Biomarker Qualification: Evidentiary Framework

Guidance for Industry and FDA Staff

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 60 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit electronic comments to https://www.regulations.gov. Submit written comments to the Dockets Management Staff (HFA-305), Food and Drug Administration, 5630 Fishers Lane, Rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document, contact (CDER) Office of New Drugs at CDER-BiomarkerQualificationProgram@fda.hhs.gov , or (CBER) Office of Communication, Outreach and Development, 800-835-4709 or 240-402-8010.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

December 2018
Drug Development Tools

Biomarker Qualification: Evidentiary Framework

Guidance for Industry and FDA Staff

Additional copies are available from:
Office of Communications, Division of Drug Information
Center for Drug Evaluation and Research
Food and Drug Administration
10001 New Hampshire Ave., Hillandale Bldg., 4th Floor
Silver Spring, MD 20993-0002
Phone: 855-543-3784 or 301-796-3400; Fax: 301-431-6353
Email: druginfo@fda.hhs.gov
http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm
and/or
Office of Communication, Outreach, and Development
Center for Biologics Evaluation and Research
Food and Drug Administration
10903 New Hampshire Ave., Bldg. 71, Room 3128
Silver Spring, MD 20993-0002
Phone: 800-835-4709 or 240-402-8010; Email: ocod@fda.hhs.gov
https://www.fda.gov/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/default.htm

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

December 2018
Drug Development Tools

TABLE OF CONTENTS

Biomarker Qualification: Evidentiary Framework Guidance for Industry and FDA Staff[1]

This draft guidance, when finalized, will represent the current thinking of the Food and Drug Administration (FDA or Agency) on this topic. It does not establish any rights for any person and is not binding on FDA or the public. You can use an alternative approach if it satisfies the requirements of the applicable statutes and regulations. To discuss an alternative approach, contact the FDA staff responsible for this guidance as listed on the title page.

I. INTRODUCTION

This guidance for biomarker[2] development stakeholders and Food and Drug Administration (FDA) staff provides recommendations on general considerations to address when developing a biomarker for qualification under the 21st Century Cures Act (Cures Act) that added new section 507of the Federal Food, Drug, and Cosmetic Act (FD&C Act), Qualification of Drug Development Tools.[3] This guidance discusses the evidentiary framework that should be used to support *biomarker qualification*, as that term is now used in section 507 of the FD&C Act, and it was informed by public workshops that predated the Cures Act.

The evidentiary framework described in this guidance identifies the recommended components of a biomarker development program, including determining the type and level of evidence sufficient to support qualification, and addresses how these components interrelate to inform the evidentiary framework. This evidentiary framework is broadly applicable to all biomarker qualification submissions, regardless of the type of biomarker or context of use (COU). Qualification of a biomarker is a determination that within the stated COU, the biomarker can be relied on to have a specific interpretation and application in drug development and regulatory review.[4] Thus, a qualified biomarker can be used across multiple drug[5] development programs under the COU for which it was qualified. Requests for qualification of a biomarker should address the evidentiary framework discussed in this document.

[1] This guidance has been prepared by the Office of New Drugs and the Office of Translational Sciences in the Center for Drug Evaluation and Research (CDER) at the Food and Drug Administration in cooperation with the Center for Biologics Evaluation and Research (CBER).

[2] Throughout this guidance, the term *biomarker* is intended to include both single entity and composite biomarkers (biomarkers consisting of several individual biomarkers whose measurements are combined in a defined algorithm to reach a single interpretive output). References in this guidance to the use of a biomarker in drug development imply making a decision in drug development based upon the measurement of the biomarker.

[3] Section 507 of the FD&C Act (21 U.S.C. 357) was added by section 3011(a) of the Cures Act (Public Law 114-255).

[4] FD&C Act section 507(e)(7).

[5] For the purposes of this guidance, all references to drugs or drug products include both human drugs and biological drug products regulated by CDER and CBER, unless otherwise specified.

Many principles discussed in this guidance could also be appropriate when considering the evidence scientifically sufficient to support the use of a biomarker in an individual drug development program (e.g., investigational new drug application, new drug application, or biologics license application submissions). The specifications for medical devices and the processes and evidence that support obtaining marketing authorization for medical devices, including the qualification of a biomarker for use in the investigation of a medical device or use with a medical device, are outside the scope of this document.

