Biomarker Validation: Context and Complexities

Lisa M. McShane

Biomarkers are key to personalizing medical care and accelerating the development of new drugs and other medical products. Sophisticated technologies permit assessment of biomarkers on unprecedented scales in terms of numbers and types of measurements. These measurements may capture biology at the DNA, RNA, or protein levels; imaging or digital monitoring may allow assessment of biomarkers reflecting physiologic structures, processes or functions. Coupling these rich data with advanced analysis methods, there would seem to be endless opportunities for biomedical discoveries leading to new therapeutic approaches. Yet, validation of a single biomarker or constellation of biomarkers (i.e., a "signature") as suitable for a particular purpose is often a painstaking process fraught with challenges. This commentary reflects on the complexities of biomarker validation, including the multitude of issues and challenges highlighted by authors of the articles in this special issue of the *JLME*.

"Biomarker validation" is a phrase often spoken but seldom adequately defined. The word "validation" seems to project some notion of credibility or credentialing, yet it is a vacuous term in the absence of context. Not until specific role and context for a biomarker have been proposed can a sensible discussion of validation occur. In biomedical research, a helpful first cut is the distinction between use of a biomarker as a medical product development tool and use of a biomarker (or more precisely, a clinical test based on it) to guide clinical care decisions. Articles in this issue touch on both of these categories. Discussion here aims to clarify concepts and emphasize the criticalness of biomarker role and context when assessing validation status of a biomarker.

As defined by the FDA-NIH BEST glossary,[1] medical product development tools are "methods, materials, or measurements used to assess the effectiveness, safety, or performance of a medical product." Examples of biomarker-based tools include those used for enrichment for clinical trial eligibility, for early detection of drug-related safety signals, and as surrogate endpoints. Several articles in this issue commented specifically on biomarkers as surrogate endpoints, which are endpoints "used in clinical trials as a substitute for a direct measure of how a patient feels, functions, or survives."[2] The U.S. Food and Drug Administration uses the term "qualification" to describe "a conclusion, based on a formal regulatory process, that within the stated context of use, a medical product development

Lisa M. McShane, Ph.D., *is with the Biometric Research Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute in Bethesda, MD.*

The Journal of Law, Medicine & Ethics, 47 (2019): 388-392. © 2019 The Author(s)
DOI: 10.1177/1073110519876169

tool can be relied upon to have a specific interpretation and application in medical product development and regulatory review."[3] Qualification has a similar flavor as the notion of validation but with specific regulatory requirements and implications. FDA has shared its current views on evidence needed for biomarker qualification in a recent publication.[4]

FDA's thinking on evidentiary criteria specifically for surrogate endpoints continues to evolve and undergo clarification. Discussions have occurred in public forums,[5] resources are in development,[6] and recently the FDA published a table of surrogate endpoints that were the basis of drug approval or licensure.[7] Important considerations mentioned in text accompanying the FDA surrogate endpoints table are: "The acceptability of these surrogate endpoints for use in a particular drug or biologic development program will be determined on a case-by-case basis. It is context dependent, relying in part on the disease, studied patient population, therapeutic mechanism of action, and availability of current treatments. A particular surrogate endpoint that may be appropriate for use in a particular drug or biologic clinical development program, should not be assumed to be appropriate for use in a different program that is in a different clinical setting." Hey and colleagues note in this special issue that it would have been helpful if the FDA's surrogates table had explicitly specified the definitive clinical endpoint for which each surrogate was considered an acceptable substitute.[8] This is a critical component of context of use for a surrogate endpoint. Further elaboration on the surrogates table would be a welcome addition. Lack of clear description of context of use and proper alignment of evidence to that use risks misapplication of biomarkers, potentially leading medical product development efforts down wrong paths or harming patients.[9]

Biomarkers underpin a large majority of clinical tests used to guide clinical management decisions for patients. Processes to validate a biomarker-based test for clinical use and to qualify a biomarker as a medical product development tool accepted by regulatory authorities have several aspects in common, but there are also important differences. Three main criteria are essential for validation of a clinical biomarker test: *analytic validity*, *clinical validity*, and *clinical utility*.[10] *Analytic validity* refers to establishing that the performance characteristics of the locked down test (i.e., completely specified) are acceptable in terms of sensitivity, specificity, accuracy, linear range, precision, reproducibility, robustness to pre-analytic factors, as applicable. *Clinical validity* refers to demonstrating that the test result is associated with a clinical outcome or characteristic of interest. *Clinical (medical) utility* refers to establishing that use of the test, when applied according to the intended use, results in a favorable benefit-to-risk balance for the patient, i.e., the patient has a better overall outcome when the biomarker test is used to guide clinical management than when it is not.

