Predictive and prognostic molecular markers for cancer medicine

Ther Adv Med Oncol
(2010) 2(2) 125–148
DOI: 10.1177/1758834009360519

Sunali Mehta, Andrew Shelling, Anita Muthukaruppan, Annette Lasham, Cherie Blenkiron, George Laking and Cristin Print

***Abstract*:** Over the last 10 years there has been an explosion of information about the molecular biology of cancer. A challenge in oncology is to translate this information into advances in patient care. While there are well-formed routes for translating new molecular information into drug therapy, the routes for translating new information into sensitive and specific diagnostic, prognostic and predictive tests are still being developed. Similarly, the science of using tumor molecular profiles to select clinical trial participants or to optimize therapy for individual patients is still in its infancy. This review will summarize the current technologies for predicting treatment response and prognosis in cancer medicine, and outline what the future may hold. It will also highlight the potential importance of methods that can integrate molecular, histopathological and clinical information into a synergistic understanding of tumor progression. While these possibilities are without doubt exciting, significant challenges remain if we are to implement them with a strong evidence base in a widely available and cost-effective manner.

***Keywords*:** biomarker, cancer, microarray, pathology, prognosis, treatment response

Introduction

Technological advances have greatly increased our understanding of the molecular basis of tumor progression and treatment response. Over the last 10 years these advances have led to the identification of numerous tumor biomarkers. These biomarkers can be divided into two types. Prognostic markers (see examples in Table 1) aim to objectively evaluate the patient's overall outcome, such as the probability of cancer recurrence after standard treatment. The presence or absence of a prognostic marker can be useful for the selection of patients for treatment but does not directly predict the response to a treatment. Predictive markers (see examples in Table 2) aim to objectively evaluate the likelihood of benefit from a specific clinical intervention, or the differential outcomes of two or more interventions, including toxicity. This is a rapidly accelerating research field that is beginning to have a significant clinical impact. Since the year 2000, there have been over 26,000 publications indexed in PubMed with the joint medical subject headings of 'neoplasm' and 'predictive marker', and almost 14,000 publications with 'neoplasm' and 'prognostic marker'. Increasingly, clinicians need to interpret molecular biomarkers and understanding the technologies that underlie them in order to make treatment decisions.

This review will summarize the current excitement in this field. It is organized around the molecular technologies used to generate DNA, epigenetic, RNA, signaling pathway, protein, and metabolic tumor biomarkers (summarized in Table 3). It will discuss the limitations and future potential of these biomarkers, and how they may be productively combined with clinicopathological data. This review will also summarize the challenges of tissue collection, the use of circulating tumor cells and metabolic imaging, and the expanding role of biomarkers in drug development. Finally, it will sound a cautionary note about the need to develop a stronger evidence base including robust clinical validation prior to commercializing predictive and prognostic markers for cancer medicine.

The growing role of DNA sequencing

DNA sequencing technologies will play a growing role in the implementation of cancer predictors and prognosticators. At present, DNA

Correspondence to:
Cristin Print
Department of Molecular Medicine & Pathology, School of Medical Sciences, University of Auckland, Private Bag 92019, Auckland, New Zealand
c.print@auckland.ac.nz

Sunali Mehta
Annette Lasham
Cherie Blenkiron
School of Medical Sciences, University of Auckland, Auckland, New Zealand

George Laking
School of Medical Sciences, University of Auckland, Private Bag 92019, Auckland, New Zealand, and Regional Cancer and Blood Service, Auckland District Health Board, Auckland, New Zealand

Andrew Shelling
Anita Muthukaruppan
Department of Obstetrics and Gynecology, School of Medicine, University of Auckland, Auckland, New Zealand

Table 1. Prognostic biomarkers for cancer medicine.

