nature biotechnology

Lost in validation

Forking over funds for large-scale validation of biomarkers could benefit healthcare payors in the long term.

As the articles in this issue attest, there has been an upswell of excitement surrounding molecular diagnostics and pharmacogenetic tests in recent months. Roche Molecular Systems has received approval for its AmpliChip, which detects polymorphisms in cytochrome P450 (CYP) *CYP2D6* and *CYP2C19* genes that commonly affect drug metabolism, and Genzyme has marketed the first test for monitoring the emergence of mutations (e.g., T315I) in *BCR-ABL* that result in resistance to Gleevec (imatinib mesylate). But it remains the case that only a handful of pharmacogenetic and diagnostic tests have been formally approved by regulatory agencies. And more importantly, even though the literature contains ~150,000 reports of disease-associated molecular markers, there are still very few validated biomarkers of proven and robust clinical utility.

Consider, for instance, cancer diagnostics. Here, as elsewhere, a smattering of companion diagnostics are used routinely to guide treatment selection or dosing: Dako's immunohistochemical assay for HER2/neu overexpression selects patients suited to treatment with Genentech's Herceptin (trastuzumab); Prometheus Laboratories' test to detect polymorphisms in thiopurine methyltransferase guides dosing and patient selection for chemotherapy with 6-mercaptopurine; and Third Wave's *UGT1A1* test predicts patient-specific toxicity of Pfizer's Camptosar (irinotecan).

In cancer screening, only a handful of markers have become widely accepted by the clinical community (these include α-fetoprotein; cancer antigens 15.3, 19.9 or 125; carcinogenic embryonic antigen; Epstein-Barr virus; T/Tn antigen; bladder-tumor–associated antigen; prostate-specific antigen; human papillomavirus and telomerase). The vast majority of putative cancer biomarkers are largely unvalidated, and yet many of them—perhaps hundreds of them—will find their way into lightly regulated 'home-brew' diagnostics. Some of these markers will eventually turn out to be highly useful authentic indicators of the emergence of disease, its progression or its regression. But at this point, it is simply impossible to tell which.

The problem is that there is no established, standardized means for validating the association between a marker (or set of markers) and clinical outcomes. At best, there will be a paper from a reputable research group published in a reputable journal that makes a case for a role of each of the markers in a particular kind of cancer. The evidence may be biochemical or physiological, but it has not, in most cases, been established by extensive clinical trials (that is, >1,000 patients) that correlate the detection of a particular marker with clinical outcomes.

There does not necessarily have to be a causal link between markers and disease. What is necessary, however, is that the association of a particular molecular lesion (detected as a nucleic acid, protein or any other biomarker) with cancer be reproducible in a statistically robust manner, can be confirmed in clinical studies within and between testing sites, and can be tested in clinical specimens before disease onset and compared with healthy controls (to assess marker levels before the appearance of cancer and to understand better how such levels relate to disease).

One rather interventionist solution to standardize biomarker validation in this manner would be to impose greater standardization of tests from above: in other words, to have the European Medicines Agency or US Food and Drug Administration step in and not only set technical standards but also describe 'Standard Operative Protocols' for the tests and determine which markers should be included. This approach would undoubtedly force test developers to generate sets of data that were more readily comparable. And, in relatively short order, the associations between clinical outcomes and marker presence might be established, albeit for a limited set of markers.

The main objection to this imposition is that it would be impractical—it would limit market freedom and require the withdrawal of many existing tests. The scheme also presupposes that regulators would know which markers are most relevant (something that noone knows at present).

An alternative approach to validate cancer-related markers would be to ensure that metadata sets can be generated retrospectively as a result of clinical experience and the operation of the free markets. At present, there is no obligation on clinicians or investigators who use the tests or on the companies who manufacture or license them to generate information other than for immediate medical uses. Even if individual data were retained, there would be no way of generating a sensible metastudy by combining data.

This situation could be greatly improved, however, if the possibility of a later metastudy were considered now. The bringing together of these disparate data could be achieved if users of tests were funded to maintain biopsies or other biological samples. Not every clinical sample tested need be retained—just a statistically relevant sample that might be, say, 5–10% of all cases. For the metastudy, disparate data could then be combined, and an additional standard test protocol could be conducted in a central laboratory on the biopsies of tissues samples. The standard test would act as a form of 'moderating' test—a way of helping align the information from the plethora of protocols used.

All very well—but who will pay for it? Not the manufacturers, and not the users—if either were to bear the cost, it would skew what is currently a rather free-form exploration of the cancer diagnostic space. No, the costs should be borne by the beneficiaries of the metastudies—health insurers or social healthcare systems. With an understanding of which markers are informative and which are essentially worthless, payors would then have a basis for deciding which tests they should pay for.