Driving forces in cancer diagnostics

By Todd R. Gander; Edward N. Brody, M.D., Ph.D.; Robert E. Mehler, M.D.; Joseph S. Heilig, Ph.D.; Britta S. Singer, Ph.D.; Larry Gold, Ph.D.

We are beginning to see the first signs of fundamental change in protein-based cancer diagnostics. The field has long been dominated by the search for highly predictive, single biomarkers that could reveal the state of the patient's health, much as hCG can indicate pregnancy or the detection of viral DNA or RNA can reveal the presence of certain pathogens. The difficulty of the search is complicated by the breadth and subtlety of the clinical questions that arise in the management of cancer patients: Does this asymptomatic patient have ovarian cancer? How aggressive is it? Are there distant metastases? What is the preferred treatment? Has the cancer recurred? Not surprisingly, the search for individual proteins with sufficient information content to answer these types of questions has generally been frustrating. Although immunohistochemical analysis of tumor proteins is common, and serum biomarkers are used in a number of cancers (Table 1), performance is not always adequate, and significant unmet needs remain, particularly in the areas of screening and post-treatment surveillance.

There is general agreement that earlier detection is preferable, since it is associated with improved survival (Figure 1). Unfortunately, early detection is the exception for most cancers. For example, patients diagnosed with localized nonsmall-cell lung cancer (NSCLC) are approximately 20 times more likely to survive five years than patients whose disease is diagnosed with distant metastases (50.5 percent survival vs. 2.6 percent). Fewer than 20 percent of NSCLC cases are diagnosed at the localized stage. Certainly, there is the potential for lead-time bias, and some cancers that are diagnosed at later stages are very aggressive — earlier detection will not always improve survival. Nevertheless, the magnitude of the survival benefit associated with earlier-stage diagnosis is striking and presents an extraordinary opportunity. Recognition of the potential impact of early detection has led to the establishment of the NCI's Early Detection Research Network (EDRN) and a proposal for the appropriate development processes for early-detection biomarkers. The current debate over screening for breast and prostate cancers is really about the quality of the available screening methodologies, rather than the merits of screening per se.

For the patient who has been treated for cancer, the need for improved surveillance tools is profound. While tracking PSA is highly informative for the patient who has undergone prostatectomy (prostate cells should not be present in such a patient), the tools are quite limited for other cancers. For women who have been treated for breast cancer, recommended surveillance and follow-up procedures include periodic histories and physical exams; mammography; and pelvic exams (for women taking tamoxifen who have not undergone hysterectomy). Routine measurement of biomarkers is not recommended. Not surprisingly, patients and physicians turn to tools without proven value in the desperate hope of detecting recurrences or metastases earlier. One study found that approximately 62 percent of surveillance costs for breast cancer survivors were attributable to “excess testing” that exceeded recommendations and was unlikely to be effective. A separate study found that only 22.6 percent of recurrences were detected at scheduled follow-up surveillance, and that symptoms were the primary indicator of relapse for 57.6 percent of cases.
post-translational modifications (e.g., initiation and termination, alternative mRNA splicing and transcriptional initiation and termination, alternative translational that result from molecular events, such as alternative tran-
glycosylation, sulfation), the number of distinct potential pro-
rations in the case of breast and ovarian cancers). Beyond these and other emerging applications, genomics is creating the potential for new diagnostic approaches by providing the “parts list” of proteins that might be usefully measured, as well as an ever-increasing sense (through gene expression studies) of proteins that might (or might not) go up or down with disease15,16 (see Table 1, p.14).

Despite the value of genetic analyses and gene expression studies, in many clinical settings they face substantial limitations that protein-based tools can address. Particularly for screening and surveillance, gene expression analyses present the challenge of sample acquisition, since mRNA must be obtained from the cells of interest. Routine ovarian biopsies certainly would not be acceptable and would be unlikely to yield material from a comparatively small tumor. The problem is the same for post-treatment surveillance for recurrence. Beyond a few exceptions such as colorectal cancer17 and hematological malignancies, the challenge of sample acquisition can be a significant barrier to effective screening and surveillance. Moreover, the measurement of proteins should be preferred in many cases, since these molecules are generally the effectors of physiology and pathophysiology and since the correlation between mRNA levels and the concentrations of the corresponding proteins is far from perfect.18,19 In addition, tumor-associated necrosis and possibly apoptosis can be detected through protein analyses, but not genetic analyses. This is the concept behind Matritech’s (Newton, MA) use of nuclear matrix proteins as markers for the detection of cancer, including an FDA-approved test for bladder cancer.20 Finally, to the extent that many genetic analyses reflect disease risk, they cannot detect the actual onset of disease.

