Available online at www.sciencedirect.com
ScienceDirect
Toxicology 245 (2008) 219–223
TOXICOLOGY
www.elsevier.com/locate/toxicol

# Strategic paths for biomarker qualification

Federico M. Goodsaid [a,*], Felix W. Frueh [a], William Mattes [b]

[a] *Genomics Group, Office of Clinical Pharmacology, Office of Translational Science, Center for Drug Evaluation and Research, FDA, United States*
[b] *Critical Path Institute*

Received 13 November 2007; accepted 17 December 2007
Available online 6 January 2008

**Abstract**

Biomarkers may be qualified using different qualification processes. A passive approach for qualification has been to accept the end of discussions in the scientific literature as an indication that a biomarker has been accepted. An active approach to qualification requires development of a comprehensive process by which a consensus may be reached about the qualification of a biomarker. Active strategies for qualification include those associated with context-independent as well as context-dependent qualifications.


*Keywords:* Biomarkers; Consortia

## 1. Introduction

Several definitions have been published for what a biomarker is (Lesko and Atkinson, 1999; Lee et al., 2006) and also for how to qualify exploratory biomarkers (Goodsaid and Frueh, 2006; Wagner et al., 2007). The focus of this review is on strategies developed to qualify biomarkers, and how these strategies are encouraging new biomarkers in drug development. Independently of the definitions we use for biomarkers and their qualification, the urgent need to improve tools available to accelerate development of new and better drugs (Orr et al., 2007; Goodsaid and Frueh, 2007a) is reflected in the intensive research on biomarkers within the pharmaceutical industry and in academic and government labs. Biomarker use requires some level of biomarker consensus, whether this consensus is generated internally within pharmaceutical companies for internal decision-making or externally among scientists and clinicians.

Acceptance of biomarkers has often not been linked to a comprehensive process for qualification. A lack of a comprehensive process for qualification results in some biomarkers in use today that have not been qualified in the specific context they are to be used. Different strategies have been proposed (Stokes et al., 2002, 2006; Wagner et al., 2007; Goodsaid and Frueh, 2007b) for biomarker qualification. Some of these propose context-independent (Stokes et al., 2002, 2006), while others propose context-dependent (Wagner et al., 2007; Goodsaid and Frueh, 2007b) qualification paths.

Current practice in biomarker acceptance is closely associated with professional debate often initiated at the level about whether qualification for specific biomarkers should be discussed at all. While a biomarker must be defined both as a test measurement as well as a preclinical or clinical interpretation of the result from this measurement, professional debate often confounds measurement with interpretation. For example, the detection of a specific molecular species is often discussed in isolation from the interpretation of this detection in a specific preclinical or clinical context.

The unstructured process by which biomarkers are currently accepted has lead to a common perception that biomarker qualification is a process that is both hopelessly complex and poorly understood. A biomarker qualification regulatory process must be clearly defined, with explicit metrics for incremental success as qualification data is generated and interpreted. New biomarkers cannot be efficiently developed and employed if biomarker data introduced through and IND or an NDA can be potentially inaccurately interpreted by regulatory reviewers. A uniform, consistent and explicit interpretation of a biomarker measure-

* Corresponding author at: Genomics Group, Office of Clinical Pharmacology, Office of Translational Science, Center for Drug Evaluation and Research, FDA, 10903 New Hampshire Avenue, MD 20903-0002, United States.
Tel.: +1 301 796 1535.
*E-mail address:* Federico.Goodsaid@fda.hhs.gov (F.M. Goodsaid).

0300-483X/$ – see front matter © 2007 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.tox.2007.12.023

ment in a specific context must be an integral part of biomarker qualification. Finally, biomarker qualification is easily justified in drug-test co-development, but efforts by individual companies to qualify biomarkers may often run into untenable costs associated with these efforts.

## 2. Context-independent qualification

Biomarkers are currently accepted through professional preclinical or clinical consensus, often after many years of debate and discussions that may focus less on the actual scientific and clinical data supporting qualification than in the complex needs of organizations, scientists and clinicians proposing their use. A qualification model independent of specific contexts required in different organizations was developed by the Interagency Coordinating Committee for the Validation of Alternative Methods (ICCVAM) (Stokes et al., 2002, 2006).

