PART I

Foundations of Biomarker Research

1

CHAPTER 1

Nuts and Bolts of Biomarker Research

Sharmistha Ghosh[1] *and Sudhir Srivastava*[2]

[1]Cancer Biomarkers Research Group, Division of Cancer Prevention, National Cancer Institute, Rockville, MD, USA

[2]Cancer Biomarkers Research Group, Division of Cancer Prevention, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA

What is a biomarker?

A biomarker is a particular characteristic, or a molecular fingerprint, which indicates manifestation of a physiological state, and which can be objectively quantified to distinguish a normal state from a pathological condition (e.g., cancer) or a response to a therapeutic intervention. The National Cancer Institute (NCI) of the National Institutes of Health (NIH) defines biomarker as: "A biological molecule found in blood, other body fluids, or tissues that is a sign of a normal or abnormal process, or of a condition or disease. A biomarker may be used to see how well the body responds to a treatment for a disease or condition. Biomarkers are also called molecular markers and signature molecules." [1]

As a normal cell undergoes a complex process of transformation into a cancerous state, it is hoped that measurable characteristics can be analyzed to derive a meaningful clinical decision – either directly, from the early-stage tumor before it is palpable or detectable by sensitive screening technologies available at this time, or as a result of an immunological response to the tumor. The characteristics include a broad range of biochemical entities such as nucleic acids (e.g., DNA, mRNA, long and small [short] non-coding RNA), proteins, post-translationally modified proteins (e.g., phosphoproteins, glycoproteins, methylated proteins, glycolipids), sugars, lipids and small metabolites, as well as whole circulating cells or biophysical characteristics of tissues.

Failure to detect an identifiable molecular marker may not be a negative predictor of malignancy, and a positive test for a molecular marker may not always be a positive predictor of malignancy. However, an ideal biomarker should indicate a reliable positive or negative correlation with the presence of the disease, which means that the clinical test for the biomarker should have high sensitivity (true positive rate – that is, the ability to correctly identify individuals with the disease) and specificity (true negative rate – that is, the ability to correctly identify individuals without the disease). The clinical value of a biomarker test is based on its positive predictive value (PPV), or how likely it is for test-positive individuals to actually have the disease, and its negative predictive value (NPV), or how likely it is for test-negative individuals to not have the disease. These again depend on the prevalence of the disease in the population of interest. Biomarkers also need to be easily accessible (e.g., by non-invasive methods for screening purposes), quantifiable, analyzable, and interpretable.

Biomarkers in Cancer Screening and Early Detection, First Edition. Edited by Sudhir Srivastava.

Why biomarker research is imperative

The development of cancer is preceded by numerous germline and somatic mutations, structural changes in chromosomes, and other genetic and epigenetic changes, which transform normal cells into benign tumors and, progressively, into malignant and metastatic forms. Cancer is a heterogeneous, multigenic group of diseases; the heterogeneity lies not only at the biochemical level (genes, proteins, metabolites), but also at the tissue and population level (e.g., [2–10]). The enormous complexity makes cancer detection, diagnosis, and treatment quite challenging. Although cancers diagnosed at earlier stages have a much better prognosis compared with cancers diagnosed at later stages, it is noteworthy that many cancer patients are diagnosed at a stage at which the cancer is too far advanced to be cured.

Currently, recommendations for early detection of cancer in average-risk individuals are available for colorectal, cervical, breast, endometrial (in menopausal women) and prostate cancers, and in high-risk individuals in the case of lung cancer. There has been a substantial increase in "cancer" incidence as a result of screening, but without a proportional decrease in mortality despite treatment. This implies that screening identifies a large reservoir of indolent cancers (overdiagnosis) [11], which would have never become symptomatic without screening, and did not require any treatment. However, because it is not known at this time which lesions are indolent, many individuals are put through intensive treatments unnecessarily, which often causes anxiety as well as substantial physical and financial harm. An extensive discussion on existing screening modalities, recommendations, and the consequences and complexities involved is beyond the scope of this chapter. A shared decision-making discussion between the patients and their physicians, based on existing data, and also taking into consideration an individual patient's values and philosophies on healthcare, is important [12].

The ability to identify tumors that are destined to progress, and which are associated with morbidity and mortality at an early stage, will allow effective treatment interventions and reduce deaths. Identification of tumor-specific molecular signatures is imperative for a new approach to early detection, diagnosis, prognosis, disease classification and risk prediction. It will also help to implement appropriate treatment decisions and therapeutic interventions, to monitor treatment response and efficacy (i.e., a measurable effect on a clinical end point), and to overcome drug resistance in a precise, patient-specific approach. Such practice of tailored "personalized medicine", based on the molecular portraits of tumor cells, allows physicians to inform individual patients of the expected outcomes – for example, whether treatments or surveillance approaches will be beneficial, and when to stop treatment based on response to drug(s). An illustration of several windows of clinical relevance in the management of cancer during its course of development is shown where different biomarker profiles can be applied to each of these windows for optimal management of cancer (Figure 1.1).

A few biomarkers discussed in this section underscore their utilization as clinical tools for facilitating diagnosis and treatment of tumors. Germline mutations in the high-penetrance genes breast cancer 1 (*BRCA1*) and breast cancer 2 (*BRCA2*), which are associated with hereditary susceptibility to breast and ovarian cancers, somatic mutations in phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha (*PIK3CA*) gene in colorectal tumors which can act as predictive biomarker for adjuvant aspirin therapy, and metastatic melanoma patients who harbor v-raf murine sarcoma viral oncogene homolog B $(BRAF)^{V600E}$ mutations in tumors and are treated with the BRAF inhibitor vemurafenib are some well-known examples [14–16]. Another example is a translocation occurring between the breakpoint cluster region (*BCR*) gene on chromosome 22 and the Abelson (*ABL*) tyrosine kinase gene on chromosome 9 in chronic myelogenous leukemia (CML), where the fusion product BCR-ABL is implicated in disease pathogenesis. Imatinib, a drug that effectively treats CML, was developed against the BCR-ABL fusion product.

