of new cases of over 220,000 and approximately 160,000 deaths in 2012. Approximately 85% of cases are non-small cell lung cancer (NSCLC), the majority of which present as advanced disease (stage IIIb or IV) at the time of diagnosis. The median survival of patients with advanced NSCLC with supportive care is approximately three to six months. Standard systemic treatment for patients with advanced NSCLC in an unselected population consists of platinum-based doublet chemotherapy with response rates of approximately 30% and a median survival of approximately ten months. Epidermal growth factor receptor (EGFR), also known as HER1, belongs to the ErbB or human epidermal receptor family of tyrosine kinase growth factor receptors; EGFR signaling mediates tumor proliferation, invasion, metastasis, resistance to apoptosis, and angiogenesis. EGFR mutations are present in approximately 10% of NSCLC patients in the U.S. Preclinical studies have shown that certain EGFR mutations are oncogenic and can transform both fibroblasts and lung epithelial cells in the absence of exogenous epidermal growth factor promoting tumor formation in immunocompromised mice. The most common EGFR-activating mutations are deletions in exon 19 (45%) and a point mutation (L858R) in exon 21 (40%–45%). Patients with EGFR-activating mutations are more likely to have distinct clinicopathologic features such as female sex, never or light smokers, Asian origin, and adenocarcinoma histology. The presence of EGFR activating mutations is highly predictive of tumor responses of large magnitude and duration to EGFR TKIs.

Brief development history
Companion diagnostic assays are required by the FDA when the identification of a specific biomarker is needed to ensure the safe and efficacious administration of a drug. The typical approval pathway involves parallel development of the targeted therapy and the diagnostic assay. However, in the case of EGFR mutation status as a biomarker for NSCLC, a somewhat circuitous path to regulatory approval was followed since EGFR TKIs had been approved prior to the realization that these mutations could serve as predictive biomarkers for sensitivity to the drugs. The steps to approval of biomarker assays for EGFR mutations thus included:

- Approval of EGFR TKIs for all patients with advanced, chemotherapy-resistant NSCLC.
- Demonstration that tumors from patients who responded to EGFR TKIs harbored specific mutations in the EGFR gene.
- Demonstrated effectiveness of EGFR TKI therapy as first-line treatment in patients with NSCLC
- A retrospective analysis showing that those patients harboring EGFR mutations experienced prolonged progression free survival (PFS) in response to EGFR TKI therapy compared to those lacking EGFR mutations.
- Development of diagnostic test kits to rapidly identify EGFR mutations in tumor tissue with high degrees of sensitive and specificity.
- Bridging study that demonstrated the utility of companion diagnostic tests as a means of identifying patients with tumors bearing EGFR mutations.
- Approval of EGFR mutation tests as companion diagnostics to approved EGFR TKIs for first-line therapy of NSCLC.
<table>
<thead>
<tr>
<th><strong>Table 1: Summary information of epidermal growth factor receptor</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Biomarker Name</strong></td>
</tr>
<tr>
<td><strong>Biomarker Classification</strong></td>
</tr>
<tr>
<td><strong>Intended/defined context(s) of use</strong></td>
</tr>
<tr>
<td><strong>Pathway to Regulatory Acceptance</strong></td>
</tr>
</tbody>
</table>
| **Sponsor(s)**                                               | The cobas® EGFR mutation test (Roche Molecular Systems) was approved as a companion diagnostic for erlotinib (Tarceva®, Astellas Pharma Inc.) for first line treatment of metastatic NSCLC whose tumors have EGFR exon 19 deletions or exon21 (L858R) substitutions mutations.  
The therascreen® EGFR RGQ PCR kit (Qiagen) was approved as a companion diagnostic for afatinib (Gilotrif®, Boehringer Ingelheim) and gefitinib (Iressa®, Astra Zeneca) for first-line treatment of metastatic NSCLC in patients with EGFR mutations. |
| **Description of tests**                                     | The cobas® EGFR Mutation Test is a real-time PCR test for the qualitative detection of exon 19 deletions and exon 21 (L858R) substitution mutations of the epidermal growth factor receptor (EGFR) gene in DNA derived from formalin-fixed paraffin-embedded (FFPET) human non-small cell lung cancer (NSCLC) tumor tissue. The test is intended to be used as an aid in selecting patients with NSCLC for whom Tarceva® (erlotinib), an EGFR tyrosine kinase inhibitor (TKI), is indicated.  
The therascreen® EGFR RGQ PCR Kit is a real-time PCR test for the qualitative detection of exon 19 deletions and exon 21 (L858R) substitution mutations of the epidermal growth factor receptor (EGFR) gene in DNA derived from formalin-fixed paraffin-embedded (FFPE) non-small cell lung cancer (NSCLC) tumor tissue. The test is intended to be used to select patients with NSCLC for whom GILOTRIF® (afatinib) or IRESSA® (gefitinib), EGFR tyrosine kinase inhibitors (TKIs), is indicated. |
The EGFR TK inhibitor, gefinitib (Iressa®, AstraZeneca), was initially approved in 2003 for all patients with advanced, chemotherapy-resistant NSCLC based on a finding that 10% of patients had a robust response to the drug. In approving gefitinib as part of a compassionate-use expanded access program, the FDA concluded that a 10% response rate was meaningful and demonstrated biologic activity of the drug. Subsequently, tissue samples from nine of the patients who responded to the drug were analyzed for clinical characteristics as well as somatic mutations in the TK domain of EGFR. The entire coding region of the gene was sequenced using PCR amplification of individual exons. Heterozygous mutations were observed in 8 of 9 patients all clustering within the tyrosine kinase domain of EGFR. No such mutations were found in NSCLC patients who were non-responders to gefitinib. The Iressa Pan-Asia study (IPASS) showed that gefitinib was associated with prolonged progression-free survival (PFS) in a patient population clinically enriched with never-smokers or light ex-smokers. In a subgroup analysis, PFS was longer in patients with EGFR exon 19 deletions and exon 21 (L858R) substitution mutations, suggesting that these mutations could be used as biomarkers to predict a response to gefitinib. Trials in patient populations enriched for these mutations confirmed this suggestion.