The genomic “parts list” will provide the foundation for recombinant protein expression and purification, and the subsequent development of specific “capture agents” such as antibodies, antibody mimics, affibodies22 and aptamers.23,24 These capture agents will allow highly sensitive and specific measurements that, when combined in multiplexed formats, will provide unprecedented information and serve as the basis for an entire generation of new tests.
Driver No. 2: Multiplexed protein measurements

Tumors are accompanied by myriad subtle changes at the molecular level. The clinician, patient and laboratorian need tools for examining a large number of potentially relevant proteins with sufficient resolution to detect even minor changes reproducibly. Today, tools such as two-dimensional polyacrylamide gel electrophoresis with mass spectroscopy analysis (2-D PAGE-MS) meet the throughput needs embodied in the former criterion, while quantitative tools in routine laboratory use, such as ELISAs, meet the latter. Multiplexed protein measurements, defined here as the simultaneous quantitative measurement of 10 or more analytes (ultimately, hundreds or thousands), will combine the best qualities of both of these technologies.

Early proteomics work using tools such as 2-D PAGE-MS has demonstrated the potential of multiplexed measurements for distinguishing relevant health states. In this approach, two-dimensional gel electrophoresis is used to separate proteins on the basis of size and charge. Features that are differentially present in two samples of interest can then be annotated as individual proteins using mass spectroscopy information. Typically, in analyses of approximately 2,000 features, 50-300 might be identified that are uniquely or differentially expressed. A significant number of studies based on 2-D PAGE have investigated the proteomes of different cancers. Some studies have used laser capture microdissection (LCM) to enrich disease-specific material and enhance the probability of discovering biomarkers. Limitations of 2-D PAGE include poor sensitivity for the analysis of low- and medium-abundance proteins; few researchers have used the technique for quantitative analyses. Alternative approaches for separating proteins prior to mass spectrometric analysis include single-dimensional and multidimensional high-pressure liquid chromatography (HPLC) and surface-enhanced laser desorption and ionization (SELDI) mass spectroscopy. SELDI has been used to identify novel biomarkers and signatures for ovarian cancer, and prostate cancer. In these cases, the protein biomarkers are essentially differential peaks observed on mass spectromgrams.

In the clinical laboratory, antibody-based assays — such as ELISAs — set the standard for protein analysis. They can provide both low limits of detection and high specificity in distinguishing closely related proteins. In order to develop a more complete picture of cancer at the molecular level, we want the performance of ELISAs, but in multiplexed formats. Low limits of detection are particularly important if tumors are to be detected when they are relatively small. The system must be scalable, since the accuracy of diagnostic tests — and the breadth of the testing menu — will improve as more markers are included in each analysis. The number of companies (Table 2) and academics developing multiplexed systems for protein analysis has grown enormously. Some products are already available, and compelling papers are appearing. Almost all work has utilized antibodies as capture agents, relying on either fluorescent analytes or secondary antibodies for detection.

Photoaptamers are the basis for a different approach that may enable superior limits of quantitation and scalability.
Because photoaptamers do not require the use of a second (detection) capture agent, they avoid the challenges of identifying secondary antibodies and screening antibody pairs for compatibility. More importantly for scaling, they obviate the reported challenges of maintaining large numbers of secondary reagents in solution and the potential for cross talk. (If a protein is labeled with a secondary antibody, it will generate a signal whether it has bound to its intended primary antibody or interacted nonspecifically with another antibody.) Finally, because photoaptamers are covalently bound to their target analytes before signal detection, vigorous washing can be used to remove background proteins, yielding superior signal-to-noise ratios and lower limits of quantitation.

We believe that analyte density is the key to the coming change of paradigm, though this multiplexing must be based on the quality of the individual analyte assays, particularly in terms of reproducibility and the limits of quantitation. Of course, if underlying cancer biology is more like pregnancy and infectious disease, and a single marker can address each clinical question, then small-scale multiplexing will suffice — 20 multiplexed ELISAs could answer 20 distinct clinical questions. Given the complexity of the disease and its management, this seems unlikely. If we need more information regarding tumor staging, potential drug responsiveness, metastases, likely aggressiveness of the tumor, etc., we will need very large multiplexed arrays just for cancer, without even considering diagnostic tools for other major diseases.

**Driver No. 3: Bioinformatics**

The new, bioinformatics-driven approach to diagnostics will rely on the integration of multiple biomarkers, which individually have only modest information content. We can state the obvious without recourse to mathematics: The use of multiple independent markers can yield diagnostics with markedly improved performance. An individual with an elevated PSA level and an enlarged, irregular prostate is more likely to have prostate cancer than a patient with only one of these signs, and is much more likely to have prostate cancer than a patient with neither of these signs. Similar thinking underlies the suggestion that measurement of the serum tumor antigen CA-125, when combined with transvaginal ultrasonography, can be used to screen for ovarian cancer in high-risk women. The coming change in protein-based diagnostics will extend the current paradigm and rely on tens of proteins to reach any given clinical conclusion, increasing the statistical power substantially. When one considers the number of possible states of the patient, it becomes clear that hundreds of proteins are likely to be involved in the complete work-up of a case.