This guidance was informed by several public workshops[6] that discussed the science to support biomarker qualification; these workshops convened before the enactment of the Cures Act. Development of this guidance was also greatly facilitated by the efforts from the biomarker development community—including FDA, National Institutes of Health (NIH), industry, academia, patient groups, and the nonprofit sector—that developed an October 2016 white paper describing a Framework for Defining Evidentiary Criteria for Biomarker Qualification.[7] In addition to considering public comments received regarding this guidance, FDA anticipates that the Agency will incorporate additional information required under the Cures Act and discussed in the reauthorized Prescription Drug User Fee Act (PDUFA VI) goals letter (PDUFA VI goals letter)[8] in a subsequent revised draft version of this guidance. Ultimately, FDA anticipates that a future revised draft guidance on this topic will meet the statutory requirement for draft guidance on a "conceptual framework describing appropriate standards and scientific approaches to support the development of biomarkers" described in section 3011(b)(1)(A) of the Cures Act and meet the commitment in section (1)(J)(6)(d) of the PDUFA VI goals letter related to publishing a draft guidance on "general evidentiary standards for biomarker qualification." As part of FDA's efforts to delineate the conceptual framework to support biomarker qualification and the general evidentiary standards for biomarker qualification, FDA also anticipates that subsequent guidance on biomarker qualification will address specific aspects of evidentiary considerations (e.g., statistical, analytical) in greater detail.

In general, FDA's guidance documents do not establish legally enforceable responsibilities. Instead, guidances describe the Agency's current thinking on a topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word *should* in Agency guidances means that something is suggested or recommended, but not required.

---

[6] Workshops convened to discuss the science to support biomarker qualification included: Institute of Medicine Workshop on Biomarker Qualification (2009), FDA co-sponsored Biomarkers Workshop with Howard Hughes Medical Institute (2013), FDA co-sponsored Brookings meeting on Advancing the Use of Biomarkers and Pharmacogenomics (2014), FDA co-sponsored workshop with M-CERSI and the Critical Path Institute on Evidentiary Considerations for Integration of Biomarkers in Drug Development (2015), NIH-FDA Workshop on Biomarker Glossary of Terms (2015), the National Biomarker Development Alliance's Workshop on Collaboratively Building a Foundation for FDA Biomarker Qualification (2015), and Foundation for the NIH-FDA Workshop on Developing an Evidentiary Criteria Framework for Safety Biomarkers Qualification (2016).

[7] Biomarkers Consortium Evidentiary Standards Writing Group: Framework for Defining Evidentiary Criteria for Biomarker Qualification. Final version 10/20/2016. Available at: https://fnih.org/sites/default/files/final/pdf/Evidentiary%20Criteria%20Framework%20Final%20Version%20Oct%2020%202016.pdf.

[8] The PDUFA VI goals letter is available at: https://www.fda.gov/downloads/ForIndustry/UserFees/PrescriptionDrugUserFee/UCM511438.pdf.

II. BACKGROUND

Historically, biomarkers gained acceptance for use in drug development after evidence from scientific and medical communities accumulated over time, leading to the recognition of the role and value of the biomarker in decision-making. This evidence was considered as part of drug-specific development efforts, and there was no formal regulatory process to assess the broader utility of the biomarker independent from its use in a specific drug program. Even after the Center for Drug Evaluation and Research established the legacy (pre-Cures Act) Biomarker Qualification Program in 2007, progress in biomarker development has been hampered by the lack of a clear, predictable, and specific regulatory framework for the type and level of evidence sufficient to support regulatory decision-making using biomarkers. This guidance is an additional step towards informing future guidances that will specifically address this need, the Cures Act requirements, and PDUFA commitments. Throughout this guidance, FDA uses certain terms that appear in the FDA-NIH Biomarker Working Group, BEST (Biomarkers, EndpointS, and other Tools) Resource. The BEST Resource includes a taxonomy of terms that can be accessed online and on which FDA welcomes comment.[9]