ICCVAM evolved to encourage the development and qualification of test methods to replace animal testing in toxicology. These test methods cover a range of biomarkers in toxicology. A model was developed by ICCVAM for test method qualification on the basis of a public process through which test method data can be shared. The results of this process are communicated to the 15 Federal agencies that are currently members of ICCVAM, and these agencies are then responsible for communicating the results to their respective regulated industries. This process is context-independent: qualification requires a consensus over applications for these biomarkers throughout different regulatory agencies. It is important for the replacement of animal testing in toxicology, where the proposed test method can be qualified across multiple contexts of use.

Much of the data required for biomarker qualification in the context of drug development is generated by scientists and clinicians associated with the pharmaceutical industry. The value of these biomarkers is closely identified with the context in which they are used in drug development. The flow of information needed for qualification of biomarkers in drug development is opposite to that expected from a context-independent model. Regulated industries can share confidential information with their regulatory agencies, but may be less likely to share this information in a public qualification process.

The International Life Science Institute Health and Environmental Sciences Institute (ILSI/HESI) assembled a technical committee for the development and application of biomarkers of toxicity. (http://www.hesiglobal.org/Committees/TechnicalCommittees/Biomarkers/#mission) This committee has focused on data generated by its members to better understand the analytical and preclinical performance of biomarkers of toxicity, with an initial focus on troponins and biomarkers of nephrotoxicity. This committee has also considered strategies for qualification and regulatory evaluation of these biomarkers, and presented an update on the committee's research programs to the FDA in April 2007. The work of this committee has shown the need to focus on the qualification of biomarkers in the context in which these are to be applied.

## 3. Context-dependent qualification

Biomarkers selected for qualification should reduce cycle time, development cost or adverse events in drug development (Lee et al., 2006). The initial impact of the application of a new biomarker on the cost of a drug development process may very well be negative. The impact of biomarkers on drug development may be expected to become increasingly clear over several development programs. Context-dependent biomarker qualification addresses the need to justify and prioritize biomarkers likely to have a major impact on drug development. Newly qualified diagnostic and mechanistic preclinical drug safety biomarkers have an immediate impact on drug development. However, any new biomarker with a positive impact on drug development is a good candidate for qualification. The prioritization of these biomarkers for qualification will depend primarily on how much these can help with more, better and safer drugs.

An accurate definition of biomarker context will contribute to a robust qualification process. The accuracy of biomarker context includes how it is measured and how it is used. We need to use an accurate measurement of the outcome for which we are trying to develop new biomarkers. Histopathology data will often be a valuable standard in the development of biomarkers for preclinical drug safety assessment. Clinical qualification will depend on the context of use for the new biomarker. Biomarkers which help guide therapeutic decisions need to be qualified around this decision-making process. Those which are being developed as improved diagnostics need to be compared against the diagnostic standard in current use.

If a biomarker is available currently to aid in drug development, it serves as a useful reference point in the development of new biomarkers. New biomarkers make sense when they provide better or different information than the biomarkers in current use. Their use may have analytical, nonclinical or clinical advantages, or may simply expand the information available from current biomarkers. If a biomarker can be qualified to show that its specificity is comparable to that for a currently available biomarker, it is also possible to determine how its sensitivity compares with that of current biomarker. It is possible that a new biomarker will provide either better diagnostic sensitivity or/and the potential for a predictive application.

This context-dependent qualification will be difficult if it is isolated from a rigorous analysis of scientific and clinical qualification data. Evidentiary standards for these data will also be context-driven. Evidence for biomarker qualification is associated, for example, with whether a biomarker is predictive, diagnostic, or mechanistic. It is likely, for example, that evidence for predictive biomarkers will be more difficult to obtain than that for mechanistic biomarkers. Predictive biomarkers are anchored on outcomes reported as a function of time and dose. Positive and negative results from predictive biomarkers have predictive value.

