Precision Medicine

Translating Promising Biomarkers into Oncology Clinical Practice

Tracy G Lively, PhD,² Michael C Sachs, PhD¹ and Barbara A Conley, MD¹

Abstract

There is a clear need for biomarker assays that can assist in the choice of the best treatment for individual cancer patients. Many manuscripts can be found in the literature that describe promising correlations of molecular features with cancer outcomes, but few of these ever translate to assays that come into common clinical use. Some reasons for this lack of development are the failure to define a clear and compelling intended use for the biomarker assay, as well as the paucity of necessary and careful validation studies: the analytical and clinical validation of a given assay, optimization for clinical use, and demonstration of the clinical utility of the assay. These steps often require resources beyond those of the original laboratory that defined the correlation of the biomarker result to clinical outcome. This manuscript details some of the matters that are faced by assay developers, as well as some solutions for improved assay development and validation.

Keywords

Tumor biomarker, clinical assay development, clinical utility, assay validation, precision medicine

Although the literature is replete with biomarker studies that show promise for predicting risk for cancer, response to treatment, or prognosis after diagnosis, the number of biomarker-based tests or assays that are useful in the practice of oncology remains distressingly small.¹ ² A biomarker can be used to predict a patient’s prognosis, which is their expected clinical outcome. A biomarker that provides accurate predictions of prognosis, regardless of treatment, is referred to as prognostic. A predictive biomarker is one that accurately predicts disease outcomes with the application of specific interventions or treatments. Predictive biomarkers are therefore useful for the selection among two or more treatment options. There are many clinical decisions that could plausibly be informed by a reliable biomarker or assay so the rationale for development of biomarkers remains great.

For example, for estimating cancer risk, we can now predict that certain heritable mutations confer increased risk for cancer, such as BRCA 1 or 2. The clinical decision that the assay informs is how to mitigate that risk, either by surgical removal of the organ most at risk, increased invasive or imaging surveillance, or long-term chemopreventive medical treatment. The inheritance of a BRCA 1 or 2 mutation conveys a risk for 18–57 % for the development of breast or ovarian cancers.³ For better preventive planning, biomarkers that could define the risk for cancer more accurately for the individual patient (not just the population of patients with a particular mutation) would also be very useful. Most cancers are not heritable, and for these, it would be desirable to have biomarkers to inform if that cancer in a particular patient will be aggressive, or indolent (prognostic biomarker). Other clinical scenarios in which informative predictive biomarkers would be welcome include the choice of the most efficacious course of adjuvant therapy after surgery for curative intent in an individual patient, the best choice of systemic or radiation therapy in patients with advanced, incurable cancers, and the risk for severe toxicity of a given treatment in that patient.

Some cancers now require the measurement of several predictive biomarkers for treatment planning.⁴ For example, breast cancer requires the assessment of estrogen and progesterone receptor and HER2 amplification. Lung cancer requires the assessment of activating EGFR mutations and ALK translocations. With the advent of next-generation sequencing, we are also faced with how to make the best decision when the data to support action in the presence of a particular molecular abnormality or set of abnormalities are sparse. Such assays may be used routinely in the near future for “personalized” or “precision” treatment of cancer patients—tailoring treatment to each patient’s unique molecular profiles.⁵
The association of a molecular abnormality or feature with a particular clinical outcome is the beginning of the process for many predictive and prognostic tests. However, further development of that insight to a reliable and accurate clinical assay is complex. Most tests or assays do not survive the process for various reasons. Significant technical knowledge and resources are required for success—such knowledge and resources must be specifically sought, and are not always present in the team that initially proposes the promising association. Effective planning must occur to develop an assay properly for use in a clinical test; and missteps can cause significant confusion that persists for years. Worst of all, if a test that is not properly developed is used in clinical care, inaccurate treatment and potential harm to patients could result. If diagnostics are essential to choosing the appropriate treatment (sometimes termed companion diagnostics), a similar degree of attention should be given to their development as is given to the development of cancer drugs. Definitions of terms used in this review are given in the Text Box at the end of this article.