III. EVIDENTIARY FRAMEWORK

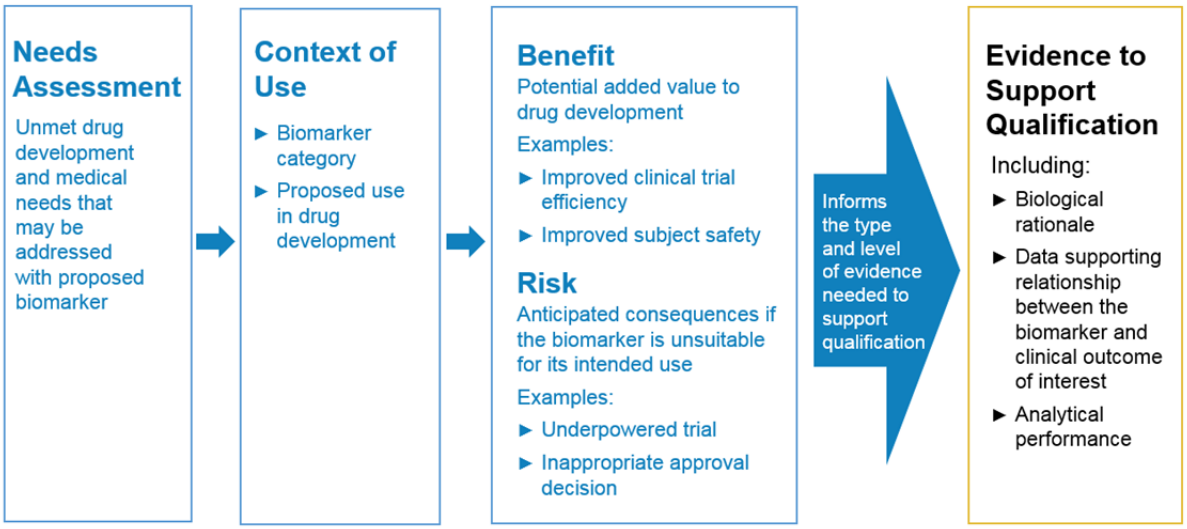
For a biomarker development effort to be successful, the biomarker should be clearly identified and characterized, including its source material or matrix and its method of measurement. The biomarker should be clearly identified based on the specific analyte (e.g., fibrinogen), anatomic feature (e.g., joint angle), or physiological characteristic (e.g., blood pressure) being measured. For composite biomarkers, it is important to list the individual biomarker components and how these components are interrelated (e.g., a description of an algorithm or scoring system). If individual components have differential weighting, the description should include the biologic rationale to support this decision. Because biomarkers are measured entities, it is important to describe the biomarker source or material for measurement, which determines the biomarker type (e.g., molecular, histologic, radiographic, physiologic characteristic). For example, a molecular biomarker obtained from a biofluid should state the sample matrix (e.g., plasma, urine), and a radiographic biomarker should include the organ or tissue imaged (e.g., kidney). For radiographic biomarkers, it may be appropriate to include the assay/imaging modality or method for interpretation (e.g., dual-energy x-ray absorptiometry, T4/T1 ratio by acceleromyography).

The evidentiary framework that should be considered when determining the type and level of evidence sufficient to support qualification of a biomarker consists of several components. The framework includes: (1) describing the drug development need, (2) defining the COU, (3) considering potential benefits if the biomarker is qualified for use, and (4) considering potential

[9] The FDA-NIH Biomarker Working Group BEST Resource is available at https://www.ncbi.nlm.nih.gov/books/NBK326791/. The BEST Resource contains a glossary intended to harmonize terms used in translational science and medical product development, with a focus on terms related to study endpoints and biomarkers. The glossary will be periodically updated with additional terms and clarifying information (last accessed March 1, 2018).

risks associated with the proposed use of the biomarker in a drug development program (see Figure 1).

Figure 1: Evidentiary Framework



A. Needs Assessment

The needs assessment describes why a biomarker is needed for drug development, including how its use might promote drug development in areas where there is an unmet medical need. The needs assessment should describe the current drug development landscape, such as the use and limitations of available biomarkers or other drug development tools, and the added value the novel biomarker could provide to the current drug development process. The needs assessment should also consider the degree to which there is an unmet medical need in the relevant condition or conditions (e.g., a greater unmet need if there is a serious condition with no or limited treatment) that can be more efficiently or effectively addressed through use of the proposed biomarker in drug development. The needs assessment can include factors that FDA may determine to be helpful for informing FDA prioritization of the review of full qualification packages, including, as applicable, the severity, rarity, or prevalence of the disease or condition targeted by the biomarker; the availability or lack of alternative treatments for such disease or condition; and the identification (by FDA or by biomedical research consortia and other expert stakeholders) of a biomarker and its proposed COU as a public health priority.[10]

B. Context of Use

According to section 507(e)(4) of the FD&C Act, "the term 'context of use' means, with respect to a drug development tool, the circumstances under which the drug development tool is to be used in drug development and regulatory review." The COU is a concise description of a biomarker's specified use in drug development. The COU includes two components: (1) the

[10] FD&C Act section 507(a)(2)(C).

biomarker category and (2) the biomarker's proposed use in drug development. Each biomarker qualification effort should identify a single COU.

