WELLCOME IMAGES

The lack of standardization in the collection and storage of medical specimens (pictured) can hinder subsequent research.

Bring on the biomarkers

The dismal patchwork of fragmented research on disease-associated biomarkers should be replaced by a coordinated 'big science' approach, argues **George Poste**.

If researchers could establish correlations between diseases and changes in biomarkers, the ability of physicians to diagnose disease and tailor treatments to individuals would be radically improved[1]. However, research into biomarkers — disease-associated molecular changes in body tissues and fluids — hasn't yet delivered on its promise. Technologies such as proteomics and DNA microarrays have contributed a voluminous literature of more than 150,000 papers documenting thousands of claimed biomarkers, but fewer than 100 have been validated for routine clinical practice. This dismal record reflects the failure of researchers to embrace a coordinated systems-based approach.

Many chronic diseases involve changes in multiple molecular pathways[2], but validating associations between diseases and large sets of biomarkers is hugely challenging. Most biomarker discovery is conducted in academic labs, which tend to lack the resources and multidisciplinary expertise needed to establish robust correlations between biomarkers and people's health status or responses to treatment. Also, getting candidate biomarkers into large-scale validation studies poses substantial logistical and regulatory challenges. Overcoming them requires the integration of diverse skills.

Changes are needed to standardize methods and obtain the large sample sizes necessary for validation trials. The traditional model of investigator-initiated research must be replaced with the collaborative approaches typical of 'big-science' projects — such as The Cancer Genome Atlas initiative of the US National Institutes of Health (NIH) to catalogue the genomic changes found in cancers. Biomarker discovery should be a component of large research networks, involving industry and experts in molecular biology, genetics, analytical chemistry, computation, engineering, clinical-trial design, epidemiology, statistics, regulation and health-care economics.

"Too many researchers rely on whatever specimens they can obtain conveniently."

Biomarkers have long been hailed as the key to better patient care and lower medical costs[3]. The American Society of Clinical Oncology, for example, estimates that routinely testing people with colon cancer for mutations in the *K-RAS* oncogene would save at least US$600 million a year[4]. It would also spare patients futile and potentially toxic treatments — for example, people with these mutations don't respond to drugs that inhibit epidermal growth factor receptors, which cost up to $100,000 per treatment.

APPLES AND ORANGES

A major impediment to progress in the hunt for biomarkers is the lack of standardization in how specimens are collected. Unless specimens are taken from people who are matched for as many variables as possible, all the subsequent steps in efforts to correlate biomarkers with people's conditions and responses to treatments are compromised. At the very least, people should be matched for sex, age, weight, ethnicity, lifestyle factors such as smoking and alcohol use, and the previous treatments they've received. In practice, differences in how patients are

assessed, and in how those assessments are reported by different physicians in different medical centres and in different countries, mean that this rarely happens.

Too many researchers rely on whatever specimens they can obtain conveniently from local institutions. These might be fixed tissues from pathology laboratories, the usefulness of which is often limited by inadequate medical records or a lack of donor consent. Much biomarker research is carried out on cultured cell lines as surrogates for patient tissue samples. These studies can be valuable, but only if researchers demonstrate that the properties of their cell lines match those of the equivalent cell type in the human body.

The standardization problem also applies to how specimens are handled and stored, which can dramatically affect the levels of biomarkers detected. Such details are rarely documented. In a 2009 NIH survey, researchers from 80% of more than 700 responding laboratories said they struggled to obtain standardized specimens for biomarker research. Alarmingly, a similar percentage did not question how specimen quality or handling conditions might affect their results[5].

Another pervasive problem is insufficient sampling. Most biomarker studies examine fewer than 100 specimens and lack the statistical power needed to demonstrate a robust association between multiple biomarkers and a particular condition[6]. Depending on how many biomarkers are profiled, hundreds, or even thousands, of matched control and disease samples may be needed to satisfy regulatory requirements and demonstrate that a test confers clinical or economic benefits.

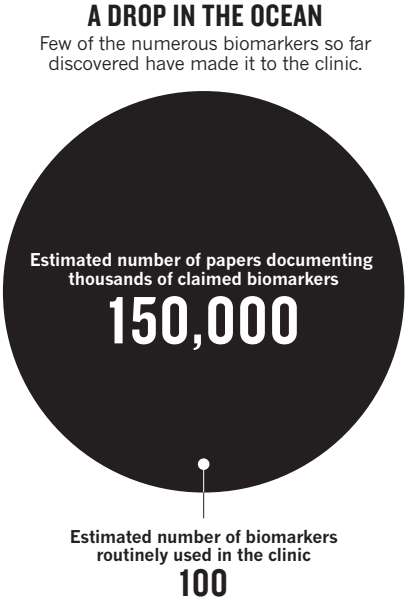
BEST PRACTICE

The recent shift to testing for multiple biomarkers (multiplex profiling) has stimulated impressive innovation in high-throughput technologies for automated parallel profiling of genes, proteins, RNA molecules and metabolites. But such advances will mean little unless organizational and funding reforms persuade the research community to adopt common standards and a cross-disciplinary, systems-based approach to biomarker discovery and validation.

