Perspective

For reprint orders, please contact: reprints@future-science.com

Bioanalysis

Adaptation of commercial biomarker kits and proposal for 'drug development kits' to support bioanalysis: call for action

Rafiqul Islam*,[1], Sumit Kar[1], Clarinda Islam[2] & Raymond Farmen[1]

[1]Celerion, Inc., Global Bioanalytical Services, 621 Rose Street, Lincoln, NE 68502 USA

[2]Somru BioScience, Inc., Innovation Way, BioCommons Research Park, Charlottetown Prince Edward Island, C1E 0B7 Canada

*Author for correspondence: Tel.: +1 402 437 4704; Fax: +1 402 939 0428; Rafiqul.Islam@celerion.com

There has been an increased use of commercial kits for biomarker measurement, commensurate with the increased demand for biomarkers in drug development. However, in most cases these kits do not meet the quality attributes for use in regulated environment. The process for adaptation of these kits can be frustrating, time consuming and resource intensive. In addition, a lack of harmonized guidance for the validation of biomarker poses a significant challenge in the adaptation of kits in a regulated environment. The purpose of this perspective is to propose a tiered approach to commercial drug development kits with clearly defined quality attributes and to demonstrate how these kits can be adapted to perform analytical validation in a regulated environment.

First draft submitted: 23 November 2017; Accepted for publication: 19 February 2017; Published online: 25 June 2018

Keywords: biomarkers • drug development kit • immunoassay

Use of commercial kits is ubiquitous

Over the last few decades the use of biomarkers in drug development has increased significantly. Biomarkers, when used under the right context, can expedite informed decision making early in the drug development process and improve the success of late-stage pivotal clinical trials [1]. The most commonly used biomarkers in drug development are pharmacodynamic, prognostic and predictive. Pharmacodynamic biomarkers can help understand the drug's ability to engage the target and its ability to modulate relevant biological function. The prognostic biomarkers can help stratify disease progression regardless of therapy. The predictive biomarker provides information about the effect of a therapeutic intervention. Regardless of the type of biomarker, from an analytical perspective, the methods need to be validated to meet the context of use.

Another driver for increased use of biomarkers in drug development is the growth in biomarker-guided adaptive trial design. Biomarkers can play a significant role in adaptive trial design by identifying the patients most likely or unlikely to benefit from the new drug. The accelerating pace of biomarker utility in clinical trials for regulatory decision making is further evidenced by the increased number and proportion of clinical trials using biomarkers [2].

The emphasis from regulatory agencies for biomarker data has resulted in an increased use of commercially available biomarker kits. These kits can potentially be more cost effective and efficient than developing an assay. However, depending on the type of kit and its intended purpose, the ease of validating a biomarker kit to drug development standards can vary.

Types of commercial kits

Commercially available biomarker kits vary in their level of performance characteristics. They range from research use grade to US Food and Drug Administration (FDA)-approved clinical grade diagnostics. While an extensive review of each type of kit is outside the scope of this paper, a brief description of each type and its relevant characteristics will serve as a useful reference.

Research use only (RUO) kits have the most variability. Its performance characteristics and characterization are vendor determined. There are no guidelines for the development or manufacturing of RUO kits. Some major

newlands press

concerns for the use of RUO kits are lot to lot variability, reagent stability, the lack of appropriate quality controls (QCs) and the calibrator used may not be traceable to a reference standard. In many cases, the user will need to optimize the assay for the intended use, including adding calibrator points, matrix QCs, reagent substitution and curve fitting. Additionally, the assay will need to have its sensitivity and stability determined. Many times a kit will have a defined LOD versus an LOQ which is a desired parameter. However, they serve as very useful tools in the drug development process, as in many cases they are the first to market for novel biomarkers and the assays can be more adaptable and optimized as a drug development tool [3].

In vitro diagnostic (IVD) kits come in several classes (Class I, II or III). As these kits are considered medical devices, they are classified according to the level of complexity and risk. All must undergo FDA review and approval through premarketing approval and the 510(k) process. The quality requirements for these kits are specified in 21 Code of Federal Regulations (CFR) 522. While the development and assay performance are clearly characterized, these kits may not have the required range or sensitivity required. Many of these kits are 'closed' systems, where substituting reagents or modifying assay parameters are not possible. As outlined in the 2013 Draft FDA guidelines, these kits must be adapted to be shown suitable for drug development studies [4]. They often consist of one- or two-point calibration curves and must be validated with at least six calibration points. The number of QCs provided in IVD kits, in most cases, are not sufficient and additional QCs need to be added. Also, the acceptance criteria based on concentration range need to be substituted with percent bias from nominal concentration.

Clinical Laboratory Improvement Amendments (CLIA) tests are designed specifically for diagnostic purposes and follow quality standards as defined by the Clinical and Laboratory Standards Institute (CLSI). These assays fall into three different categories based on the complexity of their execution: waved, moderate and high. These kits may not have sufficient assay performance characteristics beyond their approved use.

Conformité Européenne (CE)-marked kits are for products marketed in the European Economic Area. While the CE mark implies conformity to safety and manufacturing requirements established by the EU, the process is mostly self directed. There are some categories of devices that will require a third-party review. While the CE mark denotes a certain standard for manufacturing, it does not guarantee that the assay is suitable for the measurement of biomarkers. As with the other types of kits, they may need to be optimized for its context of use.