Biomarkers can be disease-related or treatment-related and should be classified by the BEST biomarker category, selected from the following (see BEST Resource for discussion of each category of biomarker[11]):

- diagnostic biomarker
- monitoring biomarker
- pharmacodynamic/response biomarker (e.g., clinical trial endpoints, including surrogate endpoints)
- predictive biomarker
- prognostic biomarker
- safety biomarker
- susceptibility/risk biomarker

The *proposed use in drug development* should include, as appropriate:

- Purpose of use in drug development (e.g., a prognostic biomarker *to support enrichment of Alzheimer's Disease clinical study/trial populations*, a safety biomarker *to evaluate drug-induced liver injury*)

- Proposed stage of drug development (e.g., phase 1 clinical trials, nonclinical safety studies)

- Clinical trial population or model system (e.g., healthy adult subjects, patients with COPD, rats, cultured mouse fibroblasts)

- Therapeutic mechanism of action (MOA) for which the biomarker is intended to have value, provided that the MOA is relevant to the biomarker's biology and intended utility (e.g., both the MOA and the biomarker are within the same biologic pathway or process)

Accumulating the data to support a biomarker for qualification can take considerable time and resources. Often, requestors do not have adequate data and/or information to support their proposed COU. One approach is to initially qualify a biomarker for a COU that is limited in scope to facilitate the integration of the biomarker in drug development, which could result in the accumulation of additional evidence that can help qualify the biomarker for a COU with a more expanded scope in the future. For example, a biomarker could be qualified first as a pharmacodynamic biomarker for use in dose selection. After additional information is accumulated, the same biomarker could ultimately be qualified as a pharmacodynamic biomarker for use as a clinical trial endpoint; if the biomarker is considered to be reasonably likely to

[11] Definition is from the BEST Glossary available at: https://www.ncbi.nlm.nih.gov/books/NBK326791/ (last accessed March 1, 2018).

predict clinical benefit or has been shown to predict clinical benefit, it could be used as a surrogate endpoint to support accelerated[12] or traditional drug approval, respectively.

C. Assessment of Benefits and Risks

Biomarker developers are expected to provide a clear and objective description of the anticipated benefits and risks of the biomarker for the proposed COU, as well as any potential risk mitigation strategies.[13] The overall balance of benefits, risks, and risk mitigation efforts are critical for determining the strength of evidence sufficient to support qualification.

The potential benefits of a biomarker for use in drug development depend on the biomarker's proposed COU and the needs assessment. Biomarker use could benefit individual patients participating in clinical trials (e.g., earlier identification of toxicity with a safety biomarker) or general drug development and regulatory decision-making (e.g., a prognostic or predictive biomarker used to enrich a patient population could reduce the sample size needed to achieve statistical significance).

The potential risks of qualifying a biomarker should address the consequences of incorrect decision-making or harm to patients if the correlation between the biomarker and the outcome of interest does not indicate what it is intended to indicate. Requestors should consider factors that might mitigate harm if the biomarker does not perform as expected. The potential risk is closely linked to biomarker category and the proposed COU. For example, if a safety biomarker fails to accurately predict early toxicity, clinical trial participants might be placed at risk for serious adverse drug reactions. Alternatively, the same safety biomarker might jeopardize the successful development of a promising new drug and prevent significant societal benefits if it erroneously identifies a risk where none exists. These risks could be mitigated, in part, by using the proposed biomarker with existing safety monitoring measures, rather than as a stand-alone assessment for the toxicity of interest. In another example, a prognostic biomarker intended for clinical trial enrichment might fail to identify patients with more rapid disease progression. In this case, mitigation strategies could include incorporating an interim analysis for sample size re-estimation.

The following questions should be used to characterize the potential benefits and risks of a biomarker for a specific COU:

1. Does the biomarker have the potential to add value to drug development?

2. What other tools are available for the biomarker's proposed use and what added value might the biomarker provide?

[12] To obtain accelerated approval for a drug, sponsors must meet the statutory criteria in section 506(c) of the FD&C Act (21 U.S.C. 356(c)). Also see 21 CFR part 314, subpart H and part 601, subpart E.

[13] The terms *benefit*, *risk*, and *risk mitigation* that are used in the context of biomarker qualification have specific meanings that are relevant to biomarker development and evaluation, and these meanings are separate and distinct from how these terms are used in the context of evaluating the safety and effectiveness of medical products.