Prognostic biomarker	Type of cancer	Clinical significance	Detection	Clinical use	References
Beta-tubulin	NSCLC	High expression of beta-tubulin confers worse prognosis	IHC	No	[Seve *et al.* 2007; Winton *et al.* 2005]
BRCA1	Breast	High expression of BRCA1 confers worse prognosis in untreated patients	IHC	Yes	[James *et al.* 2007]
	NSCLC	High expression of BRCA1 confers worse prognosis in untreated patients	qRT-PCR	No	[Rosell *et al.* 2007]
CA19-9	Pancreatic	Higher preoperative CA19-9 levels are associated with lower resectability, more advanced stage and inferior survival	IHC	No	[Ferrone *et al.* 2006]
CAIX	RCC	High expression of CAIX is associated with a better prognosis	IHC	No	[Bui *et al.* 2003]
CD44	Bladder	Expression of CD44 is associated with poor prognosis	qRT-PCR	No	[Miyake *et al.* 2002]
CEA	CRC	Elevated preoperative CEA levels in resectable colorectal cancer is associated with poor prognosis	IHC	Yes	[Wolmark *et al.* 1984; Wanebo *et al.* 1978]
c-KIT	GIST	GIST patients have a better prognosis if they harbor a mutation in exon 11 of the c-KIT gene	Pathway detection via FDG-PET	Yes	[Singer *et al.* 2002]
ColoPrint	CRC	Prognosis for colorectal cancer patients	Microarray	Yes	[Glas *et al.* 2009]
CTC (e.g. CellSearch)	Melanoma	Increased number of circulating melanoma cells is associated with poor prognosis	Circulating tumor cells	No	[Nezos *et al.* 2009]
	CRC	Colorectal patients with ≥3 CTC/7.5 ml of peripheral blood were associated with shorter PFS and OS, i.e. poor prognosis	Circulating tumor cells	Yes	[Cohen *et al.* 2009]
	Breast	Breast cancer patients with ≥5 CTC/7.5 ml of peripheral blood are associated with shorter PFS and OS, i.e. poor prognosis	Circulating tumor cells	Yes	[Cristofanilli *et al.* 2004]
	Prostate	≥5 CTC/7.5 ml of peripheral blood is associated with poor prognosis	Circulating tumor cells	Yes	[Cristofanilli *et al.* 2004]
Cyclin D1	Bladder	Expression of Cyclin D1 is associated with low grade, low stage and recurrence	IHC	No	[Liukkonen *et al.* 2000]
Cyclin E	Bladder	Expression of Cyclin E is associated with low stage and survival	IHC	No	[Richter *et al.* 2000]
E-Cadherin	Bladder	E-Cadherin is associated with poor prognosis	IHC	No	[Bringuier *et al.* 1993]
EGFR	Bladder	Overexpression of EGFR is associated with high grade and high stage	IHC	No	[Nguyen *et al.* 1994]
	NSCLC	High gene copy number of EGFR in NSCLC patients is associated with poor prognosis	FISH / SA	No	[Coate *et al.* 2009]
	NSCLC	EGFR mutation in NSCLC patients is associated with better prognosis in untreated patients			
	Rectal	Overexpression of EGFR in rectal cancers is also associated with poor prognosis	IHC	No	[Kuremsky *et al.* 2009]
ER	Breast	Patients with ER-positive breast tumors have better survival than patients with hormonal negative tumors	IHC	Yes	[Early Breast Cancer Trialists' Collaborative Group, 1998]
eXagene*BC*	Breast	Provides prognosis in node-positive or node-negative breast cancer patients	FISH	Yes	[Davis *et al.* 2007]
Her2/neu	Breast	Patients with Her2/neu-positive breast tumors are more aggressive and have a worse prognosis compared to Her2/neu-negative tumors	FISH	Yes	[Mass *et al.* 2005]
	Bladder	Overexpression of Her2/neu is associated with high grade, high stage, poor survival and metastasis in bladder cancer	IHC	No	[Lipponen *et al.* 1991]
	GIST	Overexpression of Her2/neu in advanced gastric cancer patients is associated with poor prognosis	IHC	No	[Scartozzi *et al.* 2009]
Her3	Melanoma	Correlation with increased cell proliferation, tumor progression and reduced survival in melanoma patients	IHC	No	[Reschke *et al.* 2008]
ING3	Melanoma	Reduced nuclear expression associated with poor disease-specific survival in melanoma patients	IHC	No	[Wang *et al.* 2007]

