BIOMARKERS IN CANCER STAGING, PROGNOSIS AND TREATMENT SELECTION

Joseph A. Ludwig[*‡] *and John N. Weinstein*[*]

Abstract | Advances in genomics, proteomics and molecular pathology have generated many candidate biomarkers with potential clinical value. Their use for cancer staging and personalization of therapy at the time of diagnosis could improve patient care. However, translation from bench to bedside outside of the research setting has proved more difficult than might have been expected. Understanding how and when biomarkers can be integrated into clinical care is crucial if we want to translate the promise into reality.

'Water, water, everywhere,
Nor any drop to drink'

S. Coleridge, Rime of the Ancient Mariner, 1798

The formal TNM staging system[1] (see Online links box), promulgated by the American Joint Committee on Cancer (AJCC), is based almost exclusively on the anatomical extent of disease, which is assessed using a combination of tumour size or depth (T), lymph node spread (N), and presence or absence of metastases (M). Since its inception in 1958, the TNM system has provided a standardized, anatomical basis for staging with several important functions. It provides a basis for prediction of survival, choice of initial treatment, stratification of patients in clinical trials, accurate communication among healthcare providers, and uniform reporting of outcomes. For most tumour types, disease burden and spread have been considered the most reliable predictors of survival and determinants of the type and intensity of therapy to be used. Less often, tumour grade, histological subtype or patient age has been added to TNM staging when the AJCC became convinced that such information would significantly improve the prediction of survival or response to therapy.

The anatomically based TNM staging system is, of course, most useful when local therapies (for example, surgical resection, radiofrequency ablation or regional radiotherapy) provide the only means of cure, as was the case for almost all cancers when the TNM system was established. At that time, the goals ascribed to staging could be accomplished with a single, ordinal parameter of TNM stage. It was prognostically better to be stage I than stage II, better to be stage II than stage III, and so on. The generic dictionary definition of 'stage' has the same implication: it can be defined as 'a degree of advance in a journey' or 'a period or step in a process'.

The anatomically based TNM staging system remains useful for the purposes listed above, but new factors are both complicating the situation and providing new opportunities when it comes to predicting survival and/or selection of therapy. First, individual molecular markers and patterns of markers are successfully subdividing traditional tumour classes into subsets that behave differently from each other. Second, chemotherapeutic and biological agents are more effective and more widely used than when TNM staging was introduced, especially in the adjuvant setting. Third, many new targeted agents such as imatinib (Glivec), gefitinib (Iressa), and cetuximab (Erbitux), as well as older agents (for example, tamoxifen), are effective only if their respective molecular markers are mutated or expressed at sufficient levels. In breast cancer, for example, oestrogen receptor (ER) and HER2/NEU (also known as ERBB2) status have implications for prognosis and therapy that are

[*]Genomics and Bioinformatics Group, Laboratory of Molecular Pharmacology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA.
[‡]Laboratory of Cell Biology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA.
Correspondence to J.N.W.
e-mail: weinstein@dtpax2.ncifcrf.gov
doi:10.1038/nrc1739
Published online
20 October 2005

Summary

- The TNM staging system (based on a combination of tumour size or depth (T), lymph node spread (N), and presence or absence of metastases (M)) provides a basis for prediction of survival, choice of initial treatment, stratification of patients in clinical trials, accurate communication among healthcare providers, and uniform reporting of the end result of cancer management.
- There is a dilemma in TNM staging: frequent revisions to include new biomarkers would undermine the value conferred by the stability and universality of TNM, but a static formulation of TNM risks falling behind the state of the art in diagnostic techniques, biological concepts and biomarkers.
- Biomarkers initially considered for cancer screening or risk assessment might also prove useful for cancer staging or grading.
- A biomarker for use in staging or grading need not be as specific as it must be for screening, early detection or risk assessment.
- As molecularly targeted cancer therapeutics become more common, assessing the intended target will more often be deemed necessary for prediction of clinical response, independent of TNM stage. Targeted therapies and their associated biomarkers will often 'co-evolve'.
- The ideal biomarker assay for staging should be sensitive, specific, cost-effective, fast, and robust against inter-operator and inter-institutional variability. It must also demonstrate clinical value beyond that of the other types of information that are already available at the time of diagnosis.
- Biomarker candidates must undergo clinical validation before receiving US Food and Drug Administration approval. For most candidate markers, that process is just beginning.
- Despite all of the potentially useful biomarkers — for example, those identified from microarray or mass spectrometry studies — almost none have been incorporated into formal TNM staging.

independent of TNM stage. ER positivity improves prognosis, whatever the stage, and it also makes the tumour a candidate for therapy with a targeted hormonal agent such as tamoxifen or an aromatase inhibitor. Until the advent of trastuzumab (Herceptin), which targets HER2/NEU, HER2/NEU positivity was considered simply as a negative prognostic indicator independent of TNM stage. Increasingly, we can expect to see a 'co-evolution' of biomarkers and their respective targeted therapies.

The examples of ER and HER2/NEU biomarkers raise an immediate question: what should be the relationship between formal TNM staging and newly emerging biomarkers or biomarker combinations? Should the markers be incorporated into the determination of stage or should they be thought of as supplemental features that are outside of the system but serve the same five functions as discussed above for TNM? In this article, we will adopt the broader perspective, thinking of biomarkers as serving the same purposes as TNM staging but not necessarily formally incorporated into that system. We will consider formal TNM stage as one aspect of a snapshot of the cancer's status, most often taken at the time of initial diagnosis but, in principle, taken at any time thereafter.

A second important question is: what can we expect biomarkers to add to traditional staging? An answer is implied in the truism that cancer stage provides only probabilities of the course of disease or outcome for any individual patient. There remains the mysterious heterogeneity of outcomes for patients with cancers of apparently equivalent type, stage and grade. Those differences in outcome may relate, in part, to stochastic events, such as the time at which a single cancer cell happens to undergo all of the steps necessary for successful metastasis, or they may relate to factors that can be reasonably well understood at a deterministic level. Even if molecular markers cannot eliminate the stochastic uncertainties and enable us to predict outcome definitively, they will almost certainly increase our accuracy at sub-classifying patients and their cancers. We can expect that biomarkers will help us towards more personalized medicine. Again, the example of ER status in breast cancer is paradigmatic.

Although the incorporation of biomarkers into TNM staging has been a subject of considerable discussion by both the TNM committees and the broader biomedical community, no formal consensus has been reached. We contend that uncertainties about the role of biomarkers contribute to a fundamental paradox: despite the many current and emerging markers available in the clinical research setting, few have been integrated into clinical practice[2]. In this review, we will address the current use of both US Food and Drug Administration (FDA)-approved and other biomarkers available for use in clinical oncology, with special emphasis on staging, grading, and selection of therapy at the time of diagnosis. We will then discuss the types of biomarkers that are emerging for use at clinical research centres. Finally, we will consider practical factors such as clinical validation, regulatory approval and economics that seem to be limiting the integration of biomarkers into clinical practice.

The biomarker paradox

Although fewer than 10% of cancers can be traced to Mendelian inheritance, the ability of malignant cells to proliferate and metastasize can be ascribed to genetic alterations. Those alterations typically lead to activation

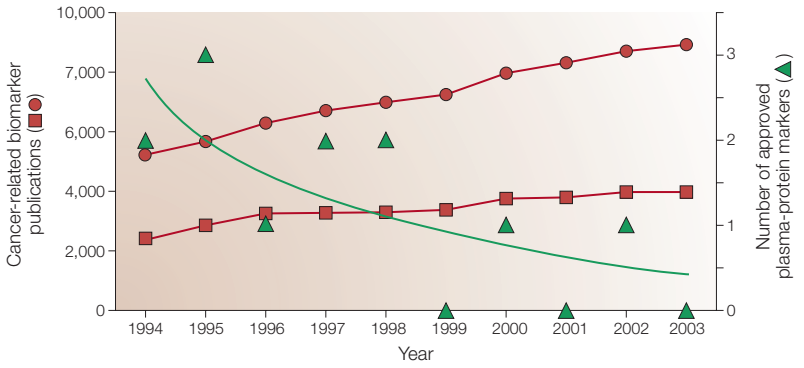

Figure 1 | **Numbers of publications on biomarkers and FDA approval of biomarkers.** Despite the increasing rates of publications on biomarkers, the number of US Food and Drug Administration (FDA)-approved plasma-protein tests is decreasing. Triangles and the associated trend line (green) represent the number of FDA-approved plasma-protein markers per year (data taken from REF. 5). Red squares and circles indicate publications under the Medline medical subject heading 'biomarker' and text word 'biomarker', respectively.

of proto-oncogenes, inactivation of tumour-suppressor genes and/or inactivation of DNA repair mechanisms. Although that genetic paradigm generally holds true, it does not account for the molecular complexity of cancers. It does not subsume epigenetic modulation of mRNA expression or differences in protein expression, post-translational modification, or function. Better appreciation of that complexity and recent advances in high-throughput technologies have provided a large inventory of candidate biomarkers with projected value for risk assessment, screening, diagnosis, prognosis, and selection and monitoring of therapy[3].

The DNA-based markers include SINGLE-NUCLEOTIDE POLYMORPHISMS (SNPs), chromosomal aberrations (such as the well-known BCR–ABL TRANSLOCATION), changes in DNA copy number, MICROSATELLITE INSTABILITY, and differential promoter-region methylation. The RNA-based biomarkers include overexpressed or underexpressed transcripts, and regulatory RNAs (for example, the microRNAs). The protein markers include cell-surface receptors such as CD20, tumour antigens such as prostate-specific antigen (PSA), phosphorylation states, carbohydrate determinants, and peptides released by tumours into serum, urine, sputum, nipple aspirates or other body fluids. Patterns of markers, particularly in serum, might prove more selective and potentially more useful than individual markers. Biomarkers at all of these levels are now, in principle, detectable by functional molecular imaging modalities based on magnetic resonance imaging (MRI), POSITRON EMISSION TOMOGRAPHY (PET) or optical imaging.