Protein-based biomarkers, such as overexpression of human epidermal growth factor receptor 2 (HER2; also known as ERBB2) in breast tumors, serve as a marker for prognosis of breast cancer, as well as an effective target for treatment with

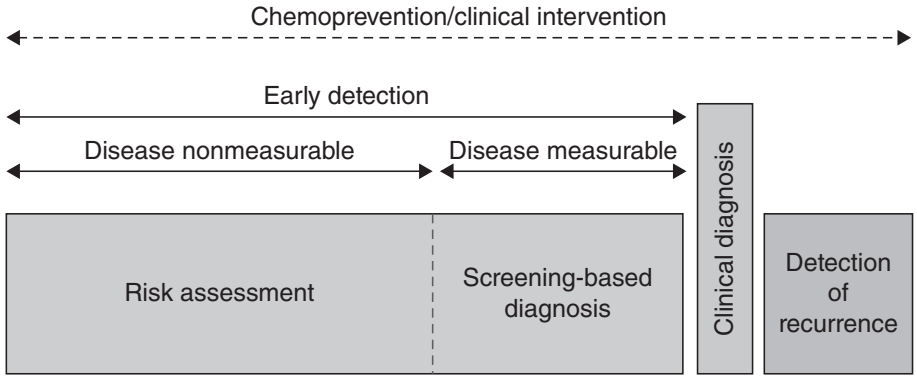


Figure 1.1 Biomarker application in the clinic. A long window of opportunity for chemoprevention or any clinical intervention is divided into sub-windows, based on whether a risk assessment is made when the disease is non-measureable, or an early diagnosis is made based on screening for measurable characteristics of the tumor, or a clinical diagnosis is made when the disease is symptomatic or has recurred. Adapted from [13]. For a color version of this figure please see color plate section.

trastuzumab, a HER2-specific monoclonal antibody [17,19]. Levels of cancer antigen 25 (CA 125) in serum can be indicative of disease progression and treatment response in ovarian cancer, but this is not definitive because of high false-positive rate [20,21]. A recently discovered biomarker, fibulin-3, in plasma and effusions of mesothelioma patients, may be promising for early detection, diagnosis and prognosis of pleural mesothelioma, if validated in well-designed prospective studies [22]. The prostate-specific antigen (PSA) is widely used as a prostate cancer screening marker, although its reliability as a screening tool is controversial. After reviewing all evidence, the US Preventive Services Task Force recommends against PSA testing for screening purposes (Grade D Recommendation since 2011), because the estimated harms outweigh benefits [23, 24]. However, PSA remains a good biomarker for the monitoring and management of patients with advanced prostate cancer.

Biomarker discovery can also start in the clinic, following a reverse course to the bench. Recently, the genome sequencing of bladder tumor of an exceptional responder to everolimus treatment showed that a loss-of-function mutation in tuberous sclerosis complex 1 (*TSC1*) had high correlation with everolimus sensitivity [25]. This demonstrates that unconventional individual patient information can be extracted to discover biomarkers of drug sensitivity, which will help the identification of specific subsets of patients who would respond to particular treatments, or provide a novel insight into the molecular mechanisms. Therefore, molecular characterization of tumors is the key to early detection, diagnosis, prognosis and development of effective treatments.

Existing screening techniques are incapable of distinguishing benign and indolent cancers from aggressive ones, and even the histopathological criteria are insufficient for this. Biomarkers, in conjunction with existing screening and imaging techniques, can also become very important diagnostic tools. Despite the widely available colonoscopy, a diagnostic screening method that can detect precancerous polyps and early-stage colon cancer, only about 40% of newly diagnosed colon cancers are localized. This is primarily because of non-compliance to colonoscopy, and missed cancers due to technical or other reasons. Among asymptomatic patients, the rate of missed cancers in the hands of experienced operators range from 2–6%; the highest being on the right side of the colon [26,27].

Non-invasive biomarker tests may have a higher probability of population-wide compliance, and can help reduce cancer burden and improve clinical outcomes by identifying individuals at moderate or high risk who must have colonoscopy. Recently, a multi-target stool DNA testing with significantly higher sensitivity (detects more cancers in average-risk, asymptomatic individuals), but lower specificity (namely, more false positive results) compared with the well-known fecal immunochemical

test, has been reported. However, whether this new test has a role in colorectal cancer screening is beyond the scope of this study [28].

This chapter aims to provide a comprehensive outline of a systematic approach to biomarker development designed to cope effectively with the US regulatory system, under which the products are brought to the market, and also to provide an insight into the various available tools that support the discovery and development of biomarkers. Subsequent chapters of this book will focus on biomarkers for early detection, diagnosis, prognosis and risk assessment (excluding genome-wide association studies) of cancer.

A systematic approach to developing clinically useful biomarkers: important tools and infrastructure to address the challenges

Successful translation of promising biomarkers from the bench to the clinic has been relatively rare [29]. Although a few molecular biomarkers have been approved by the US Food and Drug Administration (FDA) for various clinical purposes [30], none is suitable for population screening. Most biomarkers do not progress beyond the discovery stage, and it is important to understand "why" there is this disappointingly slow pace before one can think of developing effective strategies for biomarker development.

Historically, biomarker discovery efforts have been piecemeal, silo-oriented, unorganized, and lacking a systematic approach. Biomarker discovery and validation are also attributed to a host of technological and methodological problems. To name a few, there has been a lack of adequate numbers of trained personnel to collect and process biospecimens, a lack of well-annotated samples, a lack of standardized protocols or quality control, a lack of blinding of researchers and randomization of animals or patient samples, and a lack of supporting tools. All these problems are further compounded by technological limitations of detecting low-abundance signals in limited amounts of biospecimens. This has produced a patchwork of standards, a lack of reproducible data and, often, conflicts and confusion. Numerous publications, albeit published in reputable journals with rigorous peer review standards, exemplify how problems with study design, statistical deficiencies (such as overfitting of data), and lack of validation (analytic validity or clinical validity of a biomarker test) can lead to misinterpretation of the data and wrongful conclusions [31]. These are all significant impediments in the development of clinically useful biomarkers.