Unlike drugs, which must undergo regulatory review to establish safety and effectiveness for clinical use, regulatory oversight of many biomarker tests is minimal.

Unlike drugs, which must undergo regulatory review to establish safety and effectiveness for clinical use, regulatory oversight of many biomarker tests is minimal. When a biomarker test is used in clinical care decisions, CLIA regulations[11] require that the laboratory offering the test is certified to meet certain operating standards. CLIA certification implies that the laboratory has demonstrated that it operates under good laboratory practices, with appropriate quality assurance procedures and adequately trained personnel. While this should help to ensure that biomarker tests offered by these laboratories meet reasonable standards for analytic validity, CLIA certification does not focus on evidence for clinical validity or utility. Nor have the vast majority of clinical biomarker tests ever been reviewed by the US FDA, as many are offered under the setting of laboratory-developed tests where FDA often exercises enforcement discretion in its oversight; however, FDA does have the authority to halt marketing of clinical tests if problems with them come to its attention. An important class of biomarker tests that are regularly reviewed by the FDA are companion diagnostics used to select patients who are most likely to benefit from a particular drug or class of drugs. Although, regulatory oversight of these companion diagnostics is in some sense a spillover from regulatory oversight requirements for drugs. Consequently, the biomedical research community and related professional societies and guidelines bodies bear a heavy responsibility for evaluating clinical validity and utility of many biomarker tests in use. This situation has led to fragmented and sometimes inconsistent processes and standards for validation of biomarker tests.

The Journal of Law, Medicine & Ethics, 47 (2019): 388-392. © 2019 The Author(s)

https://doi.org/10.1177/1073110519876169 Published online by Cambridge University Press

Many biomarkers (or tests based on them) fall in the intersection of those with potential value as drug development tools and those with potential to underpin a biomarker test having clinical utility. Comparison of evidentiary requirements for these two situations helps to illustrate the important role of context for biomarker validation or qualification. Prognostic and predictive biomarkers often fall in this intersection. *Prognostic* biomarkers are those "used to identify likelihood of a clinical event, disease recurrence or progression in patients who have the disease or medical condition of interest."[12] *Predictive* biomarkers are those "used to identify individuals who are more likely than similar individuals without the biomarker to experience a favorable or unfavorable effect from exposure to a medical product or an environmental agent."[13] Biomarkers serving in these roles may be used in a drug development program to enrich the population of patients eligible for a clinical trial of an experimental therapy to assess whether it improves outcomes relative to standard therapy for the disease of interest. A prognostic biomarker could be used to enrich the trial population for those patients at greatest risk of unfavorable outcomes or events. This strategy not only focuses on those patients who might most need new therapies, but it can lead also to a clinical trial that is smaller or of shorter duration because the number of observed events drives the statistical power of a clinical trial with a time-to-event endpoint. The same prognostic biomarker might be used somewhat differently when incorporated into a clinical test. For example, in the adjuvant setting for some early stage cancers, there is great interest in identifying patients who have such favorable clinical outcome following surgery alone, that they could forgo adjuvant systemic therapy. The analytical performance of the biomarker test, the exact way in which the biomarker is measured and test results are interpreted, including any cut-points applied to the biomarker value for clinical decision-making, and other clinical or pathological characteristics of the patients must all be considered to ensure an appropriate benefit-to-risk balance for patients. Analogous considerations apply for predictive biomarkers, which when used for clinical trial enrichment, may increase the odds that enrolled patients have potential to benefit from the experimental therapy. If the clinical trial is positive, it is possible that further refinements would be needed before a test based on the biomarker would be deemed validated for use as a companion diagnostic in routine clinical practice. These scenarios illustrate the need for a clear understanding of context of use, availability of sufficient data of the right types, and use of analysis approaches that appropriately evaluate the risks and benefits inherent in the proposed use of the biomarker.

Fostering a biomedical research environment that is more conducive to biomarker validation will require concerted effort. Resources will be needed to develop, manage, and make available in useful form, large collections of rich data and analytic tools. Collaborations among experts representing multiple clinical and scientific disciplines, in consultation with regulatory authorities, will be essential to support and guide biomarker validation efforts to achieve more efficient medical product development and improved clinical care that leads to better outcomes for patients.

Biomarker study publications typically serve as the primary vehicle to advertise the existence of new biomarker data and discoveries, yet they often fall short of providing the type of information needed for determination of whether data and results generated by a study support validation. Key information, such as the characteristics of the study population, definitions of study endpoints, details of biomarker assay methods and statistical analyses, and clear distinction between pre-specified versus post hoc (data-driven) hypotheses are often not fully reported in published articles, yet this information is essential to assess clinical context and strength of evidence. Better reporting of biomarker studies,[14] and health research studies in general,[15] would make study results more interpretable and facilitate identification of studies that are pertinent to biomarker validation efforts.