Paradoxically, however, fewer biological tests were approved by the FDA in 2003 than during any other year in the past decade[4]. As shown in FIG. 1, for example, there has been a decrease in serum-protein biomarker approvals despite steady increases in the literature on potentially useful ones[5]. Furthermore, almost none of the FDA-approved biomarkers (TABLE 1) are used in standard clinical practice, and only two of them have made it into the TNM staging guidelines. None of them were discovered through the new high-throughput genomic or proteomic technologies or from *in silico* analysis of databases. Unless listed in TABLE 1, the clinically used markers discussed in this review have been approved in the USA as 'analyte-specific reagents' (ASRs) for research purposes only. ASRs are generally not reimbursable by governmental or private health insurers, so few have entered standard clinical practice. In part, this 'thirst' in the midst of plenty is explained by the time necessary for any diagnostic or therapeutic innovation to make its way through the steps required for acceptance, but other factors, to be discussed later, are at work as well.

Biomarker use at diagnosis

Classification. Classification of a malignancy by tissue of origin is the first step towards predicting survival and choosing therapy. Because a tumour's anatomical location usually indicates its tissue of origin, molecular markers are rarely required. Histological examination generally confirms the diagnosis and identifies the tumour subtype. However, new molecular markers might sometimes be helpful in the differential diagnosis. For example, as shown schematically in FIG. 2, we recently used a combination of high-throughput RNA, protein and tissue microarray technologies to identify markers potentially useful for distinguishing colon and ovarian abdominal carcinomas from an unknown primary location[6]. Similarly, biomarkers have been reported to distinguish primary head and neck squamous cell carcinoma (HNSCC) from metastatic lung squamous cell carcinoma (SCC)[7], to determine the site of origin for HNSCC of unknown primary location[8] and to track genetic mutations that occur with the progression of that tumour[9].

Grade. Each anatomical site has its own histological grading system, designed to classify malignancies by degree of differentiation. Low-grade, well-differentiated tumours are usually less aggressive and more favourable in prognosis than high-grade tumours, which tend to grow faster and metastasize earlier. However, tumour grade is included in formal TNM staging only when intimately linked to prognosis, as it is for soft-tissue sarcomas, prostate cancer and primary brain malignancies. Assignment of grade is inherently subjective and dependent on the skill and experience of the reviewing pathologist, but several reports indicate that biomarker patterns can correctly score tumours according to their pathologist-assigned grades[10]. COMPUTER-AIDED DIAGNOSTIC SYSTEMS (CAD systems) have been approved by the FDA for preliminary grading of cervical smears (that is, Pap smears)[11] and for assisted interpretation of radiological images such as screening mammograms[12], computerized tomography (CT) scans[13] and standard X-ray films[14]. CADs are generally designed to make routine distinctions, giving the pathologist time to focus on difficult diagnostic problems. The acceptance of CADs has been accelerated by the fact that there had previously been extensive and rigorous standardization and quality control of the underlying imaging technologies (for example, mammograms and chest X-rays).

SINGLE-NUCLEOTIDE POLYMORPHISMS
Single-nucleotide changes in DNA that differ among individuals.

BCR–ABL TRANSLOCATION
Translocation between human chromosomes 9 and 22 t(9q34;22q11), resulting in an abnormal Philadephia chromosome that codes for a fusion protein causally linked to chronic myelogenous leukaemia.

MICROSATELLITE INSTABILITY
Genetic instability in diploid tumours owing to a high mutation rate, primarily in short nucleotide repeats. This phenotype is associated with defects in DNA mismatch-repair genes.

POSITRON-EMISSION TOMOGRAPHY
Imaging technique that detects nuclides as they decay by positron emission. The emitted positron collides with a free electron, resulting in the conversion of matter to two γ-rays, which emerge in opposite directions.

COMPUTER-AIDED DIAGNOSTIC SYSTEM
A computer algorithm for interpreting digital images or laboratory tests to provide a diagnosis.

Table 1 | **US Food and Drug Administration-approved cancer biomarkers**

Biomarker	Type	Source	Cancer type	Clinical use
α-Fetoprotein	Glycoprotein	Serum	Nonseminomatous testicular	Staging
Human chorionic gonadotropin-β	Glycoprotein	Serum	Testicular	Staging
CA19-9	Carbohydrate	Serum	Pancreatic	Monitoring
CA125	Glycoprotein	Serum	Ovarian	Monitoring
Pap smear	Cervical smear	Cervix	Cervical	Screening
CEA	Protein	Serum	Colon	Monitoring
Epidermal growth factor receptor	Protein	Colon	Colon	Selection of therapy
KIT	Protein (IHC)	Gastrointestinal tumour	GIST	Diagnosis and selection of therapy
Thyroglobulin	Protein	Serum	Thyroid	Monitoring
PSA (total)	Protein	Serum	Prostate	Screening and monitoring
PSA (complex)	Protein	Serum	Prostate	Screening and monitoring
PSA (free PSA %)	Protein	Serum	Prostate	Benign prostatic hyperplasia versus cancer diagnosis
CA15-3	Glycoprotein	Serum	Breast	Monitoring
CA27-29	Glycoprotein	Serum	Breast	Monitoring
Cytokeratins	Protein (IHC)	Breast tumour	Breast	Prognosis
Oestrogen receptor and progesterone receptor	Protein (IHC)	Breast tumour	Breast	Selection for hormonal therapy
HER2/NEU	Protein (IHC)	Breast tumour	Breast	Prognosis and selection of therapy
HER2/NEU	Protein	Serum	Breast	Monitoring
HER2/NEU	DNA (FISH)	Breast tumour	Breast	Prognosis and selection of therapy
Chromosomes 3, 7, 9 and 17	DNA (FISH)	Urine	Bladder	Screening and monitoring
NMP22	Protein	Urine	Bladder	Screening and monitoring
Fibrin/FDP	Protein	Urine	Bladder	Monitoring
BTA	Protein	Urine	Bladder	Monitoring
High molecular weight CEA and mucin	Protein (Immunofluorescence)	Urine	Bladder	Monitoring

BTA, bladder tumour-associated antigen; CA, cancer antigen; CEA, carcinoembryonic antigen; FDP, fibrin degradation protein; FISH, fluorescent *in-situ* hybridization; GIST, gastrointestinal stromal tumour; IHC, immunohistochemistry; NMP22, nuclear matrix protein 22; PSA, prostate-specific antigen.

The absence of analogous standardization of biomarker platforms is an important practical problem.

If the problems can be overcome, however, the addition of either individual or PATTERN-BASED BIOMARKERS in the assessment of histological grade could increase the utility of grading for predicting response to therapy. That would be a natural extension of current practice as pathologists already have considerable experience using at least a few markers (for example, ER, the progesterone receptor (PR) and HER2/NEU) in related contexts.

Stage. The AJCC, in collaboration with the TNM Committee of the International Union Against Cancer (UICC), has defined staging criteria for most anatomical sites[15]. T, N and M are determined separately and then grouped, usually to classify the cancer into one of four main stages (stages I–IV) and subdivisions thereof. Breast cancer staging, for example, distils 30 possible TNM combinations into 5 main prognostic stages[15]. Clinical staging, which is primarily used to guide initial therapy, integrates information from physical examination with data such as those from standard X-ray, CT, MRI, PET, endoscopic examination, biopsy, and surgical exploration. Pathological staging on the basis of surgical specimens, if acquired, complements clinical staging with a precise determination of the extent of disease and additional histological information.

Increasingly, imaging agents targeted at biomarkers are being used for anatomical localization. The most common are radioisotopes, detected by standard nuclear medicine imaging, by SINGLE-PHOTON EMISSION COMPUTED TOMOGRAPHY (SPECT) or by PET. Also under study are fluorescent molecules, which are detected by optical imaging, and paramagnetic particles for enhancing MRI. The target can be any marker that delineates the cancer or its metabolism. Some tumours (for example, carcinoid, phaeochromocytoma, and

PATTERN-BASED BIOMARKER
A biomarker constructed from a pattern of individual markers that, when evaluated together, can be used for risk assessment, screening, diagnosis, staging, selection of therapy and/or monitoring of therapy. The specific markers that make up the pattern may or may not have been identified.

SINGLE-PHOTON EMISSION COMPUTED TOMOGRAPHY
Imaging technology in which a photon detector array is rotated around the body to acquire data from many angles following the injection of a γ-emitting radionuclide.

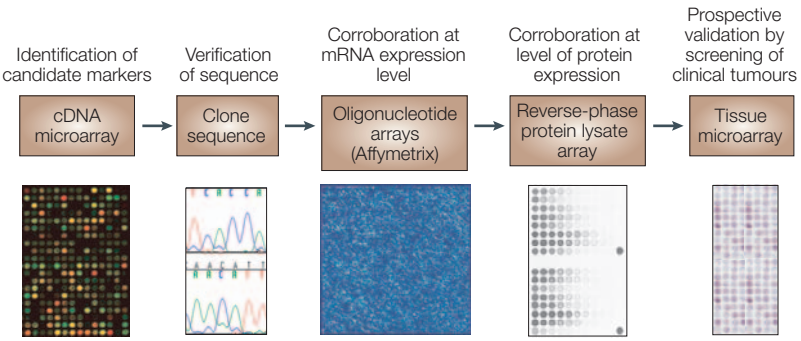

Figure 2 | **Use of multiple molecular technologies in combination to identify candidate biomarkers.** Candidate markers were initially identified from cDNA array data on the NCI-60 cancer cell line panel, then sequence-verified by re-sequencing of the clones and corroborated using Affymetrix oligonucleotide arrays. Reverse-phase proteomic arrays later showed that the selectivities of the candidate biomarkers held up at the protein level, and tissue arrays indicated the same selectivity at the level of clinical tumour specimens. Candidate biomarkers to distinguish between colon and ovarian cancers of an unknown primary location were identified and verified in this way. Modified from REF. 6.

cancers of the prostate, thyroid and colon) can be targeted by specific radiolabelled ligands. Carcinoid tumours, for example, are often localized using a radiolabelled analogue of octreotide (111-indium pentetreotide), which avidly binds to the somatostatin receptor, a protein commonly overexpressed in those tumours. Nuclear medicine-based imaging modalities are also clinically useful for evaluating tumour-related phenomena including angiogenesis[16], apoptosis[17,18], proliferation[19], metabolism[20,21], hypoxia[22] and drug resistance (such as P-glycoprotein function)[23]. Molecularly targeted functional imaging has enormous potential for staging, as it does for other aspects of cancer diagnosis and management, and it might also be easier than serum biomarkers to integrate into clinical TNM staging, given its anatomical basis.