NCI's Early Detection Research Network and the five-phase schema

Given the challenges and costs involved in developing and validating biomarkers, it is difficult, or even impossible, for a single institution or agency to undertake the work single-handedly. Validated biomarkers, however, may prove useful to many stakeholders. Therefore, for a concerted effort to accelerate the development of systematic, evidence-based discovery of cancer biomarkers, which is a highly complex process in itself, it is beneficial to recognize the power of large, well-planned organizational structures (consortia), where technological and intellectual resources are shared and integrated towards a common goal. By leveraging the strengths of multidisciplinary, multisite partners, and by sharing costs, consortia are more likely to take on challenges that individual stakeholders cannot meet [32–36]. The team environment of consortia also has the benefits of better quality control, effective validation of results, effective utilization of scientific and financial resources, ans can develop the best standardized practices, help with troubleshooting problems, and draw inspiration from team members.

The Early Detection Research Network (EDRN) of the NCI (http://edrn.nci.gov/), launched in 2000, is the first comprehensive network created to scrupulously discover and validate biomarkers for early detection of cancer. Since its inception, the EDRN has made significant progress in developing a dynamic organized infrastructure for identifying candidate biomarkers, accommodating rapidly evolving technologies, and conducting multicenter validation studies and building resources, while also fostering public-private partnerships (PPPs) with industries, and developing collaborations with other government agencies and designated cancer centers.

Table 1.1 FDA approved diagnostic tests.

Biomarker	*Clinical utility*	*Year of approval*	*EDRN principal investigator/ industrial partner*
%[-2]proPSA	Reduce the number of unnecessary initial biopsies during prostate cancer screening. Also, appears to be highly associated with increased risk of aggressive disease.	2012	D. Chan/Beckman Coulter
PCA3 (in urine)	Biopsy or re-biopsy decisions in patients at risk for prostate cancer.	2012	J. Wei/Gen-Probe
OVA1^{TM} (5 analytes in blood: CA 125, prealbumin, apolipoprotein A-1, beta2 microglobulin, and transferrin)	Prediction of ovarian cancer risk in women with pelvic mass.	2010	D. Chan/ Vermillion
Risk of ovarian malignancy (ROMA) (algorithm with CA125 and HE4 blood tests for pelvic mass malignancies)	Prediction of ovarian cancer risk in women with pelvic mass.	2011	S. Skates/Fujirebio Diagnostics
Combined panel of blood DCP (approved by FDA in 2007) and AFP-L3 (approved by FDA in 2005)	Risk assessment for development of hepatocellular carcinoma.	2011	J. Marrero/Wako Diagnostics

The valuable resources developed by EDRN that are worth mentioning include development of common data elements (CDEs), standard operating procedures (SOPs), diagnostic assays that are in use in the community, ten standard biospecimen reference sets, and a strong bioinformatics base, in collaboration with the Jet Propulsion Laboratory (JPL) of the National Aeronautics and Space Administration (NASA). The infrastructure gives researchers with promising biomarkers a platform to assess them accurately for translating discovery into clinical use. A total of five EDRN-developed or -supported biomarker-based diagnostic tests have been approved by the FDA (Table 1.1), and many diagnostic tests are currently in use in (CLIA) laboratories (Table 1.2). The EDRN Biomarker Database (BMDB) tracks all of EDRN's research progress, as well as related entities such as protocols and publications. Relevant information is available from the National Cancer Institute at http://edrn.nci.nih.gov/.

All studies follow the recommended five-phase schema [37] and the Prospective Sample Collection Retrospective Evaluation (PRoBE) guidelines [38], which address the methodological and biostatistical challenges that were not considered in the past, and evaluate the step-by-step evidence necessary to allow a thorough assessment of promising new biomarkers for diagnostic, screening, and prognostic purposes, and validate their applications in clinical settings before proclaiming use as clinical tools. The phases of biomarker development, which are analogous to the drug development process, are as follows (Table 1.3):

Phase 1: Preclinical Exploratory. The first step in biomarker development often begins with exploratory preclinical studies, which aim to identify unique characteristics or molecular signatures of tumor tissues, compared with normal tissues to develop biomarkers, with great discriminatory ability (i.e., correctly distinguish cancer from non-cancer).

Phase 2: Clinical Assay Development and Validation. Because tissues should not be procured using invasive mechanisms for screening purposes, the idea is to conduct biomarker assays using non-invasively obtained samples, such as blood, urine, saliva and so on. Clinical assays that can distinguish subjects with cancer from those without cancer are developed, optimized and validated in this phase, using non-invasively obtained biosamples. However, it is

Table 1.2 CLIA certified diagnostic tests.

Biomarker assay	*Clinical utility*	*EDRN principal investigator/CLIA laboratory*
MiPS (Mi prostate score urine test), Multiplex analysis of T2-ERG gene fusion, PCA3 and serum PSA	Detection of prostate cancer	A. Chinnaiyan/Gen-Probe
IHC and FISH for T2-ERG fusion	Detection of prostate cancer	A. Chinnaiyan/Roche
GSTP1 methylation	Repeat biopsies in prostate cancer	D. Sidransky/ OncoMethylome
Mitochondrial deletion	Detection of prostate cancer	National Institute of Standards and Technology (NIST)/Mitomics
Proteomic panel	Detection of lung cancer	W. Rom/Celera
Aptamer-based markers	Detection of lung cancer	W. Rom/Somalogic
80-gene panel	Detection of lung cancer	A. Spira/Allegro
Vimentin methylation in stool	Detection of colon cancer	S. Markowitz/Labcorp
Galectin-3 ligand	Detection of advanced adenomas and colon cancer	R. Bresalier/BG Medicine
GP73	Risk of hepatocellular carcinoma	T. Block/Beckman Coulter
8-gene panel for Barrett's esophagus	Progression prediction	Stephen Meltzer/ Diagnovus

not known at this stage whether the biomarker can be used for early detection. EDRN's standard biospecimen reference sets, which use common sets of biospecimens from well-characterized and matched cases and controls, have been carefully developed and annotated to overcome many of the logistic and design-related issues in preliminary and advanced biomarker validation.