In 2005, a second EGFR TK inhibitor, erlotinib (Tarceva®, Astellas Pharma Inc.), was also initially approved for treatment of patients with chemotherapy-resistant metastatic NSCLC based on a study that demonstrated statistically significant improvements in median survival and progression free survival.

In the OPTIMAL trial, an open-label, multicenter, randomized trial in China, 165 NSCLC patients with EGFR mutations were randomized to receive erlotinib or carboplatin plus gemcitabine. The PFS in the erlotinib arm was 13.1 months compared to 4.6 months in the chemotherapy arm. This was followed by the European Tarceva versus Chemotherapy (EURTAC) trial, which used a laboratory developed test on tumor tissue combining Sanger sequencing and PCR assessment of gene deletions. The results of these and other trials prompted efforts to develop and validate a rapid multiplex EGFR mutation assay – the cobas® EGFR mutation test (Roche Molecular Systems) – as a companion diagnostic for erlotinib. A bridging study compared results from samples that had been prospectively tested using laboratory-developed tests in the EURTAC study with results from retrospective testing of those same samples using the cobas test kit confirmed the clinical utility of the cobas test and led to its subsequent approval as a companion diagnostic for first line treatment with erlotinib of patients with metastatic NSCLC tumors bearing EGFR exon 19 deletions or exon 21 (L858R) substitutions mutations.

On May 14, 2013, the U.S. Food and Drug Administration approved erlotinib for the first-line treatment of patients with metastatic non-small cell lung cancer (NSCLC) whose tumors have EGFR exon 19 deletions or exon 21 (L858R) substitution mutations. This indication for erlotinib was approved concurrently with the cobas EGFR Mutation Test, a companion diagnostic test for patient selection. The approval was based on clinically important improvements in progression-free survival (PFS) and objective response rate (ORR) and an acceptable toxicity profile demonstrated in a multicenter randomized trial of erlotinib versus standard chemotherapy.

A third EGFR TKI, Afatinib (Gilotrip®; Boehringer Ingelheim Pharmaceuticals, Inc.), was granted Fast-Track designation from the FDA in November 2007 and was allowed to proceed with an Expanded Access Program in June 2012 after two clinical studies supported its efficacy in NSCLC patients with EGFR mutations who had not received prior treatment with a EGFR TKI. In the primary study (LUX-Lung 3 study), EGFR mutation positive patients were randomized to receive afatinib or chemotherapy as first line treatment. Mutations were characterized by the TheraScreen 29 test kit, designed to detect 29 EGFR mutations. This trial demonstrated a statistically significant improvement in PFS in patients...
randomized to afatinib, with a median PFS of 11.1 months compared to 6.9 months in the chemotherapy arm. Overall survival was not changed, with a median survival of 28 months in each arm. Treatment effects varied by the underlying mutation, with greater effects observed for those with exon 19 deletions. Also observed were apparently harmful effects for patients with uncommon EGFR mutations.\textsuperscript{19}

July 2013, FDA approved afatinib as first-line therapy for patients with metastatic lung cancer and tumors that express EGFR mutations “as detected by an FDA-approved test.” At the same time, the FDA approved Qiagen’s therascreen EGRF test as a companion diagnostic for afatinib for treating NSCLC. In July 2015, FDA also approved Qiagen’s therascreen EGRF test as a companion diagnostic for gefitinib for the first line treatment of patients with metastatic NSCLC.

**Major lessons learned from the development process**

The approval of companion diagnostics for EGRF TKIs demonstrates that both retrospective and prospective data, combined with bridging studies, may be useful in the development process using the IND pathway. Pursuing qualification would potentially allow these diagnostic tests to be used with other similar drugs without further validation. However, while there are a number of newer, third generation EGRF inhibitors in development, these newer drugs target different mutations than those targeted by earlier generation EGFR TKIs.\textsuperscript{20} As a result, currently approved EGFR mutation tests such as cobas and therascreen will not be useful as companion diagnostics for these new drugs and additional tests will need to be validated.