For most solid tumours, the primary purpose of anatomy-based staging is to discern the probability of localized, as opposed to metastatic, disease. That crucial distinction, in considerable part, predicts survival and guides the choice of initial therapy. Anatomy-based staging, however, provides only part of the answer. A more precise picture can often be obtained by incorporating tumour grade and histological subtype. The role that biomarkers can play is exemplified by HER2/NEU, which is associated with an aggressive phenotype, decreased patient survival[24,25] and response to trastuzumab.

Acknowledging the potential importance of serum-derived biomarkers for staging, the AJCC incorporated the first such markers — serum α-fetoprotein (AFP), human chorionic gonadotropin-β (β-HCG, also known as CGB) and lactate dehydrogenase (LDH) for testicular cancer — into the TNM system. That necessarily cautious shift in the staging guidelines reflects the increased scientific evidence, at least for several cancer types, of more accurate prognosis with the addition of factors independent of anatomy and histological grade. However, as stated previously, most cancers are still exclusively staged by anatomic criteria, with a few exceptions (for example, sarcomas and malignancies of the thyroid, prostate, brain and testicle).

The AJCC sometimes informally recommends supplementation of TNM staging with information about tumour grade, histological subtype or relevant immunohistochemical (IHC) markers when they have prognostic or therapeutic value. Suggested supplementary parameters for breast cancer, for example, include those with proven value in predicting response to therapy (ER, PR and HER2/NEU receptor status). IHC or reverse transcription-PCR (RT-PCR) evaluation of sentinel lymph nodes is occasionally performed in clinical trials when microscopic examination is negative and, if obtained, is included as supplemental information in AJCC breast cancer staging. However, most supplementary markers have not been clinically validated, and the significance of lymph nodes that are negative by standard pathological staining with haematoxylin and eosin but positive by IHC or PCR remains unclear. The TNM system classifies nodes with cancer cell clumps less than 0.2 mm in diameter as node-negative, even when RT-PCR detects tumour cells. Such small clumps of cancer cells usually lack markers of proliferation and rarely induce a stromal reaction that indicates tumour implantation or growth[15,26].

Any updating of a system like TNM can have advantages but also imposes a price. Ideally, one would like to incorporate the latest supportable medical science, including new biomarkers. However, the relatively static, stable character of the system enhances its utility for stratifying patients in clinical trials, for communication among physicians and for standardizing the classification of tumours among institutions and nations. If the staging criteria were changing more often, it would be difficult, for example, to determine whether earlier clinical research was pertinent to current or future patients who are, or will be, staged differently. Furthermore, information supplemental to TNM is often incomplete or sporadically noted in medical records because it is not formally part of the staging process. Adoption of a formal 'augmented' staging system that included both a relatively static TNM component (updated occasionally) and a dynamic supplemental component (revised at the pace of scientific discovery) could perhaps resolve that paradox and allow new markers to be evaluated formally without undermining the value of anatomical staging.

Prognosis and treatment selection. Tumour classification, stage and sometimes grade are used to assess prognosis. However, as noted above, there would be a cost if formal cancer staging incorporated every other parameter able to improve prognosis. Furthermore, stratification in clinical trials using all possible TNM combinations would be impractical, given limitations in patient participation and resources. Addition of markers could similarly fragment the staging process, thereby limiting its utility. More information is generally better than less information, but the advantages must be weighed against those of a stable classification with relatively few categories.

Box 1 | **New applications for established biomarkers**

Biomarkers can play roles before cancer diagnosis (in risk assessment and screening), at diagnosis (as discussed in the main text) and after diagnosis (in monitoring therapy, selecting additional therapy and detecting recurrence) (FIG. 3). This review focuses on applications at the time of diagnosis, but markers that are currently considered for risk assessment or screening may also prove useful in cancer staging or prediction of response at the time of diagnosis. For example, the *BRCA1* (breast cancer 1) gene can be used in breast cancer, both for risk assessment and as a predictor of 10-year survival[118]. For patients with HIV/AIDS, the viral load and CD4-positive T-cell count predict the likelihood of acquiring an AIDS-related cancer and the probability that such a tumour will respond to highly active anti-retroviral therapy (HAART).

For risk assessment and screening, a marker must generally be inexpensive, highly specific and minimally invasive. Those requirements do not necessarily apply to markers for staging or grading. As noted in the text, markers considered infeasible for screening because they would yield too many false positives may still be useful after diagnosis. For example, serum CEA (carcinoembryonic antigen) is increased in colon, breast and lung cancer, but also in many benign conditions. It is, in that sense, non-specific, but increases in the context of known colon cancer strongly suggest recurrence of that malignancy. Similarly, although AFP (α-fetoprotein) and β-HCG (human chorionic gonadotropin-β) can be increased for many reasons, their reliability in assessing testicular cancer burden following diagnosis accounts for their integration into staging. PSA (prostate-specific antigen), cancer antigen (CA) 125, CA19-9, and other, similar markers[119–124] may also prove useful in similar contexts, but they have not been integrated into TNM staging because their expression often fails to correlate with tumour burden.

Biomarker expression often supplants or complements tumour classification, stage and grade when biologically targeted therapeutics are under consideration. Prominent examples include CD20 positivity for treatment of lymphomas with rituximab, HER2/NEU positivity for treatment of breast cancer with trastuzumab[27], BCR–ABL translocation for treatment of chronic myelogenous leukaemia (CML) with imatinib, and KIT or platelet-derived growth factor receptor-α (PDGFRA) positivity for treatment of gastrointestinal stromal tumours (GIST) with imatinib[28]. As previously discussed, ER positivity or PR positivity is a prerequisite for treatment with tamoxifen or aromatase inhibitors. Similarly, somatic mutations in the tyrosine-kinase domain of the epidermal growth factor receptor (EGFR) have recently been shown to predict a greater efficacy of gefitinib in patients with non-small-cell lung cancer (NSCLC)[29,30]. Some of those markers are FDA-approved and in widespread clinical use (see TABLE 1); others have been assessed only in the research setting. For example, ER, PR and HER2/NEU status are routinely determined for breast cancer, whereas *EGFR* mutations are usually assessed only in clinical trials. Outside of such trials, patients with NSCLC are often given EGFR-antagonists, such as gefitinib, as salvage therapy on an empirical basis without marker studies, especially if they are more likely to have the mutation (that is, patients who are female, never-smokers, diagnosed with adenocarcinoma, or Asian)[31].

Both prognosis and prediction of response are necessary for the selection of neoadjuvant or adjuvant chemotherapy. Tissue classification, TNM staging, molecular biomarkers, grade and other factors might be used in combination for that purpose. The combinations of variables might not be easy to analyse manually, but computer DECISION SUPPORT SYSTEMS (DSS) can make the assessments automatically. For example, Adjuvant Online (see Online links box), a DSS used for breast cancer, estimates 10-year cancer recurrence and survival for women, taking into account their predicted response to adjuvant chemotherapy[32]. Markers can also be used to avoid idiosyncratic drug toxicity such as the sustained, life-threatening leukocyte suppression seen when mercaptopurine is given to leukaemia patients with homozygous mutations of the thiopurine methyltransferase (*TPMT*) gene[33,34].

Biomarkers traditionally used for risk assessment and screening are also available to enhance cancer staging, refine prognosis and estimate response to biological therapy, as summarized in BOX 1 and FIG. 3. An important point: characteristics that suit a molecular marker for one application might not do so for another. For example, a marker to be used in screening the general population must have an extremely high specificity to minimize false positives that necessitate costly or invasive follow-up studies and scare patients and their families needlessly. The same marker need not be so specific if used for high-risk populations and can be even less so once a cancer has been detected. The arguments about use of PSA for screening continue, but its value in monitoring diagnosed prostate cancer or its treatment would be hard to dispute.

Biomarkers on the horizon

Genomic and proteomic technologies have significantly increased the number of potential DNA, RNA and protein biomarkers under study. Here, we will focus on several types with promise for staging.

DNA biomarkers. Circulating DNA and tumour cells were among the first markers evaluated for cancer staging. Increased serum DNA concentrations are associated with cancer (principally metastatic cancer) and with other conditions such as sepsis and autoimmune disease[35,36]. A number of studies suggest circulating tumour cells in the blood[37–39] or bone marrow[40,41] as indicators of systemic metastasis, but the clinical sample sizes have been small and the long-term survival benefit remains to be assessed[42].

Mutations in oncogenes, tumour-suppressor genes, and mismatch-repair genes can serve as DNA biomarkers. For instance, mutations in the oncogene *KRAS* predict metastatic spread in various tumour types[43], and there are mutations in the gene that encodes the tumour suppressor p53 in more than half of sporadic cancers[44,45]. Germline inheritance of a *TP53* mutation (Li–Fraumeni syndrome) confers a risk of developing many of the same cancers. Mutations in other cancer-related genes, such as the *RAS* oncogene or the tumour-suppressor genes *CDKN2A* (cyclin-dependent kinase inhibitor A, which encodes p16INK4A), *APC* (the adenomatous

DECISION SUPPORT SYSTEM
A computerized information system that supports decision-making activities.