Phase 3: Retrospective Longitudinal Repository Studies. In this phase, the ability of the biomarker to detect preclinical disease and its promise as a screening tool for early detection (how long before a patient's clinical diagnosis the biomarker could be detected in the biospecimen) are evaluated by analyzing samples from case patients before their clinical diagnosis.

Phase 4: Prospective Screening Studies. If a biomarker shows promise as a screening tool in Phase 3, a prospective screening study is conducted, where the screen is applied to a relevant population to determine the operating characteristics of the screening test. Screen-positive individuals go through diagnostic procedures to determine the stage or nature of the cancer. A small-scale assessment of costs and survival benefits associated with screening is also done.

Phase 5: Cancer Control Studies. This phase evaluates whether screening reduces the burden of

Table 1.3 The five phases of biomarker development.

Phase	*Purpose*
Phase 1: Preclinical exploratory	Promising directions identified. It is established that the biomarker is able to distinguish between cancer cases and control subjects.
Phase 2: Clinical assay and validation	Clinical assay detects established disease.
Phase 3: Retrospective longitudinal	Biomarker detects disease early before it becomes clinical and a "screen positive" rule is defined.
Phase 4: Prospective screening	Extent and characteristics of disease detected by the test and the false referral rate are identified.
Phase 5: Cancer control	Impact of screening on reducing the burden of disease on the population is quantified.

Adapted from [37].

cancer on the population, and if there is a net benefit. Even if the biomarker detects disease early, it may not have an overall benefit for the screened population because of:

a ineffective treatments;
b poor compliance or difficulties with implementing the screening program in community practice;
c economic or morbidity-associated costs of screening itself and of the diagnostic work-up of false-positive individuals; and/or
d overdiagnosis of cancers, which would not have been detected without screening and may have caused no harm or death.

The four key components of the PRoBE study design relate to clinical context and outcomes, criteria for measuring biomarker performance, the biomarker test itself, and the study size (Pepe et al. 2008).The design also has greater implications (e.g., use in biomarker discovery, creating valuable biorepositories), as indicated by the authors.

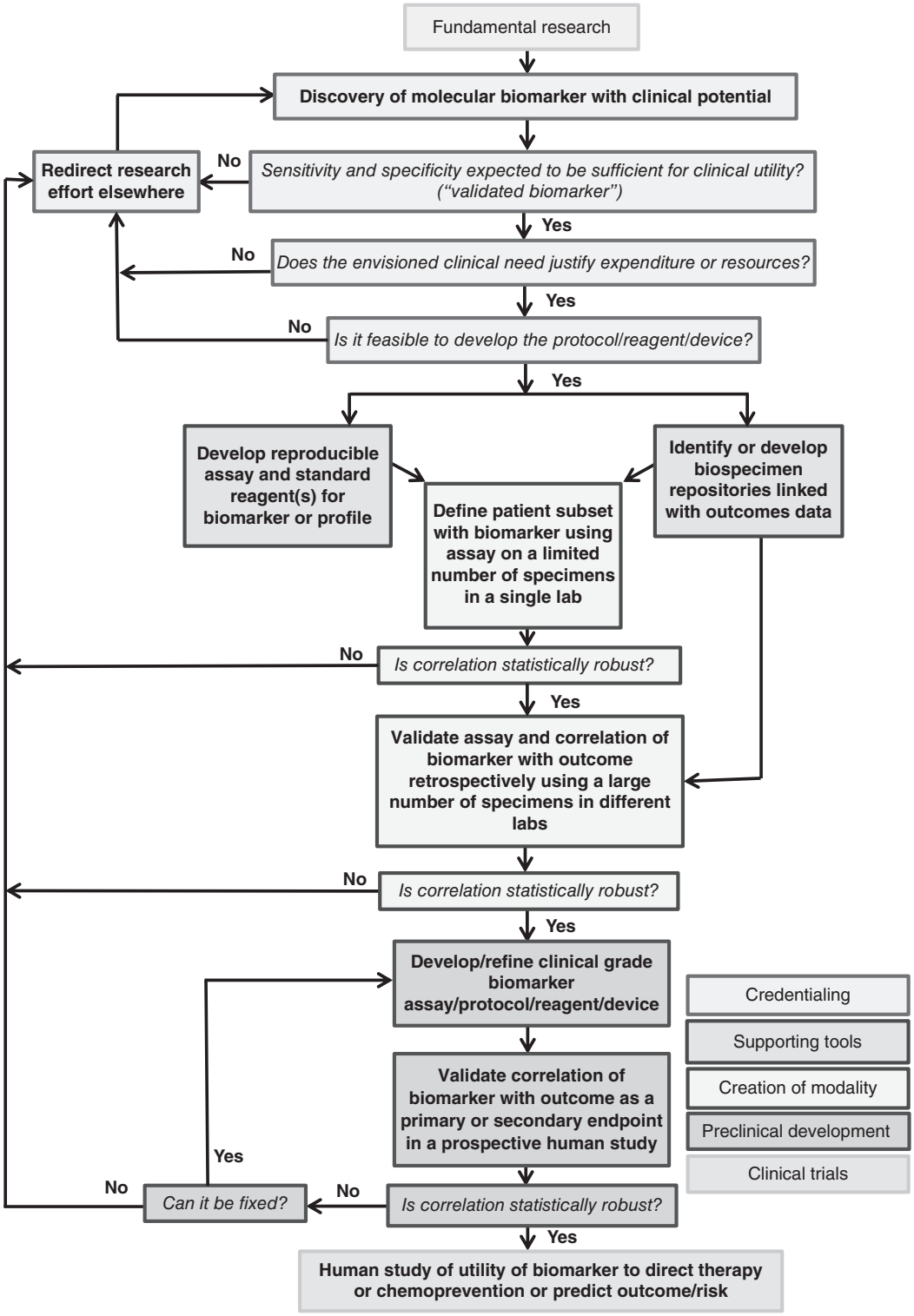
The biospecimen-based assessment modality pathway