3. What are the anticipated consequences if the biomarker is unsuitable for its proposed use?

4. What factors or other tools can mitigate the potential risks of relying on the biomarker for its proposed use if the biomarker does not perform as expected?

D. Determining Evidence That Is Scientifically Sufficient To Support COU

The evidence sufficient to qualify a biomarker depends on its COU and the potential benefits and risks associated with its use. The benefits and risks associated with a biomarker's COU drives expectations for the reliability of the biomarker to predict the outcome of interest. If the potential benefits far outweigh the potential risks and/or there are acceptable risk mitigation approaches, there could be increased tolerance for uncertainty. In such a case, the strength of evidence expected to support qualification could be lower. If the potential benefits minimally outweigh the risks of relying on the biomarker, the strength of evidence expected to support qualification should be higher.

Ultimately, whether there is sufficient evidence to support qualification of a biomarker for use in drug development depends on the selection of the appropriate biomarker for the proposed COU, the quality of the biomarker measurement, and the correlation of the biomarker with the outcome of interest. Evidence to support qualification consists of data to support clinical validation and analytical validation.

Clinical validation establishes that a biomarker's relationship with the outcome of interest is acceptable for the proposed COU. The requestor should describe what is known about the biomarker's role in the causal or outcome pathway of interest, as well as describe knowledge gaps about the pathophysiology and molecular underpinnings of the disease. Describing the biomarker's position in the disease pathway, if applicable, helps to support the biological plausibility of the biomarker's role in the proposed COU. The requestor should provide data supporting the relationship between the biomarker and a clinical outcome that reflects how an individual feels, functions, or survives. This relationship should be supported by statistical analyses (see section V.) and should come from multiple independent data sources. Together this information can establish the clinical validity of a biomarker for a specified COU.

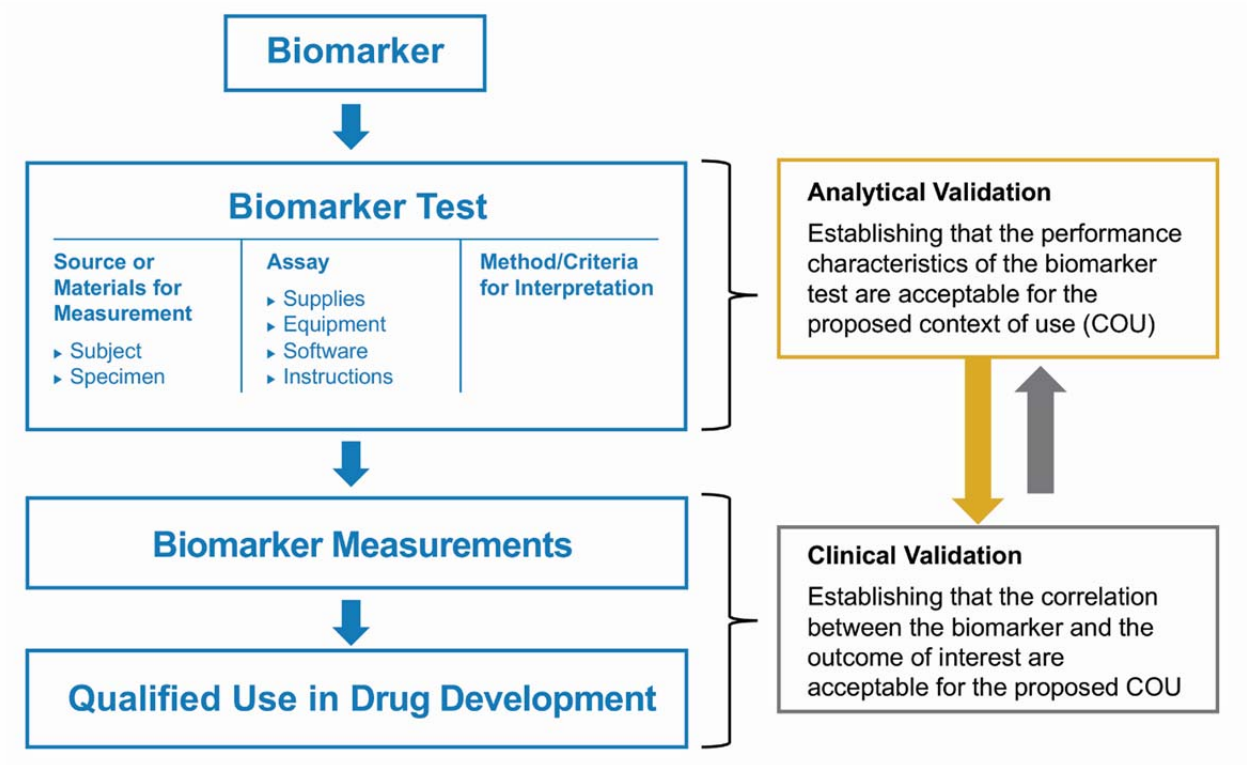
Biomarkers considered for qualification are conceptually independent of the specific method of measurement; however, a biomarker cannot be qualified without a reliable method of measurement.[14] Relevant performance characteristics of the biomarker tests used to support qualification should be assessed through analytical validation studies to ensure that biomarker data for qualification is obtained using acceptable measurement methods to support the proposed COU and that biomarker tests used in drug development for the COU (if different from the tests