Challenges in the adaptation of biomarker kits

Commercial kits are appealing to bioanalytical scientists due to their immediate availability and perceived quality. Using commercially available kits have a potential for time and cost savings by eliminating the need for developing reagents and assays 'from scratch'. However, due to the high variability in quality, it is difficult to assess the assay performance before purchase. As a result, in most cases the desired assay quality is not achieved in a cost-effective manner. In order to have an assay appropriate for the context of use, assay optimization, and in some cases assay reconstruction, are required.

Most commercial kits are developed for diagnostic purposes or drug discovery purposes, as opposed to drug development. There is an unmet need for 'well-defined' commercial kits that can be utilized in a drug development setting. While the quality of the FDA-approved IVD kits are manufactured using stringent quality systems, as illustrated by the case study below, they fail to meet drug development objectives. All commercial kits, excluding IVD kits, are marked as RUO. As previously mentioned, RUO kits do not have any harmonized quality standards and their performance characteristics vary significantly.

For example, in a recent paper evaluating commercial biomarker kits for a cardiac biomarker (human tyrosine-protein phosphatase nonreceptor type substrate 1 α [SIRPA]), the authors presented data that indicate the variability in the quality and utility of these kits [5]. Two ELISA kits for the quantitation of SIRPA were evaluated by running samples using both kits. To verify the identity of the protein being detected, the calibrators included in the kits were evaluated by western blot and MS, using a recombinant human SIRPA as a control. The results of the western blot indicated that neither kit calibrator reacted with an anti-SIRPA antibody. The results of the MS analysis showed no SIRPA being detected in either kit calibrator. While the kits generated results that corresponded with the expected values for the samples, further reagent characterization showed that the kit was not detecting the intended target. This demonstrates that the lack of analytical validation and insufficient reagent characterization are common issues with biomarker kits.

Multiplexed commercial kits that evaluate many related analytes in one assay are also becoming more common. A detailed discussion of the challenges and solutions in using multiplexed biomarker kits for drug development has been highlighted previously and is outside the scope of this discussion [6,7]. However, a recent study by Fischer

et al. highlights the significant challenges of using multiplex kits for drug development [7]. Only 18 out of 47 assays in multiplexed kits met the set bioanalytical acceptance criteria and assay sensitivity significantly diminished when using study matrix. This is another example of poor assay performance characteristics found in biomarker kits.

Some vendors have attempted to stratify the quality of their kits by using their own trademark names such as the V-Plex line of kits from Meso Scale Discovery, pharmacokinetic (PK) or anti-drug antibody (ADA) kits from Somru BioScience and Quantikine kits from R&D systems. However, the definitions of these kits and their contents are ambiguous and inconsistent from vendor to vendor.

With an increased interest in the use of biomarkers in drug development, not surprisingly, the regulatory scrutiny on the analytical validity of the biomarkers has increased. This is evidenced by the most recent FDA bioanalytical guidance document (2013) which requires stringent analytical validation requirements for biomarkers used for regulatory decision making [4,8]. Faced with an increased demand for biomarker assays and the high regulatory scrutiny, bioanalytical scientists are scrambling to develop and validate assays in quality manner, in most cases using commercial kits.

In order to overcome these challenges, we propose two solutions that will greatly alleviate this situation. First, we propose the term of drug development kits (DDK) with tiered and defined quality attributes. Second, we provide a framework to organize different regulatory and industry recommendations for the analytical validation of biomarkers to streamline the adaptation of commercial kits.

A context of use of framework for commercial kit adaptation

Detailed guidelines for the adaptation of commercial kits for biomarker bioanalytical assays have been published several times over the last decade [3,4,8–11]. For example, Khan *et al.* provides a detailed workflow of creating a biomarker work plan (BWP) to determine the fit for purpose level of validation necessary for the assay, performing a feasibility comparison of kits, and finally validating the chosen kit with necessary adaptations. These guidelines provide the most comprehensive and valid approach to adaptation of a biomarker kit to date. However, recent publications recommend using guidelines developed by the FDA and CLSI for other aspects of biomarkers used in clinical studies for their use in bioanalytical assays [12]. The recommendations in this paper focus on using the principle of fit-for-purpose context of use (COU) for adapting commercial kits for bioanalytical methods.

The FDA Biomarker Qualification Program is a mechanism to qualify and use biomarkers for use in clinical studies. Each biomarker allowed for use in a study should have a specific COU, which is 'a statement that fully and clearly describes the way the medical product tool is to be used and the medical product development-related purpose of use' [13]. The COU of the biomarker in drug development should be established early in assay development. The COU should be specific to the disease or condition being studied, and may include screening for a drug target, assessment of risk, assessment of prognostic outcome, or asses a toxic or safety concern.

For bioanalytical biomarker assays, the level of validation must be sufficient to support the biomarker's COU [13]. For commercial kits, determination of the adaptations necessary for a kit thus starts with a BWP that defines the COU (Table 1). This BWP helps guide decisions in validation as parameters such as sensitivity, precision, study matrix and determination of matrix effects must fit the COU and the disease being studied.