Based on the evidentiary framework proposed in the two seminal studies mentioned earlier [37, 38] and a systematic approach of the translational biomarker research taken by EDRN, the Translational Research Working Group (TRWG) of the NCI drafted a biospecimen-based assessment modality pathway (BM Pathway) [39], which sketches out the necessary elements in biomarker development, and also provides a framework for understanding key scientific and regulatory challenges, and guidance on how to meet the challenges and facilitate biomarker development with a programmatic and operational perspective. Such a tool is geared towards maximum effectiveness and efficiency of translational research activity. The five main phases of the BM pathway (Figure 1.2), which also overlaps significantly with phases 2 through 4 of the schema described earlier [37], are as follows:

1 *Credentialing:* The exploratory data need to be scrutinized and prioritized, to evaluate how well the key questions on clinical validity, clinical need, practicality of assay development, and so on, are addressed.
2 *Supporting tools:* Reproducible assays, standard reagents, biorepositories are some important tools. The biospecimens used for biomarker research need to be properly preserved, well annotated and clinically relevant.
3 *Creation of modality:* An assay conducted on a limited number of biospecimens to define the patient subset needs to be further validated, and correlation with clinical outcomes needs to be established using a large number of retrospectively collected specimens.
4 *Preclinical development:* The development of tests underscores the importance of systematically cataloged information on a wide range of biomarkers, even the ones that individually do not show a robust association with clinical phenomena. Clinically important tests often include a group of genes instead of individual biomarkers (e.g., Oncotype DX Breast Cancer Assay that helps treatment decisions). A prospective study will determine the performance characteristics of the test, and establish a statistically significant correlation with predetermined endpoint(s).
5 *Clinical trials:* It is important to conduct randomized controlled clinical trials before implementation of the biomarker test in clinical settings.

In addition, interactions between the BM pathway and developmental pathway for new targeted therapeutic agents where a diagnostic target can also be a therapeutic target need to be considered [39]. Although this is, theoretically, a cost-effective process, it involves logistical complications and coordination challenges, because these are often developed by different research entities with different organizational structures and business policies, such as academia and industry. However, a merger of the two pathways seems inevitable, with the advent of high-throughput 'omics' technologies, such as genomics, proteomics and metabolomics, where the generated "big data", coupled with sophisticated computational and bioinformatics methods, provides an insight into the underlying biological functions and processes [40–43]. Currently, the trend is shifting from a single biomarker to a more systematic pathway, and network-based panel of biomarkers for both diagnostic and therapeutic use.

The Cancer Genome Atlas (TCGA) (https://cancergenome.nih.gov/abouttcga), launched in 2005, uses the latest sequencing and analysis methods to characterize mutational landscapes


Fundamental research
Discovery of molecular biomarker with clinical potential
Redirect research effort elsewhere
No
Sensitivity and specificity expected to be sufficient for clinical utility? ("validated biomarker")
Yes
No
Does the envisioned clinical need justify expenditure or resources?
Yes
No
Is it feasible to develop the protocol/reagent/device?
Yes
Develop reproducible assay and standard reagent(s) for biomarker or profile
Identify or develop biospecimen repositories linked with outcomes data
Define patient subset with biomarker using assay on a limited number of specimens in a single lab
No
Is correlation statistically robust?
Yes
Validate assay and correlation of biomarker with outcome retrospectively using a large number of specimens in different labs
No
Is correlation statistically robust?
Yes
Develop/refine clinical grade biomarker assay/protocol/reagent/device
Validate correlation of biomarker with outcome as a primary or secondary endpoint in a prospective human study
Yes
No
Can it be fixed?
No
Is correlation statistically robust?
Yes
Human study of utility of biomarker to direct therapy or chemoprevention or predict outcome/risk
Credentialing
Supporting tools
Creation of modality
Preclinical development
Clinical trials

in various tumor types, and to identify molecular abnormalities that influence pathophysiology, affect outcome and constitute therapeutic targets. In a recently published article by the TCGA Pan-Cancer Group, the authors write that "Given the rate at which TCGA and International Cancer Genome Consortium projects are generating genomic data, there are reasonable chances of identifying the 'core' cancer genes and pathways and tumor-type-specific genes and pathways in the near term." [44]

Regulatory systems in the United States

To bring any diagnostic product to the market, the biomarker discovery and development process must meet the requirements set by the regulatory agencies in the United States. The Centers for Disease Control and Prevention, in partnership with the Centers for Medicare and Medicaid Services (CMS), support the Clinical Laboratory Improvement Amendments of 1988 (CLIA) to ensure that Federal standards are applied to all US clinical laboratories that test human specimens for health assessment, or to diagnose, prevent, or treat disease. The FDA regulates all medical devices, including *in vitro* diagnostic products. All procedures in nonclinical studies used to support a product that is submitted to the FDA for approval must follow Good Laboratory Practices (US FDA Bioresearch Monitoring Good Laboratory Practice: http://www.fda.gov/ICECI/EnforcementActions/BioresearchMonitoring/ucm133789.htm).

The regulations are subject to modification and evolution in the light of new trends in medical research. With the growing popularity of omics-based tests, which are highly complex in themselves, and considering some recent publicized cases of unreliable or falsified tests that may have caused harm to patients [45, 46], Federal regulators have focused on implementing rigorous standards for such tests, in order to prevent premature omics assays from entering the market and guide treatment decisions. In a recent article [47], the NCI, in collaboration with investigators representing areas of expertise relevant to omics-based test development, has defined an omics-based test as "an assay composed of or derived from multiple molecular measurements and interpreted by a fully specified computational model to produce a clinically actionable result." The FDA needs a more dynamic process for updating the regulatory policies, to keep pace with the newly developed clinical tests based on rapidly evolving modern technologies.

The changing landscape of biomarker research

Biomarker research requires continued investment in quality control, in strong technological, statistical and bioinformatics infrastructure, and in developing biospecimen repositories for generating high-quality data that meet US regulatory requirements.