Biomarker qualification precedes test qualification. Analytical validation is an integral part of biomarker qualification, but there is an excellent chance that if it is a novel biomarker there will be no off-the-shelf tests available for it. An aggressive effort for biomarker qualification cannot always depend on commer-

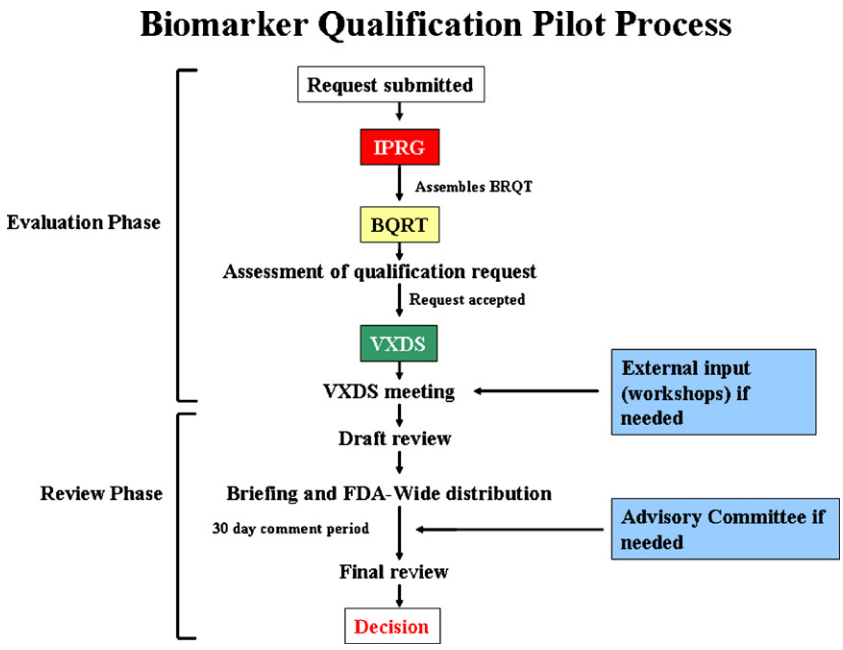


Fig. 1. Pilot process for biomarker qualification.

cially available tests, since these may not be developed before an accurate estimate can be made for their market. Proprietary data on key biomarkers in drug development may not reach publication until after data suitable for their qualification are available.

Successful biomarker contexts of interest to the pharmaceutical industry are related to their impact on better, faster and cheaper drug development and safer drugs. These biomarkers will be qualified on the basis of the data that support their use in the context in which they are qualified. Pharmaceutical biomarker contexts may be closely associated with the pharmaceutical industry and their corresponding regulatory agencies.

## 4. A Context-dependent biomarker qualification process

The FDA is testing a pilot process for biomarker qualification (Goodsaid and Frueh, 2007b, submitted for publication). Voluntary exploratory data submissions (VXDS) (Orr et al., 2007; Goodsaid and Frueh, 2007a) at the FDA have stressed the need for a regulatory path from exploratory biomarkers to biomarkers qualified in a specific context. This pilot process is focused on the specific needs of the regulatory environment to ensure scientifically accurate and clinically (or pre-clinically) useful decision-making. The goal of the pilot process for biomarker qualification is a consensus at the FDA on the interpretation of biomarker measurements submitted with IND/BLA/NDA. A process such as this is likely to be more efficient and transparent than a process of qualification on a case-by-case basis. As a pilot process, it needs to be tested with multiple qualification proposals to assess its performance (see Fig. 1). The process follows a number of steps:

(1) Submission of an initial letter defining the biomarker, its context and data sources for its qualification. Proposals can be received from Industry, Academic, Government or other scientists and clinicians.
(2) Decision to proceed to a full qualification by the biomarker qualification review team (BQRT).
(3) Full submission of qualification data for review by the BQRT.
(4) VXDS meeting to go over the qualification data and to identify potential information gaps before a full review can be completed for the qualification package.
(5) Review drafted by BQRT.
(6) Internal review at FDA.
(7) Communication of decision to sponsor.

Initial experience with this process has confirmed some of the assumptions leading to it and also underscored the caution needed in its implementation. It is very important to define the biomarker context accurately, stating the objectives for this qualification clearly. Data supporting this context should be supported by studies designed to test the hypothesis underlying this context. The analytical performance of novel biomarker assays needs to be shown as part of the qualification.