Assay Development

When developing an assay, it is essential to identify a specific intended use. An assay will likely not be equally useful for discerning cancer aggressiveness and choosing the best treatment, nor are assays that measure a particular biomarker equally useful across different cancers (e.g. breast cancer and colon cancer), since the amount of biomarker present, as well as the cutpoint for positivity may vary across tumors. The development of an assay and its intended use should address a clinical need, and a particular clinical decision. At the beginning of development, once a correlation of the assay result with a clinical outcome has been discerned, the developer must consider whether more information about the association is actually needed before committing considerable resources to continue assay development and validation. How will the assay help a clinical decision? What decision will the assay guide? Is this an important decision, not addressed by current tests or clinical judgment? The answers to these questions will help guide the clinical development of the assay, including the design of clinical trials, the patient populations, the type of specimen that will be used, the assay platform, etc.

Additional practical factors will need to be considered. Does the assay measure a prognostic or a predictive marker, or both? How large is the population of patients for which that assay will be used, and what is the prevalence of the analyte(s) that the assay detects? If the assay is for a predictive marker or a companion diagnostic, what would the effect of the targeted treatment be on patients who are “negative” for the marker? What is the turnaround time necessary for clinical use? What type of specimen is needed for the assay and what handling/preservation/storage procedures will be fit for the purpose of the assay as well as for feasibility of clinical use? What is the cost of the assay?

If there is already a “research assay,” will the research assay and use? What is the cost of the assay?

Clinical Validation

Clinical validity refers to how well the assay result, the biomarker, relates to the clinical outcome of interest, such as frequency of relapse, response to therapy, or survival. Here consideration must be given to how the assay will be used to be sure the assessment of its clinical validity is carried out in the appropriate patient population. Assessment of clinical validity relates to the both analytical performance and to the intended clinical use. The assessment of clinical validity, particularly, can be an iterative process (see Figure 1). If the analytical performance of the assay has been designed well, the analytical validation will not likely have to be repeated unless there are changes or upgrades to the assay or pre-analytic requirements. Of course, if the measurement of the analyte or signature no longer correlates at all with clinical outcome when used in a new dataset, the assay itself may need to be redeveloped.

Prospective-retrospective clinical validations refer to the type of study that assesses the performance of the assay on clinical samples from patients who participated in a (completed) clinical trial, and for whom the outcome of the treatment in the clinical trial is known. Prospective-retrospective assessments are sometimes ideal for clinical validation, if they can be performed rigorously in a sample from a well-defined patient population.
Assessment of clinical validity for an assay that measures multiple analytes, such as "omics" assays, can become complex, and the developer must understand certain pitfalls in the development of such predictors. Likewise, continuous-valued biomarkers that require a "cutpoint" that defines "positive" and "negative" may require several sets of clinical data to refine and lock down the cutpoint before use in the clinic. A common pitfall in biomarker studies is optimizing the cutpoint and then evaluating its performance using the same data. This causes the performance estimates to be optimistically biased. This type of bias, called resubstitution bias, can be avoided by carefully separating the data used to optimize the biomarker from that used to assess its clinical utility. This can be accomplished by dividing the data and using one set to optimize the biomarker, and the other set to assess its clinical utility. Adaptive study designs have been proposed in which the biomarker cutpoint is optimized and evaluated in the same trial. Resampling methods such as cross-validation are commonly used to prevent resubstitution bias and provide valid estimates of the biomarker performance.

Clinical Utility

Clinical utility is the determination of whether use of the biomarker assay for decision-making leads to improved clinical results for patients. Does the biomarker add to techniques already in use? Will the use of the biomarker result in a different, beneficial treatment for the patient? If a "biomarker-positive" assay result predicts a group of patients that benefit more from a treatment than the biomarker-negative group, but the biomarker-negative group still derives as much or more benefit from the treatment as they would with any other treatment, there would be no reason to perform the assay for such a treatment decision. On the other hand, if a biomarker predicts benefit of a treatment for one or more subgroups of patients, while defining a group that will not benefit or may even have a worse outcome, that biomarker would be very useful for guiding treatment decisions (see Figure 2).