Prognostic biomarker	Type of cancer	Clinical significance	Detection	Clinical use	References
ING4	Melanoma	Reduced levels of ING4 in melanoma patients is associated with melanoma thickness, ulceration and poor disease-specific survival and overall survival	IHC	No	[Li *et al.* 2008]
Ki-67	Bladder	Expression of Ki-67 is associated with progression and recurrence in bladder cancer	IHC	No	[Gerdes *et al.* 1984]
	Breast	Expression of Ki-67 is associated with proliferation and progression in breast cancer	IHC	No	[Dowsett *et al.* 2007]
K-ras	NSCLC	K-ras mutation is associated with poor prognosis in NSCLC patients	SA	Yes	[Zhu *et al.* 2008]
LOH at 18q	CRC	Associated with metastasis and poor prognosis in colorectal tumors	PCR	No	[Watanabe *et al.* 2001]
MammaPrint	Breast	A 70-gene prognostic assay used to identify breast cancer cases at the extreme end of the spectrum of disease outcome by identifying patients with good or very poor prognosis	Microarray	Yes	[Van 't Veer *et al.* 2002]
Mammostrat®	Breast	This standard purely prognostic test uses five antibodies with manual slide scoring to divide cases of ER-positive, lymph node negative breast cancer tumors treated with tamoxifen alone into low-, moderate- or high-risk groups	IHC	Yes	[Ring *et al.* 2006]
MMP-2	Bladder	Expression of MMP-2 is associated with poor prognosis in bladder cancer patients	PCR	No	[Xu *et al.* 2002]
MSI status	CRC	High frequency MSI colorectal tumors are associated with better prognosis and show improved relapse-free survival	IHC	No	[Kim *et al.* 2007]
NCOA3	Melanoma	Increased levels in melanoma patients correspond to poor relapse-free survival and disease-free survival	IHC	No	[Rangel *et al.* 2006]
Onco*type* DX	Breast	A 21-gene multiplex test used for prognosis to determine 10-year disease recurrence for ER-positive, lymph node negative breast cancers using a continuous variable algorithm and assigning a tripartite recurrence score	qRT-PCR	Yes	[Goldstein *et al.* 2008; Paik *et al.* 2004]
p21	Bladder	Overexpression of p21 is associated with poor prognosis	IHC	No	[Stein *et al.* 1998]
p53	Bladder	Overexpression of p53 is associated with poor prognosis	IHC	No	[Schmitz-Drager *et al.* 2000]
	NSCLC	High expression of p53 in NSCLC patients confers worse prognosis in untreated patients	IHC	No	[Tsao *et al.* 2007]
	NSCLC	TP53 mutation in NSCLC patients is associated with worse prognosis	SA	No	
PR	Breast	Patients with PR-positive breast tumors have better survival than patients with hormonal-negative tumors	IHC	Yes	[Dowsett *et al.* 2006]
Rb	Bladder	Overexpression of Rb is associated with poor prognosis	IHC	No	[Logothetis *et al.* 1992]
RRMI	NSCLC	High expression of RRMI in NSCLC patients confers better prognosis in untreated patients	AQUA	No	[Zheng *et al.* 2007]
VEGF	RCC	Overexpression of VEGF is associated with poor prognosis in clear cell renal carcinoma patients	IHC	Yes	[Oldenhuis *et al.* 2008]

AQUA, automated quantitative analysis; CA19-9, carbohydrate antigen 19-9; CAIX, carbonic anhydrase IX; CEA, carcinoembryonic antigen; CRC, colorectal tumor; CTC, circulating tumor cells; EGFR, epidermal growth factor receptor; ER, estrogen receptor; FDG, 18F-fluorodeoxyglucose; FISH, fluorescent *in situ* hybridization; GIST, gastrointestinal stromal tumor; IHC, immunohistochemistry; LOH, loss of heterozygosity; MMP-2, matrix metalloproteinase-2; MSI, microsatellite instability; NSCLC, non-small cell lung cancer; OS, overall survival; PET, Positron emission tomography; PFS, progression-free survival; PR, progesterone receptor; qRT-PCR, quantitative real time polymerase chain reaction; Rb, retinoblastoma; RCC, renal cell carcinoma; RRMI, ribonucleotide reductase messenger 1; SA, sequence analysis; VEGF, vascular endothelial growth factor.

Table 2. Predictive biomarkers for cancer medicine.