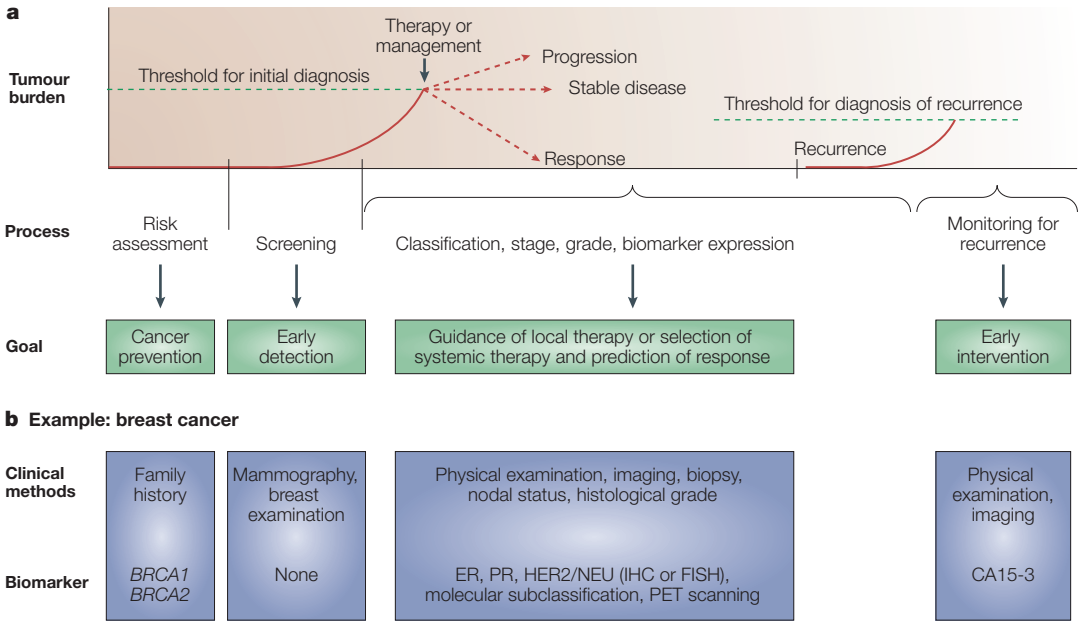

Figure 3 | **Schematic representation of the uses of biomarkers at different stages in the clinical evolution of cancer, with breast cancer biomarkers as an example . a** | Before diagnosis, markers might be used for risk assessment and screening. At diagnosis, markers can assist with staging, grading, and selection of initial therapy. Later, they can be used to monitor therapy, select additional therapy, or monitor for recurrent disease. **b** | Breast cancer biomarkers as an example. Genetic studies of *BRCA1* (breast cancer 1) and *BRCA2* are often performed in patients with a high risk of familial breast cancer. Once the breast cancer is diagnosed, biomarkers are used for subclassification and for prediction of response, particularly to targeted therapies. Cancer antigen (CA) 15-3 might on rare occasions be used to monitor for recurrence. ER, oestrogen receptor; FISH, fluorescence *in situ* hybridization; IHC, immunohistochemistry; PET, positron-emission tomography; PR, progesterone receptor.

polyposis coli gene) and *RB1* (the retinoblastoma gene), also have potential as markers for prognosis or selection of therapy. As discussed previously, the efficacy of anti-EGFR agents such as gefitinib might depend on specific *EGFR* point mutations. Second *EGFR* mutations acquired in patients with NSCLC during therapy have, in some cases, been reported to confer resistance on a previously sensitive tumour, and might therefore be useful prognostic biomarkers[46].

Epigenetic regulation of transcription and translation can also be important in carcinogenesis. Histone deacetylation, lysine-specific histone-H3 methylation, and promoter region CpG methylation can function through transcriptional abrogation of tumour-suppressor genes (for example, *CDKN2A*, *TP53*, *APC* or the breast cancer 1 gene, *BRCA1*)[47–50] or DNA mismatch-repair genes (for example, *MLH1* or the O^6-methyl-guanine-DNA methyltransferase gene, *MGMT*). They can also function through effects on apoptosis, invasion and the cell cycle[50,51]. Gene silencing by CpG methylation has received the most attention, partly because sensitive methods of measurement have become available[36,51,52]. It has been reported, for example, that differences in methylation can distinguish prostate cancer from benign prostatic hyperplasia[53]. Shedding of hypermethylated DNA into saliva from oral malignancies[54,55], into sputum[56] or bronchoalveolar lavage fluid[57] from lung cancer, and into serum from patients with lung[58,59], bladder[60] or colorectal cancer[61,62], has also been demonstrated. Pharmacogenomic effects of methylation silencing, with implications for choice of therapy, have also been shown. For example, promoter region methylation of MGMT, an enzyme that reverses 5′-guanine alkylation, predicts the response or resistance of gliomas to nitrosourea alkylating agents[63,64]. To our knowledge, epigenetic factors have not yet been used in formal staging, but their application to predict response to treatment can be expected in the future, particularly given the development of DNA-demethylating drugs such as 5-azacytidine[65] and zebularine[66], as well as novel histone-deacetylase inhibitors such as depsipeptide (FK228)[67,68].

Other potential DNA biomarkers, including SNPs, mitochondrial DNA markers and oncoviral markers, are discussed in BOX 2.

RNA biomarkers. Whereas most DNA markers are evaluated individually, many high-throughput technologies have been developed to assess mRNA expression comprehensively. Among them are Affymetrix and NimbleGen arrays that are produced by light-directed *in situ* synthesis of oligonucleotides, Rosetta-Agilent ink-jet-printed arrays, DIFFERENTIAL DISPLAY, SERIAL ANALYSIS OF GENE EXPRESSION (SAGE), and BEAD-BASED METHODS[69–71]. Quantitative real-time RT-PCR is generally considered the 'gold standard' against which other methods are validated, and it can now be performed at relatively high-throughput — for example, by using MICROFLUIDIC cards[72].

DIFFERENTIAL DISPLAY
A gel-based technique used to identify transcripts that are differentially expressed between cell or tissue samples.

SERIAL ANALYSIS OF GENE EXPRESSION
(SAGE). A technique for identification and quantitation of transcript expression levels. SAGE is based on a process in which short oligonucleotide 'tags' from defined locations within a transcript are spliced together and sequenced for identification of the transcript.

BEAD-BASED METHODS
Methods of measurement based on small or microscopic beads (as opposed, for example, to the flat surfaces characteristic of microarrays).

MICROFLUIDICS
Technology that allows the use of very small volumes of reagents, shortening reaction times and facilitating scale-up of molecular methods.

Box 2 | **DNA biomarkers under evaluation**

Single-nucleotide polymorphisms

Particular single-nucleotide polymorphisms (SNPs) are associated with increased cancer risk, and HAPLOTYPE assessment can be predictive for several cancers, including those of prostate, breast and lung. Because five well-known cancer susceptibility genes (ataxia telangiectasia mutated (*ATM*), breast cancer 1 (*BRCA1*), *BRCA2*, *RAD51* and *TP53*) show low haplotype diversity within ethnicities[125–131], perhaps as few as 10% of the SNPs will have to be sequenced for useful haplotype-based risk assessment. Genome-wide SNP analysis has been reported[132,133], but SNPs are not currently used for formal cancer staging or grading.

Oncoviral markers

Altered immune regulation in HIV/AIDS is associated with several types of virally mediated tumours, including Kaposi sarcoma (associated with human herpesvirus 8, HHV8), AIDS-related lymphomas (associated with HHV8 and Epstein–Barr virus, EBV) and cervical cancer (associated with human papillomavirus, HPV)[134]. Links between latent viral infection and cancer have also been noted for hepatocellular carcinoma (associated with hepatitis B virus, HBV, and hepatitis C virus, HCV), adult T-cell leukaemia (associated with human T-cell lymphotrophic virus type 1, HTLV1), nasopharyngeal carcinoma (associated with EBV)[135,136], Hodgkin disease (associated with EBV) and endemic Burkitt lymphoma (associated with EBV)[137]. Prophylactic immunization of women who are negative for the HPV16 L1, E6 and E7 oncoprotein markers is reported to eliminate their risk for HPV16-related cervical intraepithelial neoplasia[138]. In accordance with the goals of staging, viral markers might be used at or after diagnosis to predict treatment response and prognosis. For example, antiviral-mediated elimination of HHV8 and HIV viraemia has been associated with clinical response of Kaposi sarcoma[139].

Mitochondrial DNA aberrations

Because somatic cells contain as many as 500 mitochondria — each with multiple, usually identical mitochondrial DNA (mtDNA) molecules — the extraordinarily high gene copy number makes cancer-associated mtDNA mutations particularly easy to detect. Mutations in mtDNA occur in cancers of the colon[140], bladder, head and neck, lung, breast[141], kidney[142] and testis[141,143]. However, despite the promise of mtDNA for cancer screening, its value for tumour classification, staging or grading has yet to be determined.

HAPLOTYPE
A way of denoting the collective genotype of a number of closely linked loci on a chromosome that tend to be inherited together in a population.

SUPERVISED ALGORITHM
A method of statistical or machine learning in which a model is fitted to observations. The algorithm, in effect, learns by example.

LASER-CAPTURE MICRODISSECTION
A laser-based technology used to obtain materials from selected regions of cut tissue or tumour sections on glass slides. The method is used, for example, to obtain relatively pure populations of tumour cells from the heterogeneous mixture of cells in a tumour.

CYTOKERATIN
A protein component of intermediate filaments found in epithelial cells.

Most RNA-based biomarkers undergoing clinical evaluation consist of multi-gene molecular patterns or 'fingerprints'. Although such patterns can be more accurate than single-molecule markers, choosing which genes to include in the pattern adds an additional layer of statistical complexity, prompting new developments in biostatistics, bioinformatics and data visualization. Molecular markers and their patterns have been analysed by various SUPERVISED ALGORITHMS, most prominently by double hierarchical clustering methods that lead to colour-coded 'clustered image maps' (CIMs). We introduced CIMs in the early 1990s to illuminate patterns of similarity and difference in high-throughput DNA, mRNA, protein and pharmacological profiling studies[73], and they have since become the most familiar visual icon of 'post-genomic' biology. Various supervised statistical and machine-learning methods for classifying tumours have also been introduced. Both the experimental technologies and the methods of analysis have been reviewed extensively elsewhere[74,75].