Recent approaches for biomarker validation have also defined three tiers of validation to better fit and provide more flexibility for the different uses of biomarkers during drug development to supplement the traditional two tiers of method qualification and validation [3,12]. Shown in Table 2, these tiers are defined with their recommended test parameters for exploratory validation (Tier I), partial validation (Tier II) and full validation (Tier III). The tier of validation is dependent primarily on whether the COU involves submission to a regulatory body for decision making or is used for internal decision making. Early in the drug development process where biomarkers are used for internal decisions of screening, determination of the drug's mechanism of action and optimization of a drug, a tier I or II validation is recommended to obtain rapid yet reliable information. Later in the process, where decisions of patient safety and drug efficacy are made by a regulatory body, a full tier III validation in accordance with FDA bioanalytical guidelines must be performed regardless of whether the study is clinical or nonclinical. The addition of a middle tier (tier II) allows for greater flexibility in designing assays that fall in a gray area between discovery and regulatory labels while allowing easier validation of an assay up to tier III if the COU is changed.

It is important to note that the validation of the kit is specific to the COU and the disease being studied. Changes to the COU may change the validation tier and thus a new BWP must be drafted to evaluate if new modifications are needed to the assay and kit. This validation workflow provides an updated framework for adapting commercial biomarker kits, which are in accordance with how the biomarker is being used.

Table 1. Biomarker work plan (creatinine case study).

Parameter	
Analyte(s)	**Creatinine**
Context of use (e.g., disease sample screening, lead biomarker screening, target optimization, assessment of susceptibility, assessment of prognostic outcome, assessment of toxicity)	Determination of urine creatinine for calculation and normalization of biomarkers in human urine after dosing of tobacco products (regulatory submission)
De novo method or commercial kit	IVD commercial kit for enzymatic creatinine determination in human plasma adapted to silver DDK by vendor
Multiplexing (Yes/No)	No
Reference material	Well-characterized and traceable NIST standard will replace kit reference material. No activity, purity or concentration bridging necessary
Sample collection, handling and storage	Samples stored at -20°C. No collection and handling considerations. Stability to be determined
Clinical ranges and sensitivity required (Normal → Disease)	~150–1500 μg/ml
MRD	To be determined for urine
LLOQ and ULOQ of samples (reportable range accounting for MRD of kit)	To be determined for urine
Biological variability	Well-documented variability of creatinine in different patient populations
A priori acceptance criteria to quantify anticipated effect	Accuracy: ±20.0% (±25.0% at ULOQ and LLOQ) Precision: ±20.0% (±25.0% at ULOQ and LLOQ)
Sample matrix and volume needed for kit	Human urine. 10 μl × 2
Data interpretation	N/A. Standard BMV acceptance criteria
Summary of validation level and kit adaptations	Tier III full bioanalytical method validation of the silver DDK enzymatic creatinine kit: • 6 STDs prepared and 3 endogenous QCs in human urine. • Calibrator from NIST creatinine standard. • Matrix effect testing in urine instead of serum to determine MRD

BMV: Bioanalytical method validation; DDK: Drug development kit; MRD: Minimum required dilution; NIST: National Institute for Standards and Technology; QC: Quality control.

Table 2. Validation tests for three bioanalytical method validation tiers for commercial kits.

Parameter	BMV tier I exploratory validation	BMV tier II partial validation	BMV tier III full validation
Use for regulatory submission	No (internal decisions only)	No (internal decisions only)	Yes
Reference material	Maintain and monitor quality and activity of reference material	Maintain and monitor quality and activity of reference material. Verify specificity of assay with reference material from a second source	Reference material anchored to a WHO/international reference material when available. Purity, potency and concentration well characterized Verify specificity of assay with reference material from a second source
Calibration curve	≥6 calibrators using surrogate or study matrix spiked with reference material. Adjust calibrators in quantitation range as necessary for COU. Calibration range includes LLOQ and ULOQ concentrations		
Parallelism and/or dilution linearity	Recommended	Required	Required
Selectivity/matrix effect	Spike recovery at 2 levels with 6 lots of matrix from normal subjects	Spike recovery at 2 levels with 10 lots of matrix from normal subjects	Spike recovery at 2 levels with ≥10 lots of matrix from normal and COU disease subjects each
Specificity/cross-reactivity	Recommended	Recommend	Required: compare structurally similar compounds
Precision and accuracy	3 analytical runs	3 analytical runs	6 analytical runs on multiple days with multiple analysts
QC samples	Analytical QCs in surrogate matrix (LLOQ, low, mid, high, ULOQ)	Add endogenous QCs	Add endogenous QC pools from normal and COU disease subjects
Stability in matrix and solutions	Recommended	STS and FTS of endogenous QC	STS, FTS, LTS of stock solutions and endogenous QCs
Lot to lot variability testing	Recommended	Recommended	Screen ≥3 lots
Biomarker work plan	Required		
Validation plan/validation report	Required		

BMV: Bioanalytical method validation; COU: Context of use, FTS: Freeze-thaw stability, LTS: Long-term stability, STS: Short-term stability; QC: Quality control.

Use of multiplexed biomarker kits, while more complex, requires the same framework of forming a BWP with COU statement for determining which tier of validation is necessary for the kit. It is recommended that the number of analytes measured in a single assay be evaluated against the COU. While there is not yet consensus on the exact number, a large number of analytes in a single assay should not be used for regulatory submissions. Late-stage trials are expected to use singleplex assays [6].

Utility of CLSI & FDA guidelines for biomarkers

Another recommendation from recent guidelines is to utilize guidelines from CLSI and CLIA, which have been measuring biomarkers for over 40 years in the clinical and diagnostic laboratory environment [6,12,14].