Predictive biomarker	Type of cancer	Clinical significance	Detection	Clinical use	Reference
BRCA1	NSCLC	High expression of BRCA1 in NSCLC patients predicts resistance to cisplatin-based chemotherapy	qRT-PCR	No	[Cobo *et al.* 2008]
	Breast	High expression of BRCA1 in breast cancer can predict response to chemotherapy	IHC	Yes	[James *et al.* 2007]
CAIX	RCC	Expression of CAIX in renal cell carcinoma is predictive of sensitivity of treatment with interleukin-2 therapy	IHC	No	[Tunuguntla and Jorda, 2008]
c-KIT	GIST	GIST patients carrying the mutation on exon 11 of the c-KIT gene benefit from imatinib and sunitinib treatment, however most patients develop resistance to these over time	SA	Yes	[Dematteo *et al.* 2009]
EGFR1	NSCLC	EGFR1 mutations in patients with NSCLC are predictive for response to either gefitinib or erlotinib treatment	IHC	Yes	[Sequist *et al.* 2007]
	CRC	EGFR1 gene amplification appears to be a predictive factor for response to anti-EGFR1 antibody treatment in CRC	PCR	Yes	[Amado *et al.* 2008]
ER	Breast	High cellular expression of ER predicts benefit from tamoxifen-based chemotherapy	IHC	Yes	[Paik *et al.* 2006; Early Breast Cancer Trialists' Collaborative Group, 1998]
ERCC1	NSCLC	High expression of ERCC1 in NSCLC patients predicts resistance to cisplatin-based chemotherapy	IHC	No	[Olaussen *et al.* 2006; Arriagada *et al.* 2004]
Her2/neu	Breast	Breast cancer patients with Her2/neu overexpressing tumors benefit from treatment with trastuzumab in the metastatic as well as in the adjuvant setting	FISH	Yes	[Mass *et al.* 2005]
	Gastric	Expression of Her-2/Neu in gastric cancer is predictive of patient sensitivity towards treatment with 5-FU, doxorubicin, trastuzumab and platinum-based chemotherapy	FISH	No	[Scartozzi *et al.* 2009]
K-ras	NSCLC	K-ras mutation positivity in NSCLC patients predicts lack of benefit from adjuvant chemotherapy in early disease and resistance to treatment with EGFR TKI in advanced disease	SA	Yes	[Mascaux *et al.* 2005]
	CRC	K-ras mutation positivity in stage IV CRC patients predicts considerably less benefit from EGFR-specific antibody like cetuximab and panitumumab	PCR	Yes	[Amado *et al.* 2008]
LOH at 18q	CRC	Useful in identifying patients with resected stage III colon cancer most likely to benefit from 5-FU based adjuvant chemotherapy	PCR	No	[Watanabe *et al.* 2001]
MGMT	Glioblastoma	Methylation of MGMT promoter is predictive of sensitivity of glioblastoma to temozolomide	PCR	No	[Dunn *et al.* 2009]
NuvoSelect	Breast	A combination of several pharmacogenomic genesets used primarily to guide selection of therapy in breast cancer patients. This test also provides the ER and HER2 mRNA status	Microarray	Yes	[Rouzier *et al.* 2005; Ayers *et al.* 2004]
p53	NSCLC	High p53 expression in NSCLC patients predicts sensitivity to cisplatin-based chemotherapy, however TP53 mutation is predicitve of resistance to cisplatin-based chemotherapy	IHC/SA	No	[Tsao *et al.* 2007; Winton *et al.* 2005]
PR	Breast	High cellular expression of PR predicts benefit from tamoxifen-based chemotherapy	IHC	Yes	[Oldenhuis *et al.* 2008; Dowsett *et al.* 2006; Elledge *et al.* 2000]
Roche AmpliChip	Breast	Low expression of CYP2D6 predicts resistance to tamoxifen-based chemotherapy in breast cancer patients	Microarray	Yes	[Hoskins *et al.* 2009]