Pattern-based RNA-expression analysis of clinical breast cancers has identified previously unknown molecular subtypes that are associated with differences in survival[76–78]. That analysis has also provided increased prognostic capability[79,80], predicted response to neoadjuvant therapy[81,82], predicted the likelihood of metastasis in lymph-node negative patients[83] and correctly predicted tumour grade from LASER-CAPTURE MICRODISSECTED specimens. The transcript levels of enzymes important for drug metabolism have been used preclinically to predict the response to chemotherapy in lung[84] and colon[85] cancers. Similar approaches have led to novel discoveries for other cancers, including melanoma[86,87], leukaemias, lymphomas[88–90], and carcinomas of the lung[87,91], prostate[92] and colon[93]. Such cancer 'snapshots' taken at the time of diagnosis can be expected to further the goals of cancer staging. Extensive validation studies will be required, however, to move those developments from clinical research to standard practice in staging. Several companies have attempted to do so and have made their RT-PCR-based gene signatures available to the public for use in predicting survival. However, these and other RNA-based markers have not yet undergone rigorous, prospective clinical validation, and they have not been approved by the FDA.

Protein biomarkers. As shown in TABLE 1 and mentioned before, all but a handful of the FDA-approved cancer biomarkers in clinical use are single proteins, and most are serum-derived. AFP, β-HCG and LDH are used in the AJCC system to stage testicular cancer. Other proteins, although not formally used for staging, are important for prognosis and selection of therapy. For example, the expression of HER2/NEU and CYTOKERATINS can be used to refine the prognosis of breast cancers. HER2/NEU, EGFR and KIT are used clinically to predict if breast cancers, colon cancers or GISTs will respond to trastuzumab, cetuximab or imatinib, respectively. Similarly, expression of ER or PR is necessary for hormonal therapies to be effective against breast cancer.

Just as pattern-based RNA biomarkers frequently outperform single RNA markers in tumour classification, prognosis or prediction of response to therapy, protein-based 'fingerprints' may outperform individual protein markers. Technologies such as differential in-gel electrophoresis (DIGE)[94], two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) and multidimensional protein-identification technology (MudPIT) can be used for higher-throughput profiling with microgram quantities of protein. Other high-throughput technologies, such as the REVERSE-PHASE MICROARRAY[6,95,96] (FIG. 2) and surface-enhanced laser desorption ionization time-of-flight (SELDI-TOF) mass spectrometry, are more sensitive (in the femtomolar range) and can cover more of the 12 orders of magnitude range of serum-protein expression levels[5,75,97,98]. Emerging nanotechnologies, such as IMMUNO-PCR[99,100], FIELD EFFECT TRANSISTOR (FET)-BASED PROTEIN DETECTION[101] and QUANTUM DOTS[102–105], promise further increases in the sensitivity of protein markers, but those techniques are currently experimental.

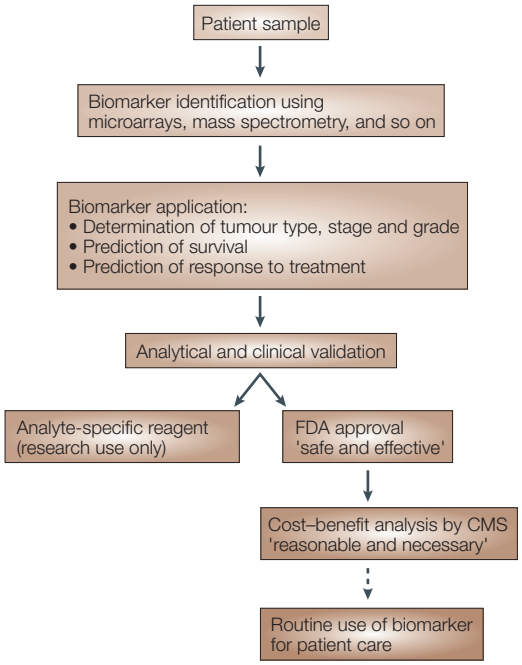

Figure 4 | **Chronology of biomarker development.** A biomarker is first identified, then evaluated for a particular clinical indication. Analytical and clinical validations must be performed before submission for US Food and Drug Administration (FDA) approval. Alternatively, the marker might bypass the FDA approval process if it is to be used for 'research purposes only'. Once a marker is FDA-approved, the Center for Medicaid and Medicare Services (CMS) might determine that it is 'reasonable and necessary' for improved patient care and, therefore, reimbursable. Because CMS decisions indirectly influence coverage by private insurance carriers, a marker is not widely used in the clinic unless all of the steps in the process have been completed.

Protein quantity by itself might not be the salient marker parameter. Protein function is instead often dependent on phosphorylation, glycosylation, other post-translational modifications, location in the cell and/or the location in the tissue. The important phosphorylation-dependent signalling cascades can be assessed, for example, using reverse-phase arrays[6,95,96]. Laser-capture microdissection and similar technologies can be used to obtain DNA, mRNA or protein from precise locations within a tumour and thereby distinguish markers inherent to the malignant cells from those in other cell types within the tumour. Microdissection has enhanced expression profiling of breast[10,106,107], ovarian[108], oral[109] and prostate[110] cancers, as well as other cancer types[97].

Biomarker validation

Regardless of their intended use, scientifically vetted biomarkers must clear a number of practical hurdles before they can be considered for clinical practice (FIG. 4). Five conceptual phases of biomarker development have been proposed: preclinical exploratory (I), clinical assay and validation (II), retrospective longitudinal (III), prospective screening (IV) and cancer control (V)[111,112]. The process is fraught with difficulties, and most candidate markers are still in the early phases of development. The clinical studies to date are generally retrospective, and the few prospective studies that have been conducted have often yielded inconsistent results. Furthermore, most include too few patients, making it problematic to alter treatment decisions based on their results. By contrast, large therapeutic clinical trials often include thousands of patients, and the results are usually replicated in separate studies before integrating their conclusions into clinical practice.

Ideally, biomarkers should be validated analogously in prospective, well-controlled clinical studies of diverse patients across multiple institutions, with well-established standards for all steps in the process. Those steps include, for example, tissue collection, purification, amplification (if necessary), hybridization or ligand-binding, data capture, normalization, statistical analysis and scoring[2]. Furthermore, there should be reliable reproducibility within and among laboratories. However, those ideal conditions rarely apply. There is widespread recognition that standards and standardization are required, but consensus is hard to achieve. There is also the continuing concern that many of the salient technologies are not yet mature enough for standard operating procedures to be set in stone.

It is beyond the scope of this article to review methods of data analysis for identifying and validating biomarkers, but a few cautionary words might be in order. The multivariate statistical and machine-learning algorithms used to define markers are prone to OVERFITTING. The more flexible and non-linear the algorithm, the greater the danger. An algorithm may perform well on the original sample set (that is, the training set) but fail when applied to independent validation samples. Careful validation is especially important for patterns of markers[113].

Multi-institutional teamwork, through participation in large collaborative oncology groups, can advance both analytical and clinical validation by creating standards and experimental designs that use limited patient samples and resources most effectively. Discussions with those aims are underway in organizations such as the Cancer Therapy Evaluation Program (CTEP; see Online links box) of the National Cancer Institute (NCI). The NCI's Cancer Diagnosis Program has established a Program for the Assessment of Clinical Cancer Tests (PACCT)[129], and the NCI's Early Detection Research Network (EDRN), in collaboration with the National Institute of Standards and Technology, continues to support "coordination among biomarker development laboratories, biomarker validation laboratories, clinical repositories, and population-screening programs"[111]. Analogous joint efforts between industry and academia combine the strengths of academic institutions (including access to archived clinical specimens and clinical study patients) with those of industry (including high-throughput drug and biomarker discovery/development programmes). The price of multi-institution studies, of course, is that they are harder to manage and coordinate.

REVERSE-PHASE MICROARRAY
A microarray spotted with numerous tissue or cell lysates and subsequently incubated with a detection ligand (usually an antibody) to quantitate protein in the lysates.

IMMUNO-PCR
A sensitive method for detection of proteins using a combination of PCR and conventional immunodetection. A bi-specific linker molecule with affinity for DNA and an antibody is used to attach a DNA marker to a specific antigen, resulting in an antigen–antibody–DNA complex that can be quantified using PCR.

FIELD EFFECT TRANSISTOR-BASED PROTEIN DETECTION
Technology for detecting proteins based on their completion of a circuit between two electrodes in a transistor, thereby resulting in a measurable increase in current.

QUANTUM DOTS
Semiconductor particles with size-dependent fluorescence-emission wavelengths visualized by laser-excitation spectrometry.

OVERFITTING
In multivariate predictive analysis, a statistical model can be overfitted if it has too many free parameters for the number and type of cases in the training set. The result can be a model that fits the training data set very well but does poorly when applied to other data.

Regulatory process. We will focus here on the situation in the United States, recognizing that similar issues arise elsewhere as well. As a result of the medical device amendments of 1976, the FDA has been charged with regulating *in vitro* diagnostic devices (IVDs), including tumour markers, to ensure their 'safety' and 'effectiveness' as defined in the Code of Federal Regulations[114,115]. The 'safety and effectiveness' of an IVD refers to the consequences expected from reliance on it to make 'clinically significant' diagnostic or treatment decisions[115]. Unless already used to detect recurrent disease or monitor the effectiveness of therapy, most tumour markers require clinical studies to support a manufacturer's claim of effectiveness and safety before they receive FDA approval for clinical laboratory use. In analogy with an investigational new drug (IND) application, which is required before trials begin with a novel agent, a pre-investigational device exemption (IDE) allows trials that are designed to evaluate a new IVD[116]. The arduous, expensive clinical validation process is predicated on a thorough analytical validation of all reagents and machines used.

A somewhat less onerous path for transition from laboratory to clinic is provided by classification as an ASR. That designation allows biomarkers to be used in an institution's in-house clinical laboratory for restricted 'research-only purposes'. Rather than the FDA, the Center for Medicaid and Medicare Services (CMS), in accordance with the clinical laboratory improvement amendments of 1988, monitors both the manufacture and laboratory use of ASRs. CMS oversight helps to ensure that biomarkers classified as ASRs are accurate, safe for patients and laboratory personnel, and available for research use while their potential clinical utility is being explored before submission for FDA-approval.