The push to harmonize CLSI diagnostic assay guidelines for biomarkers with FDA bioanalytical drug development assay guidelines can be applied to the adaptation of commercial kits. In order to harmonize these guidelines, a comparison should be made of the validation requirements for both methodologies and the pros and cons of both approaches [15]. Both the FDA bioanalytical guidance and the CLSI biomarker guidelines have similar detailed reporting, training and record-keeping requirements, as well as a quality system dedicated to ensuring traceable and quality results, and so are very similar in the spirit of the content. However, the goals of CLSI assays and FDA bioanalytical assays are largely separate. CLSI assays are validated to distinguish healthy from diseased individuals, while bioanalytical assays are used to quantitate changes for the determination of drug efficacy and safety [9]. For example, while reference ranges are necessary for clinical diagnostic purposes, reference ranges do not have the same use in drug development studies.

It is important to note that efforts to harmonize CLSI and FDA guidelines refer to the entire process of validation for all biomarker assays. However, some general principles discussed below can be applied to adapting a biomarker kit in method development and validation. Which guidelines should be used are dependent on the tier of biomarker validation. In particular, three CLSI guidelines for immunoassays are well applied to the adaptation of kits for drug development:

- Reference standard traceability: CLSI guidelines for immunoassays stipulate that assays should use traceable national or international standards for reference materials to establish accuracy of assays [16]. Since immunoassays commonly do not have developed such standards, when such material is unavailable, manufacturers should establish in-house measurement standards to support and document the assignment of values for calibrators. These recommendations are highly applicable to the development and validation of kits for drug development. Tier III biomarker validations should use WHO, National Institute for Biological Standards or National Institute for Standards and Technology reference material when available. When unavailable, assignment of reference values should be documented through characterization of purity, potency and concentration of the material by the end user and/or vendor.
- Interference tests: CLSI assays for biomarkers undergo extensive tests for interference for determination of linearity, specificity and matrix effects [17–19]. Use of patient samples and disease matrix to determine matrix effect, fortification of samples with hemoglobin or lipids, and comparison of different anticoagulants used for sample collection are all specified. These tests add valuable information for endogenous biomarkers that may be modified or are subject to interfering compounds during disease processes. Accordingly, our recommendations for the determination of matrix effect and QCs include the use of matrix from the disease specified in the COU in certain validation tiers (Table 2).
- Definition of measurement ranges: the definition and reporting for sensitivity and quantifiable range often differ from vendor to vendor for RUO kits especially in cases when a minimum required dilution is present. CLSI guidelines for IVD kits require definition of a measurement range and a reportable range [16]. The reportable range of samples may exceed the analytical measurement range when dilutions are used. For drug development kits and validation reports, the LLOQ and ULOQ should be consistently defined accounting for dilutions of samples.

Time is right for drug development kits

In most cases neither RUO) kits nor diagnostics IVD kits meet the quality attributes required for drug development. Bioanalytical scientists are left with commercial kits with 'black box' quality attributes to validate biomarker methods. The process can be frustrating, time consuming and resource intensive. Ultimately, the desired efficiency and/or productivity gains, potentially could be attained by utilizing a commercial kit, are lost.

Table 3. Tiered approach to kit development.

Parameter	Gold	Silver	Bronze
Intended use	Full validation	Full/partial validation	Exploratory validation
Reference material	Reference material must be traceable to internationally recognized (i.e., WHO, NIBSC, NIST, etc.)	Reference material must be characterized as per Table 2 section 'Reference material'	Reference material must be characterized as per Table 2 section 'Reference material'
Calibrator diluent	Matrix-based preferred; when surrogate matrix is used absence of matrix effect should be demonstrated	Surrogate matrix; absence of matrix effect should be demonstrated	Surrogate matrix; absence of matrix effect should be demonstrated
Number of calibrators	≥6	≥6	≥6
Parallelism	Optional; when traceable to internationally recognized reference materials	Should be demonstrated using endogenous analyte	Optional
Selectivity/matrix effect	Matrix effect experiments should be performed using the intended matrix	Matrix effect experiments should be performed using the intended matrix	Optional
Specificity/cross reactivity	Cross reactivity to structurally similar analyte should be performed	Cross reactivity to structurally similar analyte should be performed	Optional
Precision and accuracy	Precision and accuracy testing must be performed using the QC samples prepared in endogenous matrix	Precision and accuracy testing must be performed using the QC samples prepared in endogenous matrix	Surrogate matrix is acceptable; endogenous matrix is recommended. Sensitivity accounts for MRD (reportable range)
QC samples	At least 2 levels of QCs must be in endogenous matrix; additional QCs may be provided in surrogate matrix	At least 2 levels of QCs must be in endogenous matrix; additional QCs may be provided in surrogate matrix	Surrogate matrix QCs acceptable
Lot to lot variability of critical reagents	Lot to lot variability should be tested and controlled	Lot to lot variability should be tested and controlled	Recommended
LOQ vs LOD	LOQ should be defined using spike and recovery experiment	LOQ should be defined using spike and recovery experiment	LOD is sufficient
Documentation	Documentation should include CofA for reference material, and kit performance data outlining experimental parameters and results	Documentation should include CofA for reference material, and kit performance data outlining experimental parameters and results	Documentation should include kit performance data outlining experimental parameters and results

CofA: Certificate of analysis; MRD: Minimum required dilution; NIBSC: National Institute for Biological Standards; NIST: National Institute for Standards and Technology; QC: Quality control.