[14] Qualification of a biomarker does not connote approval or clearance of a diagnostic device or of a companion or complementary diagnostic device for use in clinical practice, and it also does not qualify the biomarker for use in clinical practice. The approval/clearance of a biomarker test by the Center for Devices and Radiological Health or by CBER also does not indicate qualification of the biomarker for use in drug development.

used to qualify the biomarker) perform as well as the tests used for biomarker qualification. Analytical considerations are discussed further in section IV. below.

Clinical validation and analytical validation are distinct processes; however, the two processes are iterative and dependent on one another. A reliable test should be used to measure the biomarker before the biomarker measurement cutoffs[15] can be established, and the cutoffs should be defined before the biomarker test can be analytically validated. Through this iterative process, experience with the biomarker and the biomarker test could lead to improvements in the technical performance of the test and the understanding of the biomarker's biological and clinical significance. It is important to have a high level of confidence in the biomarker test's analytical performance when confirming the relationship between a biomarker and clinical outcome of interest, and generally, biomarker qualification studies intended to confirm this relationship should be conducted using a validated test (see Figure 2).

Figure 2: **Biomarker Validation Approach**



The rigor of the analytical and clinical validation performed for biomarker qualification should support the utility of the proposed COU. A listing of qualified biomarkers with FDA reviews

[15] Cutoff is the value at or above which a biomarker test result is determined to be positive (or in a specific category) and the value below which the result is determined to be negative (or in a different category).

describing the evidence leading to their qualification can be found on the FDA Biomarker Qualification Program website.[16]

IV. ANALYTICAL CONSIDERATIONS

Because drug development decisions will be made based upon qualified biomarkers, any biomarker test used to measure the biomarker should be robust, sensitive, and specific enough to support the decisions defined by the COU.

Analytical validation for the purpose of biomarker qualification includes establishing that the analytical performance characteristics of a biomarker test, such as the accuracy and reproducibility, are acceptable for the proposed COU in drug development. This is validation of the test's technical performance, but is not validation of the biomarker's usefulness. The biomarker test and associated performance characteristics will vary depending on the biomarker type (molecular, histologic, radiographic, and physiologic characteristic). A biomarker test is an assessment system comprising three essential components: (1) source or materials for measurement, (2) an assay for obtaining the measurement, and (3) method and/or criteria for interpreting those measurements. All relevant components of the biomarker test should be assessed in the analytical validation studies and determined to be acceptable (see Figure 2).

Analytical validation of a biomarker test should consider the acceptability of the source or materials from which the biomarker is measured. For a molecular or histologic biomarker, for example, the source includes not only the sample, but also the sample collection, storage, and processing conditions. For a radiographic or physiologic biomarker, the source of measurement could include factors such as the patient preparation and positioning. Sample collection, preparation, and storage protocols (as applicable for the biomarker type) should be established and assessed in the analytical validation studies to determine acceptability.

A reliable biomarker test is also contingent on all components of the biomarker assay, such as supplies, equipment, software, and instructions. User instructions/protocols should be established and followed during validation testing to ensure acceptability. Additional details such as reagent versions, lot numbers, and software version should be noted to help identify modifications to the test that could alter performance.

Biomarker measurements are expressed in many ways (e.g., the concentration of molecular species in body fluids, cells, or tissues; the presence or extent of features in images obtained from microscopy or radiology; the magnitude of in vivo physiological signals). Some of these measurements are produced directly from a biomarker test, and others are determined by an interpretation of biomarker test results. Examples of these interpretations include radiographic image analysis and the combination of individual biomarker measurements in a defined algorithm to determine a composite biomarker score. The measurement interpretation, as with the other components of the biomarker test, can introduce error into the biomarker measurement;

[16] Information on qualified biomarkers is available at https://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/BiomarkerQualificationProgram/ucm535383.htm.

therefore, reliable and acceptable interpretation should be established in the analytical validation studies.