Prospective-retrospective studies may also inform about clinical utility. If a rigorous protocol for the prospective-retrospective analysis is prespecified and followed, then a high level of evidence supporting the clinical utility of the biomarker can be generated. This strategy allowed, for example, the validation of the clinical utility of detecting KRAS mutation in patients with metastatic colorectal cancer. Prospective-retrospective analysis of tissues from patients with metastatic colorectal cancer who participated in several randomized studies of the efficacy of EGFR monoclonal antibodies (cetuximab or panitumumab) showed that such patients were not likely to benefit from treatment with EGFR monoclonal antibodies if their tumor had a KRAS mutation.

Another example in which prospective-retrospective studies were used to address clinical utility was the development of the breast cancer classifier OncotypeDx. This assay uses reverse transcription polymerase chain reaction (RT-PCR) to assess the expression of 16 cancer-related genes and five reference genes in paraffin embedded early stage, node negative, estrogen receptor-positive tumors to predict the likelihood of distant recurrence. A locked down assay and prespecified study plan were used to evaluate tamoxifen-treated patients from a completed trial of adjuvant tamoxifen versus placebo in node negative, estrogen receptor-positive breast cancer patients. The assay was able to separate patients into low, intermediate, and high risk for distant recurrence-free survival at 10 years. A subsequent prospective-retrospective study evaluated whether high-risk patients, as evaluated by the assay, derived benefit from adjuvant chemotherapy. In this study, samples that had been banked from node negative, estrogen receptor-positive breast cancer patients who participated in a randomized adjuvant clinical trial to evaluate the worth of adding chemotherapy to tamoxifen treatment were used. The study showed that the patients classified as high risk by the assay derived benefit from adding chemotherapy to tamoxifen. Finally, a prospective study (Clinicaltrials.gov Identifier NCT00310180) was conducted to evaluate whether patients classified as intermediate risk by the assay had better distant disease-free survival by adding chemotherapy to hormone therapy in the adjuvant setting. This trial has completed accrual and final results are pending.

Other methods of assessing clinical utility involve the use of "convenience" samples and data. These data may be from series of patients who have not been on clinical trials, but may have been treated similarly and have outcome information, such as patients in a database of an oncology practice or group of practices. Here, one must be cautious of bias that may be present because treatment decisions not made randomly, but potentially based on some unrecorded patient characteristic, or because...
Translating Promising Biomarkers into Oncology Clinical Practice

of provider bias about treatments. If the samples and data have been collected over several decades, both the disease characteristics and treatments may have changed over time.

In any retrospective trial, tissue or other samples may not have been collected and stored in an ideal way for the biomarker assay, and so assessment of their fitness for the assay must be performed.

In the planning of a retrospective study, careful consideration should be given to the study power and sample size as precious archived specimens should not be wasted on a retrospective study that is not likely to evaluate the clinical validity of an assay.

Finally, a prospective assessment (prospective clinical trial or study) could be done to assess clinical utility, which does allow specification of tissue or sample collection and storage, but the timeline may be longer than if retrospective samples/data can be identified. A variety of study designs exist to prospectively evaluate a potentially prognostic or predictive biomarker for clinical utility.20–27 The effective integration of biomarker test development into clinical trials of investigational therapeutic agents requires both planning and multidisciplinary collaboration.22–23 The type of study used to evaluate a biomarker, such as an enrichment or all-comers design, depends on the level and quality of prior evidence, both empirical and theoretical, supporting its utility. The study should have a clear and comprehensive prespecified protocol, including specimen-handling procedures, assay procedures, and statistical analysis plans. Investigators should prespecify a statistic or set of statistics that address the intended clinical use of the biomarker. In the design, careful consideration should be given to the study power and sample size. A prospective study in which patients may be undergoing painful biopsies or unnecessary treatment should not be done unless there is a high probability of definitively answering the scientific question.