In our opinion, given the rapid growth in biomarker analysis in the drug development environment, this impasse faced by bioanalytical scientists can be a significant opportunity for kit manufacturers. There is a sufficient demand today for kit manufacturers to consider creating purpose-built DDK. We also feel that the bioanalytical community is eager to collaborate with the kit manufacturers to improve this situation. This can be a win–win proposition for the bioanalytical community and the kit manufacturers. There have been calls in the past from the industry for kits developed specifically for drug development in the same vein as kits developed specifically for diagnostics (e.g., pharmaceutical grade) [20]. However, we believe the framework and recommendations for kit contents and assay performance presented here set the foundation for these kits for discussion of the practical considerations at future meetings. This will enable delivery of quality data that will meet increased regulatory scrutiny and ultimately help bring effective and safe drugs to market to benefit patients.

A tiered framework for drug development kits

After the evaluation of currently available kits, we identified the following three key areas of deficiencies of commercial kits:

- Critical reagents including reference materials characterization;
- Verification of assay performance characteristics; and
- Documentation.

This framework for DDKs with tiered essential quality attributes is consistent with the three-tiered validation approach presented in Table 2 and the most recent industry guidelines for biomarkers. For simplicity, we decided to present the tiers as universally understood 'metal' categories: gold, silver and bronze. An example of use of these different categories is presented in Table 3. These recommendations are intended to be the minimal criteria for harmonizing critical reagents characterization, method performance and necessary documentation to enable their

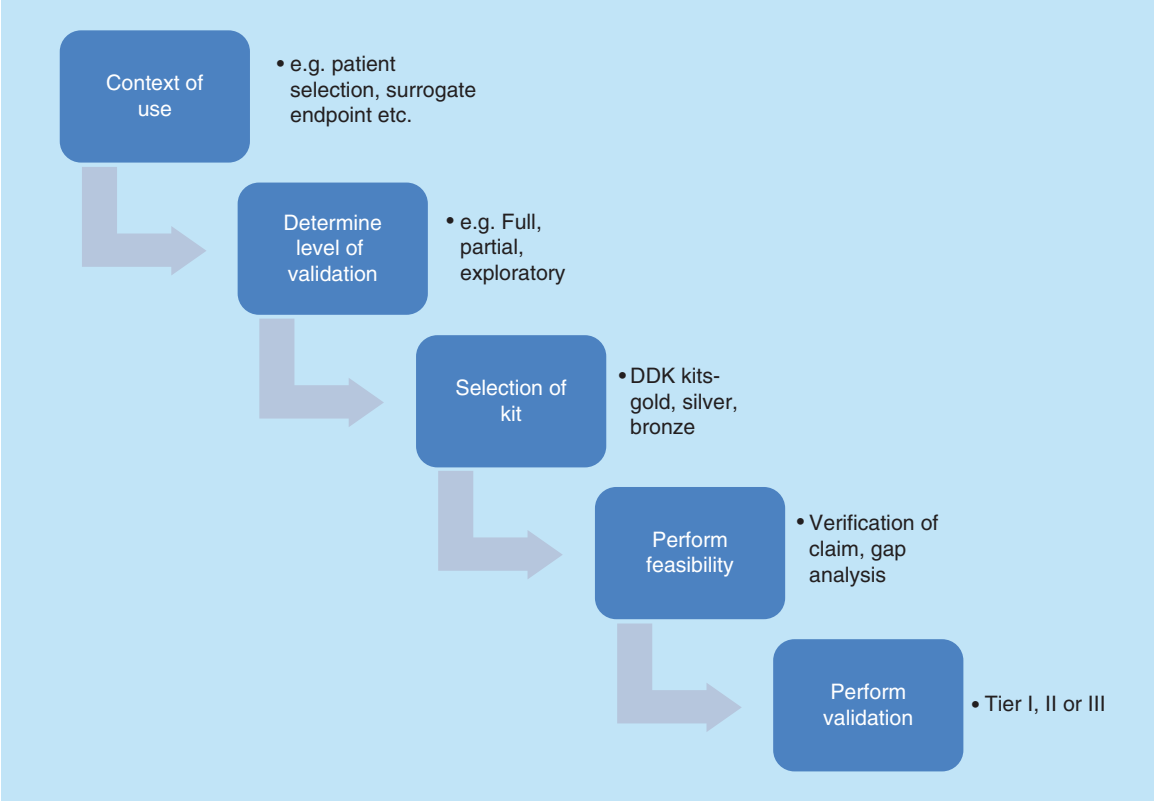

Figure 1. Proposed work flow for validation of biomarkers using commercial kits.

seamless transition and adaptation in a regulatory-compliant bioanalytical lab. Gold DDKs require kits with an international or national reference material which are applicable to a limited amount of analytes. However, this does not preclude the use of kits without international reference standards from being used for full tier III validation for regulatory work. As demonstrated in Figure 1, it is the responsibility of the bioanalytical laboratory to validate the kit chosen to the appropriate tier.