Acceptance criteria for analytical performance characteristics for a biomarker test are set according to the overarching specifications for the biomarker to support the proposed COU and according to the risks associated with limitations in the analytical performance of the test. Inadequate biomarker test performance could lead to incorrect interpretation of a biomarker's significance, thus undermining the clinical validation of the biomarker. Bias and dispersion in the biomarker test lead to uncertainty when interpreting biomarker test results and affect the value of the biomarker as a drug development tool. Requestors should consider not only the proposed COU and potential risks and benefits of the proposed biomarker, but also the following factors when specifying performance characteristic acceptance criteria for candidate tests:

- The performance characteristics of existing measurement methods
- The biological variability of the biomarker in the populations of interest, if known
- The minimum magnitude of biomarker change expected to affect decisions for the proposed COU (i.e., cutoff for separating populations or determining change from baseline)

Considerations for assessing the performance characteristics of biomarker tests for specific types of biomarkers are beyond the scope of this guidance. The FDA guidance for industry *Considerations for Use of Histopathology and Its Associated Methodologies to Support Biomarker Qualification*[17] provides general considerations regarding performance characteristics for histologic biomarker methodologies. The analytical validation studies and performance characteristics vary greatly according to the technology of the biomarker test. Many well-accepted protocols are published for examination of analytical performance characteristics for specific biomarker test methodologies. Such protocols can be selected and adapted for use in accordance with a risk-based assessment of the evidentiary stringency determined by the proposed COU.

V. STATISTICAL CONSIDERATIONS

The goal of statistical analyses in biomarker qualification is to evaluate the degree and certainty of association between a biomarker and an outcome of interest. Consideration should be given to the design and conduct of studies contributing data to support biomarker qualification, as well as the statistical analyses conducted. This section describes the potential sources of data, as well as study design and statistical considerations when assessing the association between a proposed biomarker and an outcome of interest for the purposes of biomarker qualification.

Data used to establish the relationship between a biomarker and an outcome of interest, to support biomarker qualification, can come from a variety of sources including the following:

[17] We update guidances periodically. To make sure you have the most recent version of a guidance, check the FDA guidance web page at https://www.fda.gov/RegulatoryInformation/Guidances/default.htm.

- Randomized controlled trial
- Single-arm/historical control trial
- Cohort study
- Case-control study (including nested)
- Cross-sectional study
- Case series or case reports
- Registry information
- Meta-analysis

The strongest level of evidence to support the association of a biomarker with an outcome of interest comes from prospective studies that are specifically designed and powered to assess the association. In many settings, however, data from studies conducted for other purposes are used to support biomarker qualification. Ultimately, the COU, with its associated potential benefits and risks, determines what types of data may be acceptable to support qualification; clinical trial data is not critical for all COUs. Regardless of the data sources proposed to support the biomarker's COU, biomarker developers should consider the potential methodological limitations that could lead to overestimation of any actual associations, including lack of proper control for bias, confounding, and multiplicity, and address these limitations in their analysis plan. Verification of the results with an independent data source increases the credibility of the results.

Although the recommendations provided in the ICH guidance for industry *E9 Statistical Principles for Clinical Trials*[18] are primarily intended for late-stage interventional clinical trials, many of the principles described in ICH E9 are also relevant when considering the data intended to support biomarker qualification. Specifically, the principles are as follows:

- To the extent possible, the sample size should be sufficient to ensure adequate power to assess a clinically relevant association between the biomarker and the outcome of interest with reasonable dispersion. Sample-size considerations could be based on a single study or multiple studies considered in aggregate, and it is recognized that flexibility might be needed in certain clinical contexts (e.g., rare diseases).

- The analysis plan should control for multiplicity and consider the potential for false positive results. Multiplicity commonly occurs when analyzing multiple-candidate biomarkers and could lead to overestimation of biomarker associations with the clinical outcome of interest.

- The strength of the relationship between the biomarker and the outcome of interest should be quantified appropriately. Over-reliance on p-values should be avoided.

[18] Available on the FDA Drugs guidance web page at https://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm.