Several initiatives are attempting to improve the practice of tissue storage or biobanking[7]. For example, the US National Cancer Institute's Cancer Human Biobank (caHUB) has established stringent guidelines to ensure that samples from healthy individuals and cancer patients are collected, annotated, stored and analysed under standardized conditions and accompanied by appropriate donor medical information. Such initiatives offer a 'best practice' model for biobanks in general. There is also a need for the creation of international biobanks, which would require harmonized rules to allow researchers worldwide to access the resources, and not be hindered by incompatible standards or uncoordinated regulatory barriers.

Most important, funding agencies, such as the NIH, the European Framework Programme, the Innovative Medicines Initiative (also in Europe) and the UK Wellcome Trust, should demand better returns for their investments in biomarker research. They should support only research programmes that: have access to a sufficient number of stringently characterized specimens; impose rigorous quality control in specimen acquisition, handling and storage; and possess the full spectrum of cross-disciplinary capabilities needed to translate laboratory findings to the clinic. This would encourage academic laboratories to become part of larger research networks that include clinical and industrial partners.

The large sample sizes required to validate multiplex tests are already generating massive data sets. As whole-genome sequencing becomes commonplace, the data sets will only get bigger[8]. Analysing them will require the sophisticated mathematical, statistical and computational skills of specialists in informatics. For such large-scale projects, data must also be documented, stored and analysed in standardized ways. To get candidate biomarkers into the clinic will require



making the data formats used in academic research compatible with the formats used by industry and regulators, and with the software used to produce electronic patient health records. Industrial and clinical partners can be a powerful force in driving such standardization.

As candidate biomarkers progress to expensive validation trials, industrial partners will be essential. They can provide financial support and expertise in large-scale assay production, clinical-trial design, data analysis and regulatory compliance. For complex multiplex tests, it is crucial that regulation experts are involved in the design of biomarker validation trials, to advise on what regulatory agencies are likely to demand. The hurdles that several biotechnology companies have encountered recently after marketing genetic tests for disease risk without first setting standards for their tests indicate the peril of not doing this[9].

The validating process will also need to involve experts in health-care economics. In an increasingly cost-conscious health-care environment, regulatory approval will not be a guarantee of clinical adoption. Researchers must prove that their diagnostic tests will change clinical practice and reduce costs, either by improving people's health or by eliminating ineffective, expensive treatments[10].

The obstacles to progress in biomarker research reside as much in the culture and organization of academic research as in deficiencies in analytical technologies. To become clinically useful, biomarker research must operate more like the large, collaborative networks mobilized for international genome-wide association studies and the multi-institution, multi-investigator big-science projects conducted in physics, climate change and complex-systems modelling.

The ability of biomarkers to improve treatment and reduce health-care costs is potentially greater than in any other area of current medical research. Thousands of papers have been written, but too few clinically useful biomarkers have been produced given a global public investment of several hundred million dollars over the past decade. It is time to restructure the field to focus instead on delivering tangible advances in health care. ■

George Poste *is chief scientist at the Complex Adaptive Systems Initiative, and Regents professor of health innovation at Arizona State University, Scottsdale, Arizona 85257, USA.*
e-mail: george.poste@asu.edu

1. Auffray, C., Chen, Z. & Hood, L. *Genome Med.* **1,** 2 (2009).
2. Shadt, E. E. *Nature* **461,** 218–223 (2009).
3. Davis, J. C. *et al. Nature Rev. Drug Discov.* **8,** 279–286 (2009).
4. go.nature.com/ulasjk
5. Vaught, J. *Research & Policy Initiatives in NCI's Office of Biorepositories & Biospecimen Research* (Office of Biorepositories & Biospecimen Research, 2009); available at go.nature.com/61yq9m
6. Ransohoff, D. F. & Gourlay, M. L. *J. Clin. Oncol.* **28,** 698–704 (2010).
7. Moore, H. M. *et al. Cancer Res.* **69,** 6770–6772 (2009).
8. Biesecker, L. G. *et al. Genome Res.* **19,** 1665–1674 (2009).
9. Javitt, G. *Nature* **466,** 817–818 (2010).
10. Meckley, L. M. & Neumann, P. J. *Health Policy* **94,** 91–100 (2010).

The author has declared a conflict of interest, which is available online at go.nature.com/mften7