Conceptually, the development of multiplexed assays is straightforward. Using multiplexed protein measurements, we can rapidly compare any two populations of interest (e.g., individuals with and without cancer; aggressive and indolent cancers, therapy responders and nonresponders) across a large number of protein parameters and identify the best predictors of the clinical state of interest. These protein measurements can be combined into protein signatures, validated, and used to assess the unknown patient. Although validation of the biomarkers and signatures is critical, a full understanding of the underlying biology is not. As Qu, et al. have observed (in the context of SELDI), improved knowledge about the markers does not, in itself, enhance the performance of a signature.

Of course, such analyses need not be limited exclusively to protein measurements. They can explicitly consider other variables, such as genetic risk factors, age, and clinical signs and symptoms. Protein-based diagnostics thus becomes the driver for an approach that provides new tools for integrating the available data, rather than relying solely on the pattern-recognition skills of the individual practitioner. This approach will place the laboratorian in a more central role in managing the cancer patient, alongside the physician.

Evidence for the potential of multiparameter signatures is emerging rapidly. An algorithm combining five serum tumor markers for ovarian cancer achieved substantially better performance than can be achieved with CA-125 alone. Using only protein analysis, SELDI-based work has led to multivariable signatures with improved performance for ovarian and prostate cancers. Urocor Labs (Oklahoma City, OK) offers its UroScore as a multivariable tool (Gleason score and pattern, number of positive cores, tumor location, average percent tumor involvement, PSA and patient age) for predicting organ confinement in prostate cancer. Hematopathology

---

**Table 1: CCN recommendations on the use of tumor markers in common malignances**

<table>
<thead>
<tr>
<th>SITE</th>
<th>ANALYTES</th>
<th>SCREENING</th>
<th>DIAGNOSIS AND PROGNOSIS</th>
<th>TREATMENT SELECTION AND MONITORING</th>
<th>SURVEILLANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute myelogenous leukemia</td>
<td>Immunophenotyping markers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bladder</td>
<td>NM P-22, BTA, M 344</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast</td>
<td>Her-2/neu</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colorectal</td>
<td>CEA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Esophagus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSCLC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ovarian</td>
<td>CA-125</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pancreas</td>
<td>CA 19-9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prostate</td>
<td>PSA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCLC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stomach</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Recommended in NCCN practice guidelines Marker in development (no current consensus on utility) or used in investigational settings*
uses flow cytometry to examine potentially large numbers of cell surface and intracellular antigens and has become a critical technique in the diagnosis, prognosis, management and surveillance of leukemias and lymphomas.66-69

Physicians and patients will certainly want clear reporting of the integrated data, as well as clinically relevant conclusions. Analyses will likely include summary statements of some kind: “with a 95 percent likelihood, the patient has breast carcinoma, with a 93 percent likelihood she has no distant metastases, and with a 90 percent likelihood she has no lymph node involvement.” Such a report would give the physician and the patient a sound basis upon which to discuss treatment strategies. UroScore patient reports indicate the probability associated with each possible patient state: organ-confined prostate cancer, capsular penetration, or advanced disease.65

**Early applications**

While applications will emerge across the continuum of care that extends from detection through diagnosis, prognosis, therapeutic choice, monitoring and surveillance, perhaps the greatest unmet needs lie at the ends of this continuum.

Early applications of protein signatures will probably focus on relatively tractable problems. We are likely to see surveillance of individuals at high risk of disease before screening tests are introduced. Applications might include monitoring for recurrence after treatment or surveillance of individuals predisposed to disease (e.g., women with BRCA mutations for ovarian cancer or smokers for lung cancer). Based on large, statistically powerful data sets, general population screening will follow when the accuracy of the diagnostic call for apparently healthy people reaches a high enough level to be useful for diseases of a given prevalence. Ultimately, screening will be demanded by the ultimate beneficiaries of our healthcare system — the patients.