Financial resources. A final hurdle to the widespread integration of biomarkers is their medical economics. The escalation of medical costs in the United States has focused the nation's attention on cost–benefit analysis — an often difficult and controversial hurdle to overcome, even for established biomarkers such as PSA. Some prospective benefits, such as refined prognostic accuracy and increased patient comfort, are difficult to place a value on objectively. Others, such as predicted responses to treatment with biologically targeted therapies, are more easily quantified.

In the United States, the CMS determines whether the government, through Medicare or Medicaid, will reimburse laboratories for IVD testing using a 'reasonable and necessary' standard. That decision trickles down to private insurers because most echo Medicare's reimbursement practices. Tumour marker IVDs may, of course, be used without reimbursement, but they are expensive and unlikely to be applied widely without strong evidence of clinical benefit. For that reason, most markers are used only in the research setting until approved by the FDA and supported by insurers. Similarly, markers used in other countries must successfully overcome their respective regulatory and financial hurdles before being widely used for patient care.

Conclusion

New high-throughput 'omic' technologies[144,145] in post-genomic biology have yielded many potential biomarkers and biomarker patterns, some of which may prove useful for staging and grading cancers[117]. The potential is enormous. Few markers, however, have so far been integrated into clinical practice. Metaphorically speaking, the 'water' is everywhere, but little is yet ready to drink. We have considered a number of the readily definable reasons here. Not so easy to assess or quantify are many of the practical issues, such as the queasiness of pharmaceutical companies about fractionation of their markets and their medico-legal fear of generating and possessing too much information that could later be used to their detriment. Furthermore, acceptance and adoption by practicing physicians and patients can take time. But as therapies become increasingly target specific, biomarkers will inevitably develop in tandem to play greater roles in staging, grading, and selection of therapy. The flood of potential biomarkers is opening the way to a more individualized practice of oncology, although the practical problems are many and difficult.

1. Sobin, L. H. TNM: evolution and relation to other prognostic factors. *Semin. Surg. Oncol.* **21**, 3–7 (2003). **Historical perspective of the TNM staging guidelines.**
2. Hammond, M. E. & Taube, S. E. Issues and barriers to development of clinically useful tumor markers: a development pathway proposal. *Semin. Oncol.* **29**, 213–221 (2002). **Provides useful guidelines for the discovery, evaluation and validation of clinical biomarkers.**
3. Negm, R. S., Verma, M. & Srivastava, S. The promise of biomarkers in cancer screening and detection. *Trends Mol. Med.* **8**, 288–293 (2002).
4. US Department of Health & Human Services; US Food and Drug Administration. *Challenge and Opportunity on the Critical Path of New Medical Products*. http://www.fda.gov/oc/initiatives/criticalpath/whitepaper.html (2004).
5. Anderson, N. L. & Anderson, N. G. The human plasma proteome: history, character, and diagnostic prospects. *Mol. Cell. Proteomics* **1**, 845–867 (2002).
6. Nishizuka, S. *et al.* Diagnostic markers that distinguish colon and ovarian adenocarcinomas: identification by genomic, proteomic, and tissue array profiling. *Cancer Res.* **63**, 5243–5250 (2003). **Clear demonstration that numerous genomic, proteomic and tissue-based arrays can be used in concert to distinguish between colon and ovarian carcinomas.**
7. Leong, P. P. *et al.* Distinguishing second primary tumors from lung metastases in patients with head and neck squamous cell carcinoma. *J. Natl Cancer Inst.* **90**, 972–977 (1998).
8. Califano, J. *et al.* Unknown primary head and neck squamous cell carcinoma: molecular identification of the site of origin. *J. Natl Cancer Inst.* **91**, 599–604 (1999).
9. Califano, J. *et al.* Genetic progression and clonal relationship of recurrent premalignant head and neck lesions. *Clin. Cancer Res.* **6**, 347–352 (2000).
10. Ma, X. *et al.* Gene expression profiles of human breast cancer progression. *Proc. Natl Acad. Sci. USA* **100**, 5974–5979 (2003).
11. Abulafia, O. & Sherer, D. M. Automated cervical cytology: meta-analyses of the performance of the PAPNET system. *Obst. Gynecol. Surv.* **54**, 253–264 (1999).
12. Destounis, S. V. *et al.* Can computer-aided detection with double reading of screening mammograms help decrease the false-negative rate? Initial experience. *Radiology* **232**, 578–584 (2004).
13. Awai, K. *et al.* Pulmonary nodules at chest CT: effect of computer-aided diagnosis on radiologists' detection performance. *Radiology* **230**, 347–352 (2004).
14. Erickson, B. J. & Bartholmai, B. Computer-aided detection and diagnosis at the start of the third millennium. *J. Digit. Imaging* **15**, 59–68 (2002).
15. American Joint Committee on Cancer & American Cancer Society. *AJCC Cancer Staging Handbook: from the AJCC Cancer Staging Manual* (eds Greene, F. L. *et al.*) (Springer, New York, 2002). **AJCC guidelines for cancer staging.**
16. McQuade, P. & Knight, L. C. Radiopharmaceuticals for targeting the angiogenesis marker $\alpha_v\beta_3$. *Q. J. Nucl. Med.* **47**, 209–220 (2003).
17. Blankenberg, F. G. *et al.* *In vivo* detection and imaging of phosphatidylserine expression during programmed cell death. *Proc. Natl Acad. Sci. USA* **95**, 6349–6354 (1998).