The increasing use of systems biology and bioinformatics as pivotal tools for analysis and interpretation of biological functions has become increasingly popular. Systems approaches encompassing large sets of components have unraveled a range of cellular functions and networks [48–52], so it is

Figure 1.2 Biospecimen-Based Assessment Modality (BM) Pathway. The BM pathway is depicted as a flowchart, a schematic process representation widely used in engineering. The origin of the process is at the top. Fonts in bold indicate activity steps. Fonts in italics indicate conditional tests or decision steps. Unidirectional arrows represent the direction of the activity sequence and the direction of transfer of supporting tools from their parallel development paths to the main path of modality development. For each activity or decision point, it is understood that there are many more variations that can occur, and that not all steps may occur in each instance. The pathway does not address the ways in which insights gained from late-stage clinical trials can influence the development process. Biospecimen-based assessment devices can be used for screening, early detection, diagnosis, prediction, prognosis, or response assessment. The pathways are conceived, not as comprehensive descriptions of the corresponding real-world processes, but as tools designed to serve specific purposes, including research program and project management, coordination of research efforts, and professional and lay education and communication [39]. For a color version of this figure please see color plate section.

conceivable that the availability of fully sequenced genomes and omics-based "big data" has brought mathematical and computational tool-based resolutions to cancer pathogenesis within reach. It certainly has the potential to comprehend extremely complex, nonlinear biological networks, and to provide a cogent and coherent understanding of the cellular functionalities. The Systems Biology Workbench (SWB; http://sbw.sourceforge.net/) is a useful collection of tools for simulation, analysis, and visualization of biochemical networks.

Although it may be game-changing to harness and use big data in the clinic, the existing electronic health record data systems [42] are not equipped to handle large volume of data, such as whole genome sequencing of patients, where each patient sequencing generates between 5–10 gigabytes of data. Therefore, ancillary systems supporting the use of omics-based research in the clinic need to be developed in parallel. Because of the gaining momentum of such information fusion in the field of biology, it will be also useful to bridge the gap between biologists and computer scientists, and to invest in training more young investigators to handle mathematical and computational tools.

Furthermore, there are economic considerations in the implementation and utilization of a new biomarker-based healthcare approach. Unraveling the pathogenic processes of cancer and tumor development using modern technological tools to meet the demands of "personalized medicine" results in increased complexities and, often, may result in more laboratory tests. So, what does this mean for the already escalating healthcare costs in the United States? In a microeconomic analysis of personalized medicine [53], the authors analyzed the cost-effectiveness of different types of tests from a payer perspective, and showed that, depending on the test, it may be cost-saving or cost-creating, or may even be cost-neutral per patient. Although the cost estimates are based on treatment costs in the United States, the results could be widely applicable. Also, guidelines need to be established and post-market surveillance strategies need to be enforced before such tests are implemented in the clinic in order to ensure that preferential use of "personalized" tests is not driven by any financial interests.

Last, but not least, a paradigm shift in funding and administrative approaches in academia nationwide is warranted, to show support for collaboration-driven research. Although the concept of team science is not new, and substantial investments have been made in this direction, it is still not a widely acceptable practice in biomedical research. Investigators often believe that working independently results in increased productivity and drives competitive research; however, their resistance to collaborations often also stems from the fact that building a large collaborative network requires major investment of time, could be difficult to organize, may involve geographical barriers, and may require additional finances.

In addition, collaborative efforts are likely deterred by the fact that the current system rewards individual accomplishments over collaborative efforts, in the form of awarded grants or tenure in academia. If collaborative efforts are rewarded by recognizing multi-author publications in grant applications, alongside individual work of the investigators, inculcating the significance of collaborations in young minds at the graduate level (e.g., encourage multi-advisor thesis), and revising promotion and tenure policies in support of team-based interdisciplinary research, it may encourage the investigators to invest their time and efforts in forming strong collaborative networks. It is worth mentioning that the NIH, and several institutions, have been taking strides in this direction by emphasizing the value of team science. However, significant progress remains to be achieved.

References

1. Biomarkers Definitions Working Group (2001). Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clinical Pharmacology & Therapeutics* **69**, 89–95.
2. Anderson K, Lutz C, van Delft F W, Bateman C M, Guo Y, Colman S M, Kempski H, Moorman A V, Titley I, Swansbury J, Kearney L, Enver T, Greaves M (2011). Genetic variegation of clonal architecture and propagating cells in leukaemia. *Nature* **469**, 356–61.
3. Chapman P B, Hauschild A, Robert C, Haanen J B, Ascierto P, Larkin J, Dummer R, Garbe C, Testori A, Maio M, Hogg D, Lorigan P, Lebbe C, Jouary T, Schadendorf D, Ribas A, O'Day S J, Sosman J A, Kirkwood J M, Eggermont A M, Dreno B, Nolop K, Li J, Nelson B, Hou J, Lee R J, Flaherty K T, McArthur G A, B-S Group (2011).

Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *New England Journal of Medicine* **364**, 2507–16.