Critical reagents are the essential components of ligand binding methods. In the context of kit-based assays they include reference material/calibrator, coating antibody and detection antibody. Unlike LC–MS, these methods require specific and unique biomolecular interactions between the critical reagents and the reference analyte [21]. The characteristics of these critical reagents can have a significant impact on assay reliability and reproducibility; therefore, the reagents require thorough characterization and documentation. We recommend following the tiered characterization approach taking into account not only scientific needs but also resource needs and cost implications (Table 4).

Assay performance characteristics for commercial RUO kits vary significantly from vendor to vendor and in most cases even from kit to kit within the same vendor. In the case of IVD kits, (premarket approval or 510 K) the performance characteristics are well defined. However, they are often defined for diagnostic applications with limited pathological conditions and are, in most cases, designed for closed instrument systems. This necessitates the need for well-defined performance characteristics for drug development kits that will enable bioanalytical scientists to choose the right kit for the right job. A recommendation on performance characteristics for DDKs is included in Table 3.

As with performance characteristics, the documentation included with the kits are highly variable and often lack critical documentation such as reference material characterization data, assay performance data, assay acceptance criteria and kit release criteria. This framework ensures consistency in documentation for DDKs. For gold and silver kits, documentation should include a certificate of analysis for the reference material, critical reagents and QCs. Kit performance data outlining experimental parameters, results and lot to lot bridging should also be included to assist end users during assay feasibility, gap analysis and validation.

Table 4. Tiered characterization approach for critical reagents.

Attributes	Gold	Silver	Bronze
Reference material			
Molecular weight	Yes	Yes	Yes
Concentration	Yes	Yes	Yes
Potency	Yes	Yes	Optional
Purity	Yes	Yes	Yes
Aggregation level	Yes	Optional	No
Lot to lot bridging	Yes	Yes	Optional
Traceable to internationally recognized (i.e., NIBSC, WHO, NIST, etc.) reference materials	Yes	No	No
Parallelism testing using endogenous biomarker, when possible	Yes	Optional	No
Critical reagents (i.e., coating and detection antibody)			
Concentration	Yes	Yes	Optional
Binding activity	Yes	Optional	Optional
Formulation buffer	Yes	Yes	Optional
Type of antibody including species, isotype	Yes	Optional	Optional
Functional assay	Yes	Optional	No
Stability	Yes	Yes	Yes
Cross-reactivity to structurally similar molecule	Yes	Optional	No
Conjugate incorporation ratio	Yes	Yes	Optional

NIBSC: National Institute for Biological Standards; NIST: National Institute for Standards and Technology.

There are other long-term needs for collaboration between users and vendors. As highlighted previously, some current kits are ineffective because of their lack of sensitivity in study matrices. New technologies and accompanying kits are being developed to address the need for ultrasensitivity for biomarkers [22]. These kits should also attempt to conform to the DDK guidelines. Finally, expansion of training programs that train and certify scientists and labs for various biomarker technologies may be a valuable method to ensure conformity of current and new technologies [23].

Case study: development of a DDK & its subsequent validation

Creatinine is a diagnostic assay most commonly measured in human serum for the assessment of kidney function. However, the correction of urine biomarker measurements using urine creatinine concentration is a widely used application of creatinine in drug development. Most assays for the measurement of creatinine are run using an IVD kit to distinguish diseased patients from healthy patients. These kits do not have the appropriate assay design (e.g., number of standards and QC samples, etc.) that is required to meet FDA BMV guidance necessitating kit modifications.

Using the workflow outlined in Figure 1 and Table 5, we first developed a BWP with a COU statement and determined a tier III full validation was necessary since the data generated from the assay are used for regulatory decision making (Table 1). We collaborated to develop a DDK kit (silver grade) and then adapted the kit to perform a tier III validation. Kit development work was performed by Somru Bioscience, Inc. and regulatory compliant validation work was performed by Celerion, Inc.

The reference material was characterized for MW, purity, potency and concentration. The parallelism test was performed using National Institute of Standard's reference material SRM 3667. The assay was revised to have six nonzero calibration points, and the calibration curve range took into account the target range in smoker population. Accuracy and precision of the assay were demonstrated by running QC. Two of the QC samples were prepared in urine. Additional QC samples were prepared by diluting urine samples using assay buffer. Specificity was evaluated by the fortification of creatinine, hemoglobin and ascorbic acid into QC samples. All the critical reagents were characterized and lot to lot consistency was verified.

Additional adaptations were done at the bioanalytical lab to complete a tier III validation for the specific context of use. The calibrator was switched to use National Institute for Standards and Technology SRM 3667 for better traceability. The validation also adapted the matrix effect testing and the QC samples into the intended matrix of human urine from smoker's population. Endogenous QC samples were made by screening and subsequent pooling of urine lots at appropriate concentration. Since low endogenous levels were not available, human urine was diluted

Table 5. Context of use determines the type of kits and tier of validation.