18. Kuge, Y. *et al.* Feasibility of 99mTc-annexin V for repetitive detection of apoptotic tumor response to chemotherapy: an experimental study using a rat tumor model. *J. Nucl. Med.* **45**, 309–312 (2004).
19. Krohn, K. A. Evaluation of alternative approaches for imaging cellular growth. *Q. J. Nucl. Med.* **45**, 174–178 (2001).
20. Kostakoglu, L., Agress, H. Jr & Goldsmith, S. J. Clinical role of FDG PET in evaluation of cancer patients. *Radiographics* **23**, 315–340 (2003).
21. Kostakoglu, L. & Goldsmith, S. J. 18F-FDG PET evaluation of the response to therapy for lymphoma and for breast, lung, and colorectal carcinoma. *J. Nucl. Med.* **44**, 224–239 (2003).
22. Britz-Cunningham, S. H. & Adelstein, S. J. Molecular targeting with radionuclides: state of the science. *J. Nucl. Med.* **44**, 1945–1961 (2003).
23. Agrawal, M. *et al.* Increased 99mTc-sestamibi accumulation in normal liver and drug-resistant tumors after the administration of the glycoprotein inhibitor, XR9576. *Clin. Cancer Res.* **9**, 650–656 (2003).
24. Revillion, F., Bonneterre, J. & Peyrat, J. P. ERBB2 oncogene in human breast cancer and its clinical significance. *Eur. J. Cancer* **34**, 791–808 (1998).
25. Press, M. F. *et al.* Amplification and overexpression of HER-2/neu in carcinomas of the salivary gland: correlation with poor prognosis. *Cancer Res.* **54**, 5675–5682 (1994).
26. Singletary, S. E. *et al.* Revision of the American Joint Committee on Cancer staging system for breast cancer. *J. Clin. Oncol.* **20**, 3628–3636 (2002).
27. Druker, B. J. *et al.* Activity of a specific inhibitor of the BCR–ABL tyrosine kinase in the blast crisis of chronic myeloid leukemia and acute lymphoblastic leukemia with the Philadelphia chromosome. *N. Engl. J. Med.* **344**, 1038–1042 (2001).
28. Balmain, A., Gray, J. & Ponder, B. The genetics and genomics of cancer. *Nature Genet.* **33** (Suppl.), 238–244 (2003).
29. Lynch, T. J. *et al.* Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med.* **350**, 2129–2139 (2004).
30. Paez, J. G. *et al*. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* **304**, 1497–1500 (2004).
Evidence that subtle mutations in *EGFR* can affect the clinical response to gefitinib.
31. Kaneda, H. *et al*. Retrospective analysis of the predictive factors associated with the response and survival benefit of gefitinib in patients with advanced non-small-cell lung cancer. *Lung Cancer* **46**, 247–254 (2004).
32. Ravdin, P. M. *et al*. Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. *J. Clin. Oncol.* **19**, 980–991 (2001).
33. Relling, M. & Dervieux, T. Pharmacogenetics and cancer therapy. *Nature Rev. Cancer* **1**, 99–108 (2001).
34. Lenz, H. J. The use and development of germline polymorphisms in clinical oncology. *J. Clin. Oncol.* **22**, 2519–2521 (2004).
35. Leon, S. A., Shapiro, B., Sklaroff, D. M. & Yaros, M. J. Free DNA in the serum of cancer patients and the effect of therapy. *Cancer Res.* **37**, 646–650 (1977).
36. Sidransky, D. Emerging molecular markers of cancer. *Nature Rev. Cancer* **2**, 210–219 (2002).
In-depth review of diverse biomarkers undergoing clinical evaluation.
37. Kahn, H. J. *et al*. Enumeration of circulating tumor cells in the blood of breast cancer patients after filtration enrichment: correlation with disease stage. *Breast Cancer Res. Treat.* **86**, 237–247 (2004).
38. Stathopoulou, A. *et al*. Real-time quantification of CK-19 mRNA-positive cells in peripheral blood of breast cancer patients using the lightcycler system. *Clin. Cancer Res.* **9**, 5145–5151 (2003).
39. Hoon, D. S. *et al*. Molecular markers in blood as surrogate prognostic indicators of melanoma recurrence. *Cancer Res.* **60**, 2253–2257 (2000).
40. Muller, V. & Pantel, K. Bone marrow micrometastases and circulating tumor cells: current aspects and future perspectives. *Breast Cancer Res.* **6**, 258–261 (2004).
41. Pantel, K. & Brakenhoff, R. H. Dissecting the metastatic cascade. *Nature Rev. Cancer* **4**, 448–456 (2004).
42. Meng, S. *et al*. Circulating tumor cells in patients with breast cancer dormancy. *Clin. Cancer Res.* **10**, 8152–8162 (2004).
43. Brennon, J. A. *et al*. Molecular assesment of histopathological staging in squamous-cell carcinoma of the head and neck. *N. Engl. J. Med.* **332**, 429–435 (1995).
44. Levine, A. J. p53, the cellular gatekeeper for growth and division. *Cell* **88**, 323–331 (1997).
45. Nigro, J. M., *et al*. Mutations in the p53 gene occur in diverse tumor types. *Nature* **342**, 705–708 (1989).
46. Kobayashi, S. *et al*. EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med.* **352**, 786–792 (2005).
47. Merlo, A. *et al*. 5′ CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers. *Nature Med.* **1**, 686–692 (1995).
48. Herman, J. G. *et al*. Silencing of the *VHL* tumor-suppressor gene by DNA methylation in renal carcinoma. *Proc. Natl Acad. Sci. USA* **91**, 9700–9704 (1994).
49. Jenuwein, T. & Allis, C. Translating the histone code. *Science* **293**, 1074–1080 (2001).
50. Esteller, M., Corn, P. G., Baylin, S. B. & Herman, J. G. A gene hypermethylation profile of human cancer. *Cancer Res.* **61**, 3225–3229 (2001).
51. Herman, J. G., Graff, J. R., Myohanen, S., Nelkin, B. D. & Baylin, S. B. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proc. Natl Acad. Sci. USA* **93**, 9821–9826 (1996).
52. Frommer, M. *et al*. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual dna strands. *Proc. Natl Acad. Sci. USA* **89**, 1827–1831 (1992).
53. Jeronimo, C. *et al*. Quantitation of GTSP1 methylation in non-neoplastic prostatic tissue and organ-confined prostate adenocarcinoma. *J. Natl Cancer Inst.* **93**, 1747–1752 (2001).
54. Rosas, S. L. *et al*. Promoter hypermethylation patterns of p16, o6-methylguanine-DNA-methyltransferase, and death-associated protein kinase in tumors and saliva of head and neck cancer patients. *Cancer Res.* **61**, 939–942 (2001).
55. Belinsky, S. A. *et al*. Aberrant methylation of p16INK4a is an early event in lung cancer and a potential biomarker for early diagnosis. *Proc. Natl Acad. Sci. USA* **95**, 11891–11896 (1998).
56. Palmisano, W. A. *et al*. Predicting lung cancer by detecting aberrant promoter methylation in sputum. *Cancer Res.* **60**, 5954–5958 (2000).
57. Ahrendt, S. A. *et al*. Molecular detection of tumor cells in bronchoalveolar lavage fluid from patients with early stage lung cancer. *J. Natl Cancer Inst.* **91**, 332–339 (1999).
58. Ramirez, J. L. *et al*. Methylation patterns and K-ras mutations in tumor and paired serum of resected non-small-cell lung cancer patients. *Cancer Lett.* **193**, 207–216 (2003).
59. Usadel, H. *et al*. Quantitative adenomatous polyposis coli promoter methylation analysis in tumor tissue, serum, and plasma DNA of patients with lung cancer. *Cancer Res.* **62**, 371–375 (2002).
60. Valenzuela, M. T. *et al*. Assessing the use of *p16INK4a* promoter gene methylation in serum for detection of bladder cancer. *Eur. Urol.* **42**, 622–630 (2002).
61. Nakayama, H. *et al*. Molecular detection of p16 promoter methylation in the serum of recurrent colorectal cancer patients. *Intl J. Cancer* **105**, 491–493 (2003).
62. Yamaguchi, S., Asao, T., Nakamura, J. I., Ide, M. & Kuwano, H. High frequency of DAP-kinase gene promoter methylation in colorectal cancer specimens and its identification in serum. *Cancer Lett.* **194**, 99–105 (2003).
63. Esteller, M. *et al*. Inactivation of the DNA-repair gene *MGMT* and the clinical response of gliomas to alkylating agents. *N. Engl J. Med.* **343**, 1350–1354 (2000).
64. Weinstein, J. N. Pharmacogenomics — teaching old drugs new tricks. *N. Engl. J. Med.* **343**, 1408–1409 (2000).
65. Esteller, M. DNA methylation and cancer therapy: new developments and expectations. *Curr. Opin. Oncol.* **17**, 55–60 (2005).
66. Cheng, J. C. *et al*. Preferential response of cancer cells to zebularine. *Cancer Cell* **6**, 151–158 (2004).
67. Piekarz, R. & Bates, S. A review of depsipeptide and other histone deacetylase inhibitors in clinical trials. *Curr. Pharm. Des.* **10**, 2289–2298 (2004).
68. Piekarz, R. L. *et al*. T-cell lymphoma as a model for the use of histone deacetylase inhibitors in cancer therapy: impact of depsipeptide on molecular markers, therapeutic targets, and mechanisms of resistance. *Blood* **103**, 4636–4643 (2004).
69. Mischel, P. S., Cloughesy, T. F. & Nelson, S. F. DNA-microarray analysis of brain cancer: molecular classification for therapy. *Nature Rev. Neurosci.* **5**, 782–792 (2004).
70. Gray, J. W. & Collins, C. Genome changes and gene expression in human solid tumors. *Carcinogenesis* **21**, 443–452 (2000).
71. Velculescu, V. E., Zhang, L., Vogelstein, B. & Kinzler, K. W. Serial analysis of gene expression. *Science* **270**, 484–487 (1995).
72. Weigl, B. H., Bardell, R. L. & Cabrera, C. R. Lab-on-a-chip for drug development. *Adv. Drug. Deliv. Rev.* **55**, 349–377 (2003).
73. Weinstein, J. N. *et al*. An information-intensive approach to the molecular pharmacology of cancer. *Science* **275**, 343–349 (1997).
74. Papin, J. A, Hunter, T., Palsson, B. O. & Subramaniam, S. Reconstruction of cellular signalling networks and analysis of their properties. *Nature Rev. Mol. Cell Biol.* **6**, 99–111 (2005).
75. Petricoin, E. F., Zoon, K. C., Kohn, E. C., Barrett, J. C. & Liotta, L. A. Clinical proteomics: translating benchside promise into bedside reality. *Nature Rev. Drug Discov.* **1**, 683–695 (2002).
Comprehensive review of clinical proteomic technologies and associated pattern-based protein markers.
76. Perou, C. *et al*. Molecular portraits of human breast tumours. *Nature* **406**, 747–752 (2000).
77. Sorlie, T., Perou, C., Brown, P., Botstein, D. & Borresen-Dale, A. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl Acad. Sci. USA* **98**, 10869–10874 (2001).
78. Sotiriou, C. *et al*. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *Proc. Natl Acad. Sci. USA* **100**, 10393–10398 (2003).
79. van de Vijver, M. J. *et al*. A gene-expression signature as a predictor of survival in breast cancer. *N. Engl. J. Med.* **347**, 1999–2009 (2002).
80. van't Veer, L. J. *et al*. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415**, 530–536 (2002).
81. Chang, J. C. *et al*. Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. *Lancet* **362**, 362–369 (2003).
82. Sotiriou, C. *et al*. Gene expression profiles derived from fine needle aspiration correlate with response to systemic chemotherapy in breast cancer. *Breast Cancer Res.* **4**, R3 (2002).
83. Ramaswamy, S., Ross, K., Lander, E. & Golub, T. A molecular signature of metastasis in primary solid tumors. *Nature Genet.* **33**, 49–54 (2003).
84. Rosell, R. *et al*. Ribonucleotide reductase messenger RNA expression and survival in gemcitabine/cisplatin-treated advanced non-small cell lung cancer patients. *Clin. Cancer Res.* **10**, 1318–1325 (2004).
85. Salonga, D. *et al*. Colorectal tumors responding to 5-fluorouracil have low gene expression levels of dihydropyrimidine dehydrogenase, thymidylate synthase, and thymidine phosphorylase. *Clin. Cancer Res.* **6**, 1322–1327 (2000).
86. Bittner, M. *et al*. Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature* **406**, 536–540 (2000).
87. Carr, K. M., Bittner, M. & Trent, J. M. Gene-expression profiling in human cutaneous melanoma. *Oncogene* **22**, 3076–3080 (2003).
88. Golub, T. R. *et al*. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **286**, 531–537 (1999).
89. Alizadeh, A. A. *et al*. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* **403**, 503–511 (2000).
90. Lossos, I. S. *et al*. Transformation of follicular lymphoma to diffuse large-cell lymphoma: alternative patterns with increased or decreased expression of *c-myc* and its regulated genes. *Proc. Natl Acad. Sci. USA* **99**, 8886–8891 (2002).
91. Garber, M. E. *et al*. Diversity of gene expression in adenocarcinoma of the lung. *Proc. Natl Acad. Sci. USA* **98**, 13784–13789 (2001).
92. Best, C. J. *et al*. Molecular differentiation of high- and moderate-grade human prostate cancer by cDNA microarray analysis. *Diagn. Mol. Pathol.* **12**, 63–70 (2003).
93. Zou, T. T. *et al*. Application of cDNA microarrays to generate a molecular taxonomy capable of distinguishing between colon cancer and normal colon. *Oncogene* **21**, 4855–4862 (2002).
94. Zhou, G. *et al*. 2D differential in-gel electrophoresis for the identification of esophageal scans cell cancer-specific protein markers. *Mol. Cell. Proteomics* **1**, 117–124 (2002).
95. Nishizuka, S. *et al*. Proteomic profiling of the NCI-60 cancer cell lines using new high-density reverse-phase lysate microarrays. *Proc. Natl Acad. Sci. USA* **100**, 14229–14234 (2003).
96. Paweletz, C. P. *et al*. Reverse phase protein microarrays which capture disease progression show activation of pro-survival pathways at the cancer invasion front. *Oncogene* **20**, 1981–1989 (2001).
97. Verma, M., Wright, G. L. Jr, Hanash, S. M., Gopal-Srivastava, R. & Srivastava, S. Proteomic approaches within the NCI early detection research network for the discovery and identification of cancer biomarkers. *Ann. NY Acad. Sci.* **945**, 103–115 (2001).