4. Gerlinger M, Rowan A J, Horswell S, Larkin J, Endesfelder D, Gronroos E, Martinez P, Matthews N, Stewart A, Tarpey P, Varela I, Phillimore B, Begum S, McDonald N Q, Butler A, Jones D, Raine K, Latimer C, Santos C R, Nohadani M, Eklund A C, Spencer-Dene B, Clark G, Pickering L, Stamp G, Gore M, Szallasi Z, Downward J, Futreal P A, Swanton C (2012). Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *New England Journal of Medicine* **366**, 883–92.
5. Ljungman M, Lane D P (2004). Transcription – guarding the genome by sensing DNA damage. *Nature Reviews Cancer* **4**, 727–37.
6. Navin N, Kendall J, Troge J, Andrews P, Rodgers L, McIndoo J, Cook K, Stepansky A, Levy D, Esposito D, Muthuswamy L, Krasnitz A, McCombie W R, Hicks J, Wigler M (2011). Tumour evolution inferred by single-cell sequencing. *Nature* **472**, 90–4.
7. Shah S P, Morin R D, Khattra J, Prentice L, Pugh T, Burleigh A, Delaney A, Gelmon K, Guliany R, Senz J, Steidl C, Holt R A, Jones S, Sun M, Leung G, Moore R, Severson T, Taylor G A, Teschendorff A E, Tse K, Turashvili G, Varhol R, Warren R L, Watson P, Zhao Y, Caldas C, Huntsman D, Hirst M, Marra M A, Aparicio S (2009). Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature* **461**, 809–13.
8. Shibata D, Schaeffer J, Li Z H, Capella G, Perucho M (1993). Genetic heterogeneity of the c-K-ras locus in colorectal adenomas but not in adenocarcinomas. *Journal of the National Cancer Institute* **85**, 1058–63.
9. Sottoriva A, Spiteri I, Shibata D, Curtis C, Tavare S (2013). Single-molecule genomic data delineate patient-specific tumor profiles and cancer stem cell organization. *Cancer Research* **73**, 41–9.
10. Xu X, Hou Y, Yin X, Bao L, Tang A, Song L, Li F, Tsang S, Wu K, Wu H, He W, Zeng L, Xing M, Wu R, Jiang H, Liu X, Cao D, Guo G, Hu X, Gui Y, Li Z, Xie W, Sun X, Shi M, Cai Z, Wang B, Zhong M, Li J, Lu Z, Gu N, Zhang X, Goodman L, Bolund L, Wang J, Yang H, Kristiansen K, Dean M, Li Y, Wang J (2012). Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. *Cell* **148**, 886–95.
11. Esserman L J, Thompson I M Jr, Reid B (2013). Overdiagnosis and overtreatment in cancer: an opportunity for improvement. *JAMA* **310**, 797–8.
12. Elmore J G, Kramer B S (2014). Breast cancer screening: toward informed decisions. *JAMA* **311**, 1298–9.
13. Hassanein M, Rahman J S, Chaurand P, Massion P P (2011). Advances in proteomic strategies toward the early detection of lung cancer. *Proceedings of the American Thoracic Society* **8**, 183–8.
14. Davies H, Bignell G R, Cox C, Stephens P, Edkins S, Clegg S, Teague J, Woffendin H, Garnett M J, Bottomley W, Davis N, Dicks E, Ewing R, Floyd Y, Gray K, Hall S, Hawes R, Hughes J, Kosmidou V, Menzies A, Mould C, Parker A, Stevens C, Watt S, Hooper S, Wilson R, Jayatilake H, Gusterson B A, Cooper C, Shipley J, Hargrave D, Pritchard-Jones K, Maitland N, Chenevix-Trench G, Riggins G J, Bigner D D, Palmieri G, Cossu A, Flanagan A, Nicholson A, Ho J W, Leung S Y, Yuen S T, Weber B L, Seigler H F, Darrow, T L, Paterson H, Marais R, Marshall C J, Wooster R, Stratton M R, Futreal P A (2002). Mutations of the BRAF gene in human cancer. *Nature* **417**, 949–54.
15. Flaherty K T, Puzanov I, Kim K B, Ribas A, McArthur G A, Sosman J A, O'Dwyer P J, Lee R J, Grippo J F, Nolop K, Chapman P B (2010). Inhibition of mutated, activated BRAF in metastatic melanoma. *New England Journal of Medicine* **363**, 809–19.
16. Nazarian R, Shi H, Wang Q, Kong X, Koya R C, Lee H, Chen Z, Lee M K, Attar N, Sazegar H, Chodon T, Nelson S F, McArthur G, Sosman J A, Ribas A, Lo R S (2010). Melanomas acquire resistance to B-RAF(V600E) inhibition by RTK or N-RAS upregulation. *Nature* **468**, 973–7.
17. Baselga J (2006). Targeting tyrosine kinases in cancer: the second wave. *Science* **312**, 1175–8.
18. Druker B J (2004). Imatinib as a paradigm of targeted therapies. *Advances in Cancer Research* **91**, 1–30.
19. Yeon C H, Pegram M D (2005). Anti-erbB-2 antibody trastuzumab in the treatment of HER2-amplified breast cancer. *Investigational New Drugs* **23**, 391–409.
20. Kabawat S E, Bast R C Jr, Bhan A K, Welch W R, Knapp R C, Colvin R B (1983). Tissue distribution of a coelomic-epithelium-related antigen recognized by the monoclonal antibody OC125. *International Journal of Gynecologic Pathology* **2**, 275–85.
21. Spitzer M, Kaushal N, Benjamin F (1998). Maternal CA-125 levels in pregnancy and the puerperium. *Journal of Reproductive Medicine* **43**, 387–92.
22. Pass H I, Levin S M, Harbut M R, Melamed J, Chiriboga L, Donington J, Huflejt M, Carbone M, Chia D, Goodglick L, Goodman G E, Thornquist M D, Liu G, de Perrot M, Tsao M S, Goparaju C (2012). Fibulin-3 as a blood and effusion biomarker for pleural mesothelioma. *New England Journal of Medicine* **367**, 1417–27.
23. Marcus P M, Kramer B S (2012). Screening for Prostate Cancer with Prostate-Specific Antigen: What's the Evidence? *American Society of Clinical Oncology Educational Book* **32**, 96–100.
24. McNaughton-Collins M F, Barry M J (2011). One man at a time – resolving the PSA controversy. *New England Journal of Medicine* **365**, 1951–3.