The recent work by Petricoin, et al. provides a practical illustration of how tests might evolve. The serum protein signature that they used provided sensitivity and specificity of 100 percent and 95 percent, respectively, for distinguishing ovarian cancer from benign ovarian disease. For the sample set, the positive-predictive value (PPV) was 94 percent versus 35 percent for CA 125, based on testing of the same samples. The authors noted, however, that the test did not provide acceptable performance for population-based screening,60 and it has been suggested that the PPV in such a context would be less than 1 percent.64 Continued research to identify markers and the associated patterns may ultimately lead to tests with PPVs that are appropriate for use with the general population. Petricoin and colleagues have already reported a second-

*Continues on page 16*
generation ovarian cancer signature with sensitivity and specificity of 97 percent and 100 percent, respectively.\(^4\) (99 percent specificity would yield a 3.7 percent PPV for the general population while 99.9 percent specificity would yield a 28 percent PPV, based on the researchers’ prevalence estimates.) Moving from research into the clinic, multiplexed arrays using specific capture agents represent an alternative approach for measuring these same markers, and it has been suggested that such assays would be the preferred approach if SELDI profiling is not developed as a clinical tool.\(^4\)

The near future

Although today’s clinicians are overwhelmed by data, a lack of relevant information continues to be a hallmark of cancer management. There are few reliable tests for assessing whether the asymptomatic patient might have cancer; prognosis remains difficult, and treatment choices are complicated; and finally, patients must live with the dread of recurrence.

New assays, enabled by genomics, multiplexed protein measurements and bioinformatics will change this situation dramatically and are already having an impact. As new applications become available, laboratorians will answer the questions that are most relevant to clinicians and their patients. Patient management will continue to improve, and early detection will allow us to use our most effective tools — particularly surgery — when their impact is the greatest. Longer term, suppressive therapies may be indicated. For patients, there will be greater confidence in the selection of treatment options and less anxiety as the quality of clinical information improves.

**Note:** In this article, the term “diagnostics” is used broadly to include screening, diagnosis, prognosis, therapy selection, and surveillance based on clinical testing.

---

**Todd R. Gander** is a consultant and former senior director of strategic planning at Somalogic, Inc., in Boulder, CO. Dr. Edward N. Brody is the chief medical officer at Somalogic. Dr. Robert E. Mehler and Dr. Britta S. Singer are members of Somalogic’s medical group. Dr. Joseph S. Heilig is senior director of science at Somalogic. Dr. Larry Gold, Somalogic’s chairman and CEO, has been an academic scientist and the founder of several biotech enterprises, and is a member of the National Academy of Sciences.

**References**


---

**Table 2: Selected companies commercializing microarrays for protein measurements**

<table>
<thead>
<tr>
<th>COMPANY</th>
<th>TECHNOLOGY FOCUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affibody (Stockholm, Sweden)</td>
<td>Affibody capture agents</td>
</tr>
<tr>
<td>Beckman-Coulter (Fullerton, CA)</td>
<td>Multiplexed arrays printed in individual wells of 96-well microtiter plates</td>
</tr>
<tr>
<td>Becton-Dickinson Clontech Ab</td>
<td>Two-color (nonquantitative) system for comparing quantities of prelabeled protein from two samples (378 analytes)</td>
</tr>
<tr>
<td>Micorarrays (Franklin, Lakes, NJ)</td>
<td>Surface plasmon resonance detection</td>
</tr>
<tr>
<td>Biacore (Uppsala, Sweden)</td>
<td>Bead-based multiplexing of ELISA-format assays; technology licensed from Luminex</td>
</tr>
<tr>
<td>Bio-Rad Laboratories (Hercules, CA)</td>
<td>Arrays derived from current point-of-care technology</td>
</tr>
<tr>
<td>Biosite (San Diego, CA)</td>
<td>Antibody development</td>
</tr>
<tr>
<td>Dyaq (Cambridge, MA)</td>
<td>Surface plasmon resonance detection</td>
</tr>
<tr>
<td>HTS (Hopkinton, MA)</td>
<td>Bead-based multiplexing of ELISA-format assays (instruments and bead sets); multiplexed kits available through multiple vendors</td>
</tr>
<tr>
<td>Luminex (Austin, TX)</td>
<td>Flow-through chips</td>
</tr>
<tr>
<td>Molecular Staging (New Haven, CT)</td>
<td>Polyclonal antibody arrays</td>
</tr>
<tr>
<td>Packard BioChip Technologies (Billerica, MA)</td>
<td>Detection based on rolling-circle amplification technology (RCAT)</td>
</tr>
<tr>
<td>Phylos (Lexington, MA)</td>
<td>Hydrogel coated slides for protein arrays</td>
</tr>
<tr>
<td>Pierce Endogen (Woburn, MA)</td>
<td>Antibody mimics based on PROfusion technology</td>
</tr>
<tr>
<td>Randox (Crumlin, UK)</td>
<td>Multiplexed ELISA arrays printed in individual wells of 96-well microtiter plates</td>
</tr>
<tr>
<td>Somalogic (Boulder, CO)</td>
<td>ELISA arrays for clinical diagnostic applications</td>
</tr>
<tr>
<td>SurroM ed (M Mountain View, CA)</td>
<td>Aptamer and photoaptemer capture agents</td>
</tr>
<tr>
<td>Zyomyx (Hayward, CA)</td>
<td>Bead-based multiplexing of ELISA-format assays and multiparameter flow cytometry</td>
</tr>
</tbody>
</table>