98. Wulfkuhle, J. D., Liotta, L. A. & Petricoin, E. F. Proteomic applications for the early detection of cancer. *Nature Rev. Cancer* **3**, 267–275 (2003).
99. Sano, T., Smith, C. L. & Cantor, C. R. Immuno-PCR: very sensitive antigen detection by means of specific antibody–DNA conjugates. *Science* **258**, 120–122 (1992).
100. Nam, J. M., Thaxton, C. S. & Mirkin, C. A. Nanoparticle-based bio-bar codes for the ultrasensitive detection of proteins. *Science* **301**, 1884–1886 (2003).
101. Service, R. F. American Chemical Society Meeting. Tiny transistors scout for cancer. *Science* **300**, 242–243 (2003).
102. Seydel, C. Quantum dots get wet. *Science* **300**, 80–81 (2003).
103. Wu, X. *et al*. Immunofluorescent labeling of cancer marker Her2 and other cellular targets with semiconductor quantum dots. *Nature Biotechnol.* **21**, 41–46 (2003).
104. Ben-Ari, E. T. Nanoscale quantum dots hold promise for cancer applications. *J. Natl Cancer Inst.* **95**, 502–504 (2003).
105. Igor, M. *et al*. Self-assembled nanoscale biosensors based on quantum dot FRET donors. *Nature Mater.* **2**, 630–638 (2003).
106. Wulfkuhle, J. D. *et al*. Proteomics of human breast ductal carcinoma *in situ*. *Cancer Res.* **62**, 6740–6749 (2002).
107. Wulfkuhle, J. D. *et al*. New approaches to proteomic analysis of breast cancer. *Proteomics* **1**, 1205–1215 (2001).
108. Jones, M. B. *et al*. Proteomic analysis and identification of new biomarkers and therapeutic targets for invasive ovarian cancer. *Proteomics* **2**, 76–84 (2002).
109. Patel, V., Leethanakul, C. & Gutkind, J. S. New approaches to the understanding of the molecular basis of oral cancer. *Crit. Rev. Oral Biol. Med.* **12**, 55–63 (2001).
110. Ornstein, D. K. *et al*. Proteomic analysis of laser capture microdissected human prostate cancer and *in vitro* prostate cell lines. *Electrophoresis* **21**, 2235–2242 (2000).
111. Pepe, M. S. *et al*. Phases of biomarker development for early detection of cancer. *J. Natl Cancer Inst.* **93**, 1054–1061 (2001).
112. Barker, P. E. Cancer biomarker validation: standards and process: roles for the National Institute of Standards and Technology (NIST). *Ann. NY Acad. Sci.* **983**, 142–150 (2003).
113. Ransohoff, D. F. Rules of evidence for cancer molecular marker discovery and validation. *Nature Rev. Cancer* **4**, 309–314 (2004).
Effectively summarizes the problem of data overfitting in discovery-based research and the need for 'rules of evidence' to ensure clinical validity.
114. Gutman, S. Regulatory issues in tumor marker development. *Semin. Oncol.* **29**, 294–300 (2002).
Examines the regulatory environment surrounding biomarker approval by the FDA and CMS.
115. Taube, S. E. & Freiberg, G. P. Regulatory issues related to marker development. *Urol. Oncol.* **5**, 214–216 (2000).
116. Hackett, J. L. & Lesko, L. J. Microarray data — the US FDA, industry and academia. *Nature Biotechnol.* **21**, 742–743 (2003).
117. Mendelsohn, A. R. & Brent, R. Postgenomic protein analysis: the next bend in the river. *Nature Biotechnol.* **16**, 520–521 (1998).
118. Robson, M. E. *et al*. A combined analysis of outcome following breast cancer: differences in survival based on *BRCA1*/*BRCA2* mutation status and administration of adjuvant treatment. *Breast Cancer Res.* **6**, R8–R17 (2004).
119. Einspahr, J. G. *et al*. Surrogate end-point biomarkers as measures of colon cancer risk and their use in cancer chemoprevention trials. *Cancer Epidemiol. Biomarkers Prev.* **6**, 37–48 (1997).
120. Togashi, K. Ovarian cancer: the clinical role of US, CT, and MRI. *Eur. Radiol.* **13** (Suppl. 4), L87–L104 (2003).
121. Schlieman, M. G., Ho, H. S. & Bold, R. J. Utility of tumor markers in determining resectability of pancreatic cancer. *Arch. Surg.* **138**, 951–956 (2003).
122. Depres-Brummer, P. *et al*. The usefulness of CA15.3, mucin-like carcinoma-associated antigen and carcinoembryonic antigen in determining the clinical course in patients with metastatic breast cancer. *J. Cancer Res. Clin. Oncol.* **121**, 419–422 (1995).
123. Duncan, J. L., Price, A. & Rogers, K. The use of CA15.3 as a serum tumour marker in breast carcinoma. *Eur. J. Surg. Oncol* **17**, 16–19 (1991).
124. Gion, M. *et al*. Tumor markers in breast cancer monitoring should be scheduled according to initial stage and follow-up time: a prospective study on 859 patients. *Cancer J.* **7**, 181–190 (2001).
125. Chang, B. L. *et al*. Polymorphisms in the *CYP1A1* gene are associated with prostate cancer risk. *Intl J. Cancer* **106**, 375–378 (2003).
126. Lee, S. G. *et al*. Genetic polymorphisms of *XRCC1* and risk of gastric cancer. *Cancer Lett.* **187**, 53–60 (2002).
127. Petrij-Bosch, A. *et al*. *BRCA1* genomic deletions are major founder mutations in Dutch breast cancer patients. *Nature Genet.* **17**, 341–345 (1997).
128. Newman, B. *et al*. Frequency of breast cancer attributable to BRCA1 in a population-based series of American women. *JAMA* **279**, 915–921 (1998).
129. Wu, X. *et al*. p53 genotypes and haplotypes associated with lung cancer susceptibility and ethnicity. *J. Natl Cancer Inst.* **94**, 681–690 (2002).
130. Bonnen, P. E., Wang, P. J., Kimmel, M., Chakraborty, R. & Nelson, D. L. Haplotype and linkage disequilibrium architecture for human cancer-associated genes. *Genome Res.* **12**, 1846–1853 (2002).
131. Johnson, G. & Todd, J. Haplotype tagging for the identification of common disease genes. *Nature Genet.* **29**, 233–237 (2001).
132. Hoque, M. O., Lee, C. C., Cairns, P., Schoenberg, M. & Sidransky, D. Genome-wide genetic characterization of bladder cancer: a comparison of high-density single-nucleotide polymorphism arrays and PCR-based microsatellite analysis. *Cancer Res.* **63**, 2216–2222 (2003).
133. Janne, P. A. *et al*. High-resolution single-nucleotide polymorphism array and clustering analysis of loss of heterozygosity in human lung cancer cell lines. *Oncogene* **23**, 2716–2726 (2004).
134. Boshoff, C. *et al*. Kaposi's sarcoma-associated herpesvirus infects endothelial and spindle cells. *Nature Med.* **1**, 1274–1278 (1995).
135. To, E. W. H. *et al*. Rapid clearance of plasma Epstein-Barr virus DNA after surgical treatment of nasopharyngeal carcinoma. *Clin. Cancer Res.* **9**, 3254–3259 (2003).
136. Leung, S. F. *et al*. Pretherapy quantitative measurement of circulating Epstein-Barr virus DNA is predictive of posttherapy distant failure in patients with early-stage nasopharyngeal carcinoma of undifferentiated type. *Cancer* **98**, 288–291 (2003).
137. Mueller, N. Overview: viral agents and cancer. *Environ. Health Perspect.* **103** (Suppl. 8), 259–261 (1995).
138. Koutsky, L. A. *et al*. A controlled trial of a human papillomavirus type 16 vaccine. *N. Engl. J. Med.* **347**, 1645–1651 (2002).
139. Pellet, C. *et al*. Virologic and immunologic parameters that predict clinical response of AIDS-associated Kaposi's sarcoma to highly active antiretroviral therapy. *J. Invest. Dermatol.* **117**, 858–863 (2001).
140. Polyak, K. *et al*. Somatic mutations of the mitochondrial genome in human colorectal tumours. *Nature Genet.* **20**, 291–293 (1998).
141. Bianchi, N. O., Bianchi, M. S. & Richard, S. M. Mitochondrial genome instability in human cancers. *Mutat. Res.* **488**, 9–23 (2001).
142. Welter, C., Kovacs, G., Seitz, G. & Blin, N. Alteration of mitochondrial DNA in human oncocytomas. *Genes Chromosomes Cancer* **1**, 79–82 (1989).
143. Fliss, M. S. *et al*. Facile detection of mitochondrial DNA mutations in tumors and bodily fluids. *Science* **287**, 2017–2019 (2000).
144. Weinstein, J. N. Fishing expeditions. *Science* **282**, 628–629 (1998).
145. Weinstein, J. N. 'Omic' and hypothesis-driven research in the molecular pharmacology of cancer. *Curr. Opin. Pharmacol.* **2**, 361–365 (2002).

Acknowledgements

We thank Maria Chan (FDA) for guidance regarding Table 1 and Anna Maria Calcagno (NCI) for her constructive scientific input. This work was supported by the Intramural Research Program of the NIH, the National Cancer Institute and the Center for Cancer Research. We are grateful to Anna Barker and the reviewers whose constructive comments and suggestions significantly improved the manuscript.

Online links

DATABASES

The following terms in this article are linked online to:

Entrez Gene: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene

ABL | AFP | BCR | CGB | EGFR | ER | HER2/NEU | p53 | PR | PSA

National Cancer Institute: http://www.cancer.gov

breast cancer | colon cancer | HNSCC | leukaemia | NSCLC | lymphomas | ovarian cancer | prostate cancer | soft-tissue sarcomas | testicular cancer

FURTHER INFORMATION

Adjuvant Online: http://www.adjuvantonline.com/

Cancer Therapy Evaluation Program: http://ctep.cancer.gov/

Early Detection Research Network: http://www3.cancer.gov/prevention/cbrg/edrn/

National Institute of Standards and Technology: http://www.nist.gov/

NCI Genomics and Bioinformatics group: http://discover.nci.nih.gov

Program for the Assessment of Clinical Cancer Tests: http://www.cancerdiagnosis.nci.nih.gov/assessment/

TNM staging system: http://www.cancerstaging.org

Access to this interactive links box is free online.