Advances in technologies to measure and collect biomarker data have led to generation of large volumes of increasingly complex data that may provide key biological insights, yet these data present challenges for capture, storage, and retrieval for use in biomarker validation efforts and other studies.

The Journal of Law, Medicine & Ethics, 47 (2019): 388-392. © 2019 The Author(s)

Increased openness to (and requirements for) sharing of research data, more resources to store and maintain biomedical data and make it accessible in an ethically appropriate manner (including data from routine clinical care records), and development of new analytical tools capable of handling large and complex data sets, all represent steps in the right direction. Li and Sim[16] make a strong case for how platforms for sharing data from clinical trials, which are generally the most comprehensive and high quality data, can advance the study of biomarkers. Research funders can play an important role in making clinical trial data available, as evidenced by the recently launched U.S. National Cancer Institute NCTN/NCORP Data Archive.[17] The Data Archive currently has individual patient-level data available for request from dozens of phase III cancer clinical trials, including some in the process of being linked with biomarker and imaging data. In contrast, biomarker investigations conducted on "convenience" specimen sets, often lacking any semblance of study design or having basis in a defined clinical cohort, tend to produce data and results of low value for translational research and biomarker validation efforts.[18] Unfortunately, many of the earliest studies from which large-scale biomarker data such as omics data were made publicly available utilized convenience specimen sets, and studies of convenience remain quite common in biomarker research.

Simply making more data and analysis tools available may not be sufficient to move many biomarkers toward validation. Statistically significant associations between biomarkers and clinical outcomes together with suggestion of an intriguing clinical or biological hypothesis may be sufficient to support a publication; however, a huge chasm exists between published biomarkers and validated (or qualified) biomarkers. Successful validation or qualification of a biomarker requires a carefully crafted context of use and access to data for the biomarker measured in the right way, at the right time, on the right specimens from the right patients, and accompanied by key clinical and pathologic variables, treatment information, and clinical outcome data. Mavergames and colleagues[19] describe efforts of the Cochrane Collaboration to curate and synthesize evidence which may support biomarker validation, explaining how Cochrane's Linked Data Project could help advance understanding of surrogate endpoints. Shrager and colleagues[20] argue that "big data" and "artificial intelligence" methods applied naively are not likely to lead to meaningful medical advances, in part due to the complexity of some diseases like cancer and the "curse of dimensionality." Instead, they favor informed data reduction by capture of treatment rationales that incorporate expert knowledge and clinical context. These discussions emphasize that biomarker validation is a complex evidence evaluation process that requires more than volumes of data and advanced data analysis tools.

Fostering a biomedical research environment that is more conducive to biomarker validation will require concerted effort. Resources will be needed to develop, manage, and make available in useful form, large collections of rich data and analytic tools. Collaborations among experts representing multiple clinical and scientific disciplines, in consultation with regulatory authorities, will be essential to support and guide biomarker validation efforts to achieve more efficient medical product development and improved clinical care that leads to better outcomes for patients.

Note

The author has no conflicts to disclose. The views presented in this commentary are those of the author and should not be viewed as official opinions or positions of the National Cancer Institute, National Institutes of Health, or U.S. Department of Health and Human Services.