25. Iyer G, Hanrahan A J, Milowsky M I, Al-Ahmadie H, Scott S N, Janakiraman M, Pirun M, Sander C, Socci N D, Ostrovnaya I, Viale A, Heguy A, Peng L, Chan T A, Bochner B, Bajorin D F, Berger M F, Taylor B S, Solit D B (2012). Genome sequencing identifies a basis for everolimus sensitivity. *Science* **338**, 221.
26. Bressler B, Paszat L F, Chen Z, Rothwell D M, Vinden C, Rabeneck L (2007). Rates of new or missed colorectal cancers after colonoscopy and their risk factors: a population-based analysis. *Gastroenterology* **132**, 96–102.
27. Rex D K, Rahmani E Y, Haseman J H, Lemmel G T, Kaster S, Buckley J S (1997). Relative sensitivity of colonoscopy and barium enema for detection of colorectal cancer in clinical practice. *Gastroenterology* **112**, 17–23.
28. Imperiale TF, Ransohoff DF, Itzkowitz SH, Levin TR, Lavin P, Lidgard GP, Ahlquist DA, Berger BM (2014). Multitarget Stool DNA Testing for Colorectal-Cancer Screening. *New England Journal of Medicine* **370**, 1287–97.
29. Nass S J, Moses H L (eds, 2007). *Cancer Biomarkers: The Promises and Challenges of Improving Detection and Treatment*. National Academies Press, Washington, DC.
30. Dunn B K, Wagner P D, Anderson D, Greenwald P (2010). Molecular markers for early detection. *Seminars in Oncology* **37**, 224–42.
31. Diamandis E P (2010). Cancer biomarkers: can we turn recent failures into success? *Journal of the National Cancer Institute* **102**, 1462–7.
32. Chin-Dusting J, Mizrahi J, Jennings G, Fitzgerald D (2005). Outlook: finding improved medicines: the role of academic-industrial collaboration. *Nature Reviews Drug Discovery* **4**, 891–7.
33. Croft SL (2005). Public-private partnership: from there to here. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **99** Suppl 1, S9–14.
34. Kettler H, White K, Jordan S (2003). *Valuing Industry Contributions to Public-Private Partnerships for Health Product Development*. The Initiative on Public-Private Partnerships for Health, Global Forum for Health Research, Geneva, Switzerland.
35. Nishtar S (2004). Public-private 'partnerships' in health – a global call to action. *Health Research Policy and Systems* **2**, 5.
36. Schwartz K, Vilquin J T (2003). Building the translational highway: toward new partnerships between academia and the private sector. *Nature Medicine* **9**, 493–5.
37. Pepe M S, Etzioni R, Feng Z, Potter J D, Thompson M L, Thornquist M, Winget M, Yasui Y (2001). Phases of biomarker development for early detection of cancer. *Journal of the National Cancer Institute* **93**, 1054–61.
38. Pepe M S, Feng Z, Janes H, Bossuyt P M, Potter J D (2008). Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: standards for study design. *Journal of the National Cancer Institute* **100**, 1432–8.
39. Srivastava S, Gray J W, Reid B J, Grad O, Greenwood A, Hawk E T, Translational Research Working Group (2008). Translational Research Working Group developmental pathway for biospecimen-based assessment modalities. *Clinical Cancer Research* **14**, 5672–7.
40. Hanash S (2004). Integrated global profiling of cancer. *Nature Reviews Cancer* **4**, 638–44.
41. Souchelnytskyi S (2005). Proteomics of TGF-beta signaling and its impact on breast cancer. *Expert Review of Proteomics* **2**, 925–35.
42. Sreekumar A, Poisson L M, Rajendiran T M, Khan A P, Cao Q, Yu J, Laxman B, Mehra R, Lonigro R J, Li Y, Nyati M K, Ahsan A, Kalyana-Sundaram S, Han B, Cao X, Byun J, Omenn G S, Ghosh D, Pennathur S, Alexander D C, Berger A, Shuster J R, Wei J T, Varambally S, Beecher C, Chinnaiyan A M (2009). Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. *Nature* **457**, 910–4.
43. Tainsky M A (2009). Genomic and proteomic biomarkers for cancer: A multitude of opportunities. *Biochimica Et Biophysica Acta – Reviews on Cancer* **1796**, 176–193.
44. Kandoth C, McLellan M D, Vandin F, Ye K, Niu B, Lu C, Xie M, Zhang Q, McMichael J F, Wyczalkowski M A, Leiserson M D, Miller C A, Welch J S, Walter M J, Wendl M C, Ley TJ, Wilson R K, Raphael B J, Ding L (2013). Mutational landscape and significance across 12 major cancer types. *Nature* **502**, 333–9.
45. Buchen L (2011). Cancer: Missing the mark. *Nature* **471**, 428–32.
46. Samuel Reich E (2011). Cancer trial errors revealed. *Nature* **469**, 139–40.
47. McShane L M, Cavenagh M M, Lively T G, Eberhard D A, Bigbee W L, Williams P M, Mesirov J P, Polley M Y, Kim K Y, Tricoli J V, Taylor J M, Shuman D J, Simon R M, Doroshow J H, Conley B A (2013). Criteria for the use of omics-based predictors in clinical trials. *Nature* **502**, 317–20.
48. Chen KC, Calzone L, Csikasz-Nagy A, Cross FR, Novak B, Tyson JJ (2004). Integrative analysis of cell cycle control in budding yeast. *Molecular Biology of the Cell* **15**, 3841–62.
49. Duarte N C, Herrgard M J, Palsson B O (2004). Reconstruction and validation of Saccharomyces cerevisiae iND750, a fully compartmentalized genome-scale metabolic model. *Genome Research* **14**, 1298–309.

50. Feist A M, Henry C S, Reed J L, Krummenacker M, Joyce A R, Karp P D, Broadbelt L J, Hatzimanikatis V, Palsson B O (2007). A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Molecular Systems Biology* **3**, 121.
51. Klipp E, B Nordlander, R Kruger, P Gennemark, S Hohmann (2005). Integrative model of the response of yeast to osmotic shock. *Nat Biotechnol* **23**, 975–82.
52. Tyson J J, Chen K, Novak B (2001). Network dynamics and cell physiology. *Nature Reviews Molecular Cell Biology* **2**, 908–16.
53. Davis J C, Furstenthal L, Desai A A, Norris T, Sutaria S, Fleming E, Ma P (2009). The microeconomics of personalized medicine: today's challenge and tomorrow's promise. *Nature Reviews Drug Discovery* **8**, 279–86.