A nonclinical submission for qualification needs individual animal data as well as summary tables. Unbiased histopathology results are important in order to establish the correlation between histopathology and the new biomarker measurements. Data received from consortia should share a common lexicon and metrics between different members, as well as a scoring consensus between pathologists from these consortia. Receiver operating characteristic (ROC) curves for sensitivity and specificity data are basic metrics to assess biomarker performance. ROC curves generated for different histopathology score ranges are excellent metrics to help summarize biomarker performance.

## 5. Consortia and qualification

Several consortia have in their goals (Wagner et al., 2007; Goodsaid et al., submitted for publication) context-dependent biomarker qualification. One of them is the NIH/FDA Biomarkers Consortium (Wagner et al., 2007). This Consortium has focused on providing grants for the development and generation of data for qualification of clinical biomarkers in a subset of therapeutic areas.

Another Consortium is the Predictive Safety Testing Consortium (PSTC). This Consortium has developed the legal framework needed share the cost of qualification and to protect intellectual property associated with biomarker qualification. Pre-competitive sharing of qualification data is a cost-effective process with which to quickly reach a data threshold for qualification. Qualification includes data generated by companies in industries regulated by specific regulatory agencies. Proprietary qualification data can be shared between companies in regulated industries and their regulatory agencies. These data may or may not be published.

The PSTC was announced by the C-Path Institute and the FDA in March 2006 (see http://www.fda.gov/bbs/topics/news/2006/NEW01337.html). C-Path developed over the first year of this predictive safety testing consortium the legal framework needed for data sharing between its members. The PSTC focus has been to develop the ways and means to overcome these and other hurdles in biomarker qualification. Its goals include:

### *5.1. Definition of context dependence for exploratory biomarkers*

A clear definition of application context requires an accurate understanding of what a biomarker measurement is for and the scientific, preclinical or clinical evidence supporting this measurement. The PSTC will work on the accurate definition of application context for the exploratory biomarkers that the consortium is working on.

### *5.2. Collaboration with regulatory agencies in the development of a process for biomarker qualification*

The PSTC is working with regulatory agencies to identify structural conditions required for an efficient and comprehensive biomarker qualification process. The goal is to replace the complex, unstructured and open-ended process associated thus far with biomarker acceptance with a process that will work to qualify biomarkers in narrow contexts within drug development and regulatory review.

### *5.3. Uniform interpretation of biomarker qualification context by reviewers across regulatory agencies*

The biomarker qualification data generated by the PSTC is submitted at the same time to the FDA and EMEA to allow both Agencies to review and discuss the biomarker data in the context in which the biomarker was qualified. These reviews will be useful examples for a uniform and efficient biomarker qualification process.

### *5.4. Contribution by multiple pharmaceutical companies with data and samples for biomarker qualification*

PSTC members share data and samples to accelerate biomarker qualification by reducing the cost per company to a feasible level.

The FDA and EMEA provide representatives that serve as observers and consultants for both the PSTC steering committee as well as each of its working groups. PSTC members complement each other in the information and samples they can provide for biomarker qualification through different working groups. The PSTC currently has working groups in nephrotoxicity, hepatotoxicity, vascular injury, myopathy, and genotoxic/non-genotoxic carcinogenicity. These working groups are supported by teams focusing on data management and (clinical) translational strategies. Key to the PSTC's structure is a consortium agreement that as a legal document addresses key concerns such as membership, anti-trust issues, governance, funding, information sharing, confidentiality, publicity and intellectual property. Project agreements represent the specific legal documents covering Working group research projects. Governance of the consortium is handled by an advisory committee where each member company has one vote.

Each of the working groups consider prior experience in developing programs to qualify promising biomarker assays, and the approach of each working group is influenced by the particular nature of the scientific question. The vascular injury working group, in addressing the disparities between pathologies seen in various pre-clinical species and the absence of clear relevance to human disease, has been focusing on biomarkers that not only correlate with the observed pathology but also can be assayed in several species. The carcinogenicity working group has critically examined certain published genomic signatures of non-genotoxic carcinogenicity by evaluating their performance with member company genomic data. The goal in this effort is to develop a robust test that predicts the occurrence of liver tumors in the rodent two-year carcinogenicity bioassay from gene expression measurements made on short-term (14 days or less) studies. The results of this examination have been encouraging enough to suggest that the signatures be re-assessed using a common gene expression platform (e.g. quantitative RT-PCR). The Hepatotoxicity working group has shared data on several assays, and is initiating a cross-qualification effort on four enzymatic assays where extensive internal data has indicated promise for detection of liver injury with more sensitivity and specificity than standard tests. The nephrotoxicity working group has examined a panel of 23 urinary protein assays where the performance was extensively compared with that of standard markers (e.g. BUN and serum creatinine) and histopathology.