Context of use	Description	Kit type	Validation status
Discovery	Identification of novel biomarkers associated with a disease or condition and may be screened for many potential uses, including as a target for intervention to prevent, treat or mitigate a disease or condition	RUO kit	Tier I
Early product development – lead optimization	Biomarkers used for compound screening, target validation, etc.	Bronze	Tier 1
Early product development – clinical	Biomarkers used for safety assessments and subject selection based on diseased state for clinical trials	CLIA/FDA-approved kit for diagnostics	Tier II
Early product development – clinical-pharmacodynamic	Pharmacodynamic end point	Silver	Tier II
Surrogate end points	Biomarkers used in late-stage clinical trials and biomarkers used for regulatory decision making	Gold	Tier III
Clinical practice	Biomarkers used by clinicians for uses such as risk stratification, disease prevention, screening, diagnosis and prognosis. The data used as definitive for clinician's decision making	CLIA/FDA-approved kit	Tier III
Clinical practice (late-stage development)	Biomarkers used by clinicians for uses such as risk stratification, therapeutic monitoring and post-treatment surveillance. The data are used as supporting information for clinician's decision making	CLIA or silver	Tier II
Comparative efficacy and safety (i.e., bioequivalence or biosimilars)	Biomarkers used in clinical studies looking at the relative efficacy of two drugs. The data are used as definitive for regulatory approval	Gold	Tier III
Comparative efficacy and safety (i.e., bioequivalence or biosimilars)	Biomarkers used in clinical studies looking at the relative efficacy of two drugs. The data are used in a supporting role for regulatory approval	Silver	Tier II

CLIA: Clinical Laboratory Improvement Amendments; RUO: Research use only.

in assay buffer to create low QC samples. The ULOQ QC was generated by spiking endogenous QC samples with kit calibrator. Accuracy, precision and matrix effect evaluation was performed as per FDA Bioanalytical Method Validation guideline. Finally, short term, freeze–thaw and long-term stability was evaluated using QC samples prepared using endogenous QC samples.

The FDA compliant validation demonstrated the appropriateness of the use of the study matrix, showing an acceptable recovery of creatinine during accuracy and precision. No significant matrix effect was observed in any of the ten human urine lots that were fortified at 618 µg/ml after a minimum required dilution of four. Three QC samples in human urine and ULOQ and LLOQ QC samples in assay buffer across the entire range of the assay were utilized for determining accuracy and precision. Accuracy was shown to range from -4.3 to 0% bias, and precision was shown to range from 3.9 to 5.6%. The LLOQ and ULOQ was reliably shown to be 30 and 600 µg/ml (measurement range) with a reportable range of 120 to 2400 µg/ml. Both long-term and short-term stabilities, bracketing the time period between samples collection and sample analysis, were demonstrated. In addition, incurred sample reanalysis, as required by FDA guidelines, was performed on various studies and results were reproducible for >90% of samples.

This case study illustrates the importance for collaboration between kit manufacturers and the bioanalytical laboratory to achieve successful regulatory compliant assay validation that is appropriate for the COU.

Conclusion

In this article, we present a tiered framework for biomarker commercial kits and their validation based on a biomarker's COU to serve as a starting point for a call for action. The case study of a creatinine DDK by the vendor and subsequent tier III validation by the validating lab presented above demonstrates the feasibility of this approach. However, harmonization and finalization of details (e.g., terminology, specific validation tests for each tier) will require repeated discussions between industry partners and regulatory agencies at future meetings. It is the hope of the authors that this article will help stimulate the discussions necessary among all relevant parties to remove the barriers and to establish a standardized process for the development and adaptation of commercial kits.

Future perspective

Today, and in the next era of drug development, we expect increased use of biomarkers in drug development to expedite scientific and regulatory decision making. Since biomarkers can predict drug pharmacological mechanism,

efficacy and safety more quickly than conventional clinical end points, they hold the potential to reduce drug development costs and substantially reduce the product development timeline. In fact, in some estimates just a 10% improvement in the ability to predict drug failures before clinical trials could save US$100 million in development costs per drug [24].

However, there are two major hurdles we need to overcome in order to reap the benefit of biomarkers. First, we need to overcome the lack of a standardized approach to the validation of biomarkers that take into account technological innovation, and context of use under harmonized regulatory guidelines. In this article, we provide a framework for validation of biomarkers and biomarker kits that take into account the most recent FDA Bioanalytical guidelines, various recent discussions between industry partners and regulatory authorities, and applicable CLSI guidelines.

Second is the lack of standardization for reagents and assay quality attributes for commercial kits. We present a recommendation for the development of 'drug development kits' with standardized quality attributes. We believe an initial investment by manufacturers to design and test three tiers of kits prior to release, similarly tiered characterization of critical reagents and sharing of data through documentation provides feasible steps to increase the utility of commercial kits. A well-defined biomarker kit standardization strategy can be a way for a company to differentiate its product and itself in a highly competitive commercial kit market. Due to intense regulatory demand for quality biomarker data, we believe that in the coming decades, in order for a kit manufacturer to have a significant impact in this market, it will be essential to improve kit quality and implement some form of standardization. It is possible that if kit manufacturers do not make these changes, they put themselves at risk of becoming irrelevant or becoming obsolete. Ultimately, overcoming these hurdles will bring the benefit of biomarkers to patients by accelerating cost-effective drug development.

Executive summary

Challenges in adaptation of commercial kits

- There has been increased use of commercial kits for bioanalysis of biomarkers commensurate with the increased demand for biomarkers in regulated drug development settings.
- The adoption of commercial kits has been hindered due to highly variable quality of commercial kits and lack of harmonized guidance on analytical validation of biomarkers.
- Here we propose a framework for analytical validation of biomarkers that harmonizes the recommendation from US FDA and Clinical and Laboratory Standards Institute. We also propose a framework for the development of drug development kits – a new class of kits with defined quality attributes.