- Potential sources of bias should be identified and strategies to minimize bias should be described. For example, when possible, the biomarker analysis plan should be developed before unblinding of the data and access to subjects' biomarker status for purposes of the analysis. In some situations, clinical outcome data might have already been unblinded and analyzed, but the initial analyses did not include the biomarker data (i.e., if samples were collected for later use) or the analyses recommended to support qualification were not performed. Although such data could be used to support qualification, the analyses intended to support biomarker qualification should be specified in an analysis plan with a prospective-retrospective design before analyzing the data.

- Consideration should be given to sample and data collection methods, including strategies to minimize and account for the effect of missing data, and these methods should be included in the analysis plan. When collection of biomarker data is only from a subset of clinical sites, groups, or treatments, this non-randomized sampling (convenience sampling) might be statistically problematic if the subset is somehow partial to the outcomes being studied, yielding biased estimates with unknown characterization of the bias.

- Innovative statistical approaches such as adaptive designs and Bayesian designs, including prior information and hierarchical models, can be considered for qualification of biomarkers.

Data supporting biomarker qualification are often based in part on the published literature and, in some situations, could be exclusively based on the published literature. It is critical for the biomarker developer to identify the limitations and gaps in these data and address how they affect the interpretability of the results. In addition, when using published literature, the criteria for study inclusion should be specified a priori in a systematic study protocol of the published literature, to avoid publication or selection bias.

When assessing whether the association between a biomarker and an outcome of interest is acceptable for the proposed COU, a key consideration is how to define the outcome of interest. In some settings, there might not be a current standard outcome, or a standard outcome with known limitations is used for comparative purposes. For example, changes in serum creatinine are widely used in biomarker development as the current standard for predicting drug-induced kidney injury. However, changes in serum creatinine levels are neither highly sensitive nor highly specific for drug-induced kidney injury. In a setting in which the current standard outcome has significant limitations or a current standard outcome does not exist, it is important to consider the totality of all available data that may provide sufficient support to establish that the biomarker can be acceptably relied upon for the proposed COU. Each biomarker qualification submission has unique challenges that call for careful clinical and statistical considerations that may lead to distinct solutions.

There are multiple statistical approaches to assessing the association between a biomarker and clinical outcome measures. For binary outcome measures, such as the presence or absence of disease, results can be evaluated using clinical sensitivity and specificity, and positive and negative predictive values, or by evaluating receiver operating characteristic curves. For

continuous outcome measures, such as disease progression, results can be evaluated using regression models. When appropriate, adjusted or composite biomarkers can be considered with adequate justification, including biomarkers derived from composite measurements, covariate-adjusted measurements, change from baseline measurements, and repeated measurements.

When continuous data will be dichotomized, the relationship between the clinical outcome and the biomarker could be initially established quantitatively. Expressing biomarker measures quantitatively increases the statistical power compared to dichotomization when establishing such a relationship. Once a relationship between a biomarker and an outcome of interest has been established, several cutoffs on a continuous biomarker can be considered. The most appropriate cutoffs can then be selected by comparing the clinical outcomes of *at risk* subjects with each different biomarker cutoff. The choice of a cutoff can also be informed by the benefit-risk tradeoff of the decisions made based on the biomarker and the proposed COU (e.g., selecting a cutoff that gives more weight to clinical sensitivity versus a cutoff that gives more weight to clinical specificity). In some instances, selecting a specific cutoff might not be appropriate, and describing a spectrum of threshold values for the biomarker could be more informative. For example, in the case of an enrichment biomarker, submissions might describe a spectrum of cutoffs in a model representing the potential increase in power to be gained from enrichment, which should be considered against potential enrollment challenges resulting from a narrowed patient population.

There are no set quantitative criteria for determining whether the relationship between the biomarker and the clinical outcome is sufficiently strong to support biomarker qualification. Criteria based on parameters used to quantify the relationship, such as the threshold values for sensitivity and specificity, and the presence of a gradient (e.g., clinical performance change as function of biomarker quantity) can provide confidence that a finding is likely to be relevant, reliable, and statistically robust. Additional considerations that support the biomarker's association with the clinical outcome should also be assessed, such as whether there is a strong biological rationale supporting the role of the biomarker in the proposed COU and whether the findings are supported by more than one investigation or analysis set or there are multiple lines of evidence (e.g., experimental models and human studies). Together, the strength of the data supporting the association and additional considerations should be evaluated to determine whether the evidence supporting the relationship between the biomarker and the clinical outcome is adequate to support biomarker qualification. This determination will be dependent on the evidentiary framework assessment for each individual submission described in section III.