An efficient and comprehensive process for biomarker qualification would constitute an engine for delivering new tools for

both regulators and drug developers. The key to the establishment of this process is the availability of actual test cases that will allow development and refinement of robust procedures. To that end, the nephrotoxicity working group of the PSTC submitted a full biomarker qualification package for seven biomarkers of nephrotoxicity to the pilot process for biomarker qualification at the FDA and EMEA in July, 2007. The qualification package was discussed with these regulatory agencies through a voluntary exploratory data submission (VXDS) meeting. Initial review suggested additional information may be needed for this submission, including a review of clinical studies of five of the biomarkers. The pilot process for biomarker qualification is working as a dialog between the regulatory agencies and the pharmaceutical industry and it can be expected that it will continue to evolve as more and diverse datasets are submitted. It is also hoped that this process will encourage collaboration between different consortia developing biomarker qualification data.

## 6. Conclusion

There are several possible strategies for biomarker qualification. Some of these are context-independent, while others are context-dependent. Proprietary data for qualification of biomarkers associated with regulated industries may be accessed in biomarker qualification processes where regulatory agencies share these data with their regulated industries. The pilot process for biomarker qualification currently tested by the FDA is a context-dependent process. Several consortia have generated qualification data to be submitted to test this process.

## References

Goodsaid, F., Frueh, F., 2006. Process map proposal for the validation of genomic biomarkers. Pharmacogenomics 7, 773–782.

Goodsaid, F., Frueh, F., 2007a. Implementing the US. FDA guidance on pharmacogenomic data submissions. Environ. Mol. Mutagen. 48, 354–358.

Goodsaid, F., Frueh, F., 2007b. Biomarker qualification pilot process at the US Food and Drug Administration. AAPS J. 9, E105–E108.

Goodsaid F., Frueh F. submitted for publication. Questions and answers about the pilot process for biomarker qualification at the FDA. Drug Discov. Today Technol.

Goodsaid F., Frueh F., Mattes, W. submitted for publication. The predictive safety testing consortium: a synthesis of the goals, challenges and accomplishments of the critical path. Drug Disc. Today Technol.

Lee, J.W., Devanarayan, V., Barrett, Y.C., Weiner, R., Allinson, J., Fountain, S., Keller, S., Weinryb, I., Green, M., Duan, L., Rogers, J.A., Millham, R., O'Brien, P.J., Sailstad, J., Khan, M., Ray, C., Wagner, J.A., 2006. Fit-for-purpose method development and validation for successful biomarker measurement. Pharm. Res. 23, 312–328.

Lesko, L.J., Atkinson Jr., A.J., 1999. Use of biomarkers and surrogate endpoints in drug development and regulatory decision making: criteria, validation, strategies. Annu. Rev. Pharmacol. Toxicol. 41, 347–366.

Orr M.S., Goodsaid F., Amur S., Rudman A., Frueh F. 2007. The experience with voluntary genomic data submissions at the FDA and a vision for the future of the voluntary data submission program.

Stokes, W.S., Schechtman, L.M., Hill, R.N., 2002. The interagency coordinating committee on the validation of alternative methods (ICCVAM): a review of the ICCVAM test method evaluation process and current international collaborations with the European centre for the validation of alternative methods (ECVAM). Altern. Lab. Anim. 30 (Suppl. 2), 23–32.

Stokes, W.S., Schechtman, L.M., Rispin, A., Gupta, K., Hamernik, K., Tice, R., Wind, M., 2006. The use of test method performance standards to streamline the validation process. ALTEX 23 (Suppl.), 342–345.

Wagner, J.A., Williams, S.A., Webster, C.J., 2007. Biomarkers and surrogate end points for fit-for-purpose development and regulatory evaluation of new drugs. Clin. Pharmacol. Ther. 81, 104–107.