Other Considerations

Some necessary steps in test development may be in tension with other goals in the design of clinical trials. For example, a biomarker may be proposed as a predictive marker. If a well-characterized assay with an established cutpoint is available to identify the patient population of interest, it can be used even in early-phase trials to enrich enrollment for likely responders and limit the numbers of patients exposed to an investigational agent.21 But if the assay itself is under development, then establishing that cutpoint and the relevant clinical sensitivity and specificity of the test requires assessing the treatment outcome in at least some marker-negative as well as marker-positive cases.21,22 Even for a truly predictive biomarker, care must be taken not to place excessive confidence in a research grade assay for restricting enrollment to a trial, because of the effects that an imperfect assay can have on development efficiency. False positive cases will decrease statistical power by decreasing the effect size; false negative test results will slow accrual by screening out patients who should have been entered. In some cases it may be necessary to decide whether to prioritize sensitivity (minimizing false negatives) or specificity (minimizing false positives) for a clinical trial.

As another example, there can be tension between the goal of proceeding as quickly as possible from one phase to the next in the evaluation of an especially promising agent, and the need to allow time for the necessary steps in assay development to take place. Two-stage trial designs have been proposed to streamline drug and biomarker co-development, designs in which a biomarker is tested but not used to restrict enrollment in stage one and then, only if needed and if a well-validated assay has been developed, becomes an eligibility criterion in stage two.21 These designs are most applicable to the transition between a randomized phase II and a phase III trial, that is, for later stage trials, and it may not be realistic to expect a fully "seamless" transition from stage one to stage two within the same trial. Such an expectation presupposes that the assay has been fully developed and validated to the point where it is fit for use as an eligibility criterion, and has an Investigational Device Exemption from the US Food and Drug Administration, if necessary,22 even though it might never be needed for stage two. Time may also be required for assessment of interlaboratory comparability if multiple testing laboratories will be involved.

Other aspects of biomarker test development can be accelerated if efforts are initiated in early phase trials specifically for this purpose. Information about biomarker prevalence, assay performance in the tissue of interest, and specimen-handling requirements can be acquired even
before phase I from studies on nontrial specimens or tumor xenografts. As biomarkers are selected for inclusion as correlative laboratory studies in a trial, consideration can be given to how readily a research-grade assay, if developed further, could be incorporated into clinical practice. These considerations include not only the prospect of integrating the assay into the standard clinical workflow, but also the level of interest from commercial partners and anticipated regulatory requirements.

Conclusions

In summary, bringing predictive and prognostic biomarkers in to clinical use requires a thoughtful development plan beginning with the careful consideration of how the assay will inform clinical decision-making, through assay development, and ending with demonstration of clinical utility. Careful planning and execution can assure that the assay is analytically validated, meaning that it measures the biomarker with accuracy and reproducibility. Once analytically validated, it is of interest to show that assay results correlate with the outcome of interest; that it has clinical validity. Throughout the whole process, investigators should keep in mind the clinical need that they intend to address. To show clinical utility, feasibility in the clinic must be considered, including properties of the sample that will be used and the platform, and that the assay use could bring a benefit to the patient.

Definitions

Analytical validity: The extent to which the test detects the analyte of interest, in the clinical context (i.e. stage of disease, specimen type)

Assay: A process by which a molecular feature or set of features is measured in a biospecimen

Clinical utility: The extent to which the use of the test improves the clinical outcome of interest

Molecular biomarker: A molecular feature that “is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention”

Predictive biomarker: A biomarker that predicts response/outcome to a particular treatment; e.g. HER2 amplification and trastuzumab

Prognostic biomarker: A biomarker that predicts outcome regardless of treatment, e.g. OncotypeDX for estrogen receptor-positive early stage breast cancer