References

1. FDA-NIH Biomarker Working Group, *BEST (Biomarkers, EndpointS, and other Tools) Resource* (Silver Spring, MD: Food and Drug Administration [US]; Bethesda, MD: National Institutes of Health [US], 2016-), *available at* <https://www.ncbi.nlm.nih.gov/books/NBK326791/> (last visited July 22, 2019).
2. *Id.*
3. *Id.*
4. C. Leptak, J.P. Menetski, J.A. Wagner, J. Aubrecht, L. Brady, M. Brumfield, W.W. Chin, S. Hoffmann, G. Kelloff, G. Lavezzari, R. Ranganathan, J-M. Sauer, F.D. Sistare, T. Zabka, and D. Wholley, "What Evidence Do We Need for Biomarker Qualification?" *Science Translational Medicine* 9, no. 417 (2017): eaal4599.
5. "Biomarkers Consortium — Workshop: Defining an Evidentiary Criteria Framework for Surrogate Endpoint Qualification," Foundation for the National Institutes of Health, *available at* <https://fnih.org/what-we-do/biomarkers-consortium/programs/biomarkers-consortium-workshop-defining-evidentiary-criteria-framework-surrogate-endpoint> (last visited July 22, 2019).
6. "Surrogate Endpoint Resources for Drug and Biologic Development," U.S. Food and Drug Administration, *available at* <https://www.fda.gov/drugs/development-resources/surrogate-endpoint-resources-drug-and-biologic-development> (last visited July 22, 2019).
7. "Table of Surrogate Endpoints that were the Basis of Drug Approval or Licensure," U.S. Food and Drug Administration, *available at* <https://www.fda.gov/drugs/development-resources/table-surrogate-endpoints-were-basis-drug-approval-or-licensure> (last visited July 22, 2019).
8. S.P. Hey, W.B. Feldman, E. Jung, E. D'Andrea, and A.S. Kesselheim, "Surrogate Endpoints and Drug Regulation: What is Needed to Clarify the Evidence," *Journal of Law, Medicine & Ethics* 47, no. 3 (2019): 381-387.
9. S.P. Hey, E. D'Andrea, E. Jung, F. Tessema, J. Lou, B. Gyawali, and A.S. Kesselheim, "Challenges and Opportunities for Biomarker Validation," *Journal of Law, Medicine & Ethics* 47, no. 3 (2019): 357-361.
10. S.M. Teutsch, L.A. Bradley, G.E. Palomaki, J.E. Haddow, M. Piper, N. Calonge, W.D. Dotson, M.P. Douglas, and A.O.

The Journal of Law, Medicine & Ethics, 47 (2019): 388-392. © 2019 The Author(s)

Berg on behalf of the EGAPP Working, "The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Initiative: Methods of the EGAPP Working Group," *Genetics in Medicine* 11, no. 1 (2009): 3-14; L.M. McShane and D.F. Hayes, "Publication of Tumor Marker Research Results: The Necessity for Complete and Transparent Reporting," *Journal of Clinical Oncology* 30, no. 34 (2012): 4223-4232; D.R. Parkinson, R.T. McCormack, S.M. Keating, S.I. Gutman, S.R. Hamilton, E.A. Mansfield, M.A. Piper, P.A. DeVerka, F.W. Frueh, J.M. Jessup, L.M. McShane, S.R. Tunis, C.C. Sigman, and G.J. Kelloff, "Evidence of Clinical Utility: An Unmet Need in Molecular Diagnostics for Cancer Patients," *Clinical Cancer Research* 20, no. 6 (2014): 1428-1444.

11. "Clinical Laboratory Improvement Amendments (CLIA)," Centers for Medicare & Medicaid Services (CMS), *available at* <http://www.cms.gov/CLIA/05_CLIA_Brochures.asp> (last visited July 22, 2019).
12. *Supra* note 1.
13. *Supra* note 1.
14. L.M. McShane, D.G. Altman, W. Sauerbrei, S.E. Taube, M. Gion, and G.M. Clark for the Statistics Subcommittee of the NCI-EORTC Working Group on Cancer Diagnostics, "REporting recommendations for tumor MARKer prognostic studies (REMARK)," *Journal of the National Cancer Institute* 97, no. 16 (2005): 1180-1184; D.G. Altman, L.M. McShane, W. Sauerbrei, and S.E. Taube, "Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK): Explanation and Elaboration," *PLoS Medicine* 9, no. 5 (2012): e1001216.
15. "EQUATOR Network: Enhancing the QUAlity and Transparency of Health Research," Equator Network, *available at* <http://www.equator-network.org/> (last visited July 22, 2019).
16. R. Li and I. Sim, "How Clinical Trial Data Sharing Platforms Can Advance the Study of Biomarkers," *Journal of Law, Medicine & Ethics* 47, no. 3 (2019): 369-373.
17. National Cancer Institute NCTN/NCORP Data Archive, *available at* <https://nctn-data-archive.nci.nih.gov/> (last visited July 22, 2019).
18. R.M. Simon, S. Paik, and D.F. Hayes, "Use of Archived Specimens in Evaluation of Prognostic and Predictive Biomarkers," *Journal of the National Cancer Institute* 101, no. 21 (2009): 1446-1452.
19. C. Mavergames, D. Beecher, L. Becker, A. Last, and A. Ali, "Cochrane's Linked Data Project: How it can Advance our Understanding of Surrogate Endpoints," *Journal of Law, Medicine & Ethics* 47, no. 3 (2019): 374-380.
20. J. Shrager, M. Shapiro, and W. Hoos, "Is Cancer Solvable? Towards Efficient and Ethical Biomedical Science," *Journal of Law, Medicine & Ethics* 47, no. 3 (2019): 362-368.

The Journal of Law, Medicine & Ethics, 47 (2019): 388-392.

https://doi.org/10.1177/1073110519876169 Published online by Cambridge University Press