Harmonization of FDA & Clinical and Laboratory Standards Institute guidelines

- New recommendations attempt to harmonize bioanalytical, Clinical and Laboratory Standards Institute and FDA bioanalytical guidelines and the FDA Biomarker Qualification Program.
- We highlight the specific practices that can be adapted to drug development biomarkers.

Proposal for drug development kits

- We propose a three-tiered approach to drug development kits: gold, silver and bronze.
- The details as outlined in Table 3 can significantly improve the acceptance of commercial kits in regulated environment and enable seamless adaptation of the kits to support drug development.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

References

Papers of special note have been highlighted as: • of interest; •• of considerable interest

1. Nowatzke W, Cole TG, Bowsher RR. Systematic analytical validation of commercial kits for the determination of novel biomarkers for clinical drug development. *Bioanalysis* 2(2), 237–247 (2010).
2. Hayashi K. Analyzing global trends of biomarker use in drug interventional clinical studies. *Drug Discov. Ther.* 6(2), 102–107 (2012).
3. Khan MU, Bowsher RR, Cameron M *et al.* Recommendations for adaptation and validation of commercial kits for biomarker quantification in drug development. *Bioanalysis* 7(2), 229–242 (2015).

4. US FDA. *Guidance for Industry: Bioanalytical Method Validation* (2013). www.fda.gov/downloads/Drugs/Guidance/ucm070107.pdf
5. Patel P, Kuzmanov U, Mital S. Avoiding false discovery in biomarker research. *BMC Biochem.* 17(1), 17 (2016).
6. Song A, Lee A, Garofolo F *et al.* 2016 White Paper on recent issues in bioanalysis: focus on biomarker assay validation (BAV): (Part 2 – hybrid LBA/LCMS and input from regulatory agencies). *Bioanalysis* 8(23), 2457–2474 (2016).
7. Saloumeh FK, Zhu Y. Commercial biomarker assays: friend and foe. *Bioanalysis* 8, 2351–2362 (2016).
8. Arnold ME, Booth B, King L, Ray C. Workshop report: Crystal City VI – bioanalytical method validation for biomarkers. *AAPS J.* 18(6), 1366–1372 (2016).
9. Lee JW, Devanarayan V, Barrett YC *et al.* Fit-for-purpose method development and validation for successful biomarker measurement. *Pharm. Res.* 23(2), 312–328 (2006).
10. Jani D, Allinson J, Berisha F *et al.* Recommendations for use and fit-for-purpose validation of biomarker multiplex ligand binding assays in drug development. *AAPS J.* 18(1), 1–14 (2016).
11. Hougton R, Gouty D, Allinson J *et al.* Recommendations on biomarker bioanalytical method validation by GCC. *Bioanalysis* 4, 2439–2446 (2012).
12. Piccoli SP, Neoteric Consulting & John Michael Sauer, Critical Path Institute. *Points to Consider Document: Scientific and Regulatory Considerations for the Analytical Validation of Assays Used in the Qualification of Viomarkers in Biological Matrices*. Biomarker Assay Collaborative Evidentiary Considerations Writing Group (2017). https://healthpolicy.duke.edu/sites/default/files/atoms/files/white_paper_6_14_17_0.pdf

•• **From the US FDA and pharma presents the most recent considerations for bioanalytical use of biomarkers using context of use and forms the basis for the framework for biomarker kits proposed here.**

13. FDA-NIH Biomarker Working Group. *BEST (Biomarkers, EndpointS, and other Tools) Resource Glossary*. Silver Spring, MD, USA (2016).
14. Garofolo W. The decennial index of the White Papers in bioanalysis: 'a decade of recommendations (2007–2016)'. *Bioanalysis* 9(21), 1681–1702 (2017).
15. Lynch KL. CLSI C62-A: a new standard for clinical mass spectrometry. *Clin. Chem.* 62(1), 24–29 (2016).
16. *Clinical Evaluation of Immunoassays; Approved Guideline (2nd Edition).* Clinical and Laboratory Standards Institute, NJ, USA (2008).
17. *Evaluation of Matrix Effects; Approved Guideline (2nd Edition).* Clinical and Laboratory Standards Institute, NJ, USA (2005).
18. *User Verification of Precision and Estimation of Bias; Approved Guideline (3rd Edition).* Clinical and Laboratory Standards Institute, NJ, USA (2008).
19. *Interference Testing in Clinical Chemistry; Approved Guideline (3rd Edition).* Clinical and Laboratory Standards Institute, NJ, USA (2008).
20. Bowsher RR, Nowatzke W, Sailstad JM, Khan MU. Application of commercial research-grade biomarker assays in drug development: is it time to create 'pharmaceutical-grade' kits? *Bioanalysis* 4(20), 2427–2430 (2012).
21. Nowatzke W, Woolf E. Best practices during bioanalytical method validation for the characterization of assay reagents and the evaluation of analyte stability in assay standards, quality controls, and study samples. *AAPS J.* 9(2), E117–E122 (2007).
22. Simon S, Ezan E. Ultrasensitive bioanalysis: current status and future trends. *Bioanalysis* 9, 753–764 (2017).
23. Lynch HE, Sanchez AM, D'Souza MP *et al.* Development and implementation of a proficiency testing program for Luminex bead-based cytokine assays. *J. Immunol. Methods* 409, 62–71 (2014).
24. Myoshko D. Pharma R&D is now biomarker-driven. PharmaVOICE (2016). www.pharmavoice.com/article/2016--01-biomarkers/