Predictive biomarker	Type of cancer	Clinical significance	Detection	Clinical use	Reference
Rotterdam Signature	Breast	A 76-gene assay used to predict recurrence in ER-positive breast cancer patients treated with tamoxifen	Microarray	Yes	[Wang *et al.* 2005]
RRM1	NSCLC	High expression of RRM1 in NSCLC patients predict resistance to cisplatin-based chemotherapy	qRT-PCR	No	[Rosell *et al.* 2004]
TP	GIST	Predicitve of senstivity of treatment to 5-FU- and capcetabine-based chemotherapy in gastric cancer patients	IHC/PCR	No	[Scartozzi *et al.* 2009]
	CRC	Expression of TP in metastatic colorectal patients is predictive of sensitivity of treatment to 5-FU and capcetabine based chemotherapy	IHC/qRT-PCR	No	[Meropol *et al.* 2006; Metzger *et al.* 1998]
PTEN	Breast	PTEN mutation can result in reduced sensitivity of treatment with trastuzumab in breast cancer patients	IHC	No	[Nagata *et al.* 2004]

CAIX, carbonic anhydrase IX; CRC, colorectal tumor; EGFR, epidermal growth factor receptor; ER, estrogen receptor; ERCC1, excision repair cross-complementation group 1; FISH, fluorescent *in situ* hybridization; GIST, gastrointestinal stromal tumor; IHC, immunohistochemistry; LOH, loss of heterozygosity; MGMT, O6-methylguanine-DNA methyltransferase; NSCLC, non-small cell lung cancer; PCR, polymerase chain reaction; PR, progesterone receptor; RRMI, ribonucleotide reductase messenger 1; qRT-PCR, quantitative real-time polymerase chain reaction; RCC, renal cell carcinoma; SA, sequence analysis; TK1, tyrosine kinase inhibitor; TP, thymidine phosphorylase.

sequencing is more common in research than in the clinic, however it is likely to become more important for patient care over the next 5 years. While high costs may prevent whole-genome sequence analyses from becoming routine for all cancer patients, 'exome' sequencing (targeted sequencing of protein-coding regions [Ng *et al.* 2009]) may become a cost-effective way to provide access to the entire transcribed genome of individuals and their tumors. To date, only a small number of whole-genome tumor sequencing projects have been undertaken. For example, Ley and colleagues sequenced the tumor genome of a patient with acute myeloid leukemia, and identified more than 500 nonsynonymous somatic mutations, including 10 mutations within protein-coding genes [Ley *et al.* 2008]. Recently, the International Cancer Genome Consortium (ICGC) announced a project to sequence the full genomes of 25,000 tumor samples from 50 different cancer types, with matching samples from healthy individuals (http://www.icgc.org/), to accelerate research into the causes and control of cancer.

Recently, several technological advances have increased the utility of genotyping in cancer. 'Deep' DNA sequencing, also known as 'next generation' or 'massively parallel' sequencing, is rapidly taking over from traditional 'Sanger' DNA sequencing. Essentially, in next-generation sequencing, an entire genome can be fragmented and then sequenced in parallel. As the technology can only read short lengths of DNA (<250 bp), the genome must be covered several times to ensure reliability. These advances have led to a surge in sequencing-based studies, and the full genomes of over 50 people, including related individuals with rare disorders, are now either published or in preparation [Hayden, 2009a]. Three companies are currently using different technologies to provide these next generation DNA sequences: the 454 system from Life Sciences (Branford, Connecticut, USA), Solexa from Illumina (San Diego, California, USA), and SOLiD from Applied Biosystems (Foster City, California, USA). However, even faster and less expensive 'single-molecule' sequencing technologies are now arriving which will further revolutionize the molecular analysis of neoplasia [e.g. Bowers *et al.* 2009].

Tumor genomic heterogeneity is not well captured by simply sequencing tumor tissue, since the profiles of small regions within tumors, which contain additional mutations and may eventually overgrow the tumor, are diluted. However, new

Table 3. The range of tumor biomarker technologies used to generate cancer predictors and prognosticators, with references to examples discussed in this review.

Measurement target	Technology	Reference
Single genes	FISH	[Lebeau *et al.* 2001]
	CISH and SISH	[Penault-Llorca *et al.* 2009]
Sets of genes	FISH	[Davis *et al.* 2007]
	DNA sequencing	[Ley *et al.* 2008]
	DNA SNP analysis	[Tenesa and Dunlop, 2009; Easton *et al.* 2007]
Methylated gene promoters	Methylation-specific quantitative PCR	[Huang *et al.* 2009]
Pairs of mRNAs	Quantitative PCR	[Ma *et al.* 2004]
Sets of mRNAs	Quantitative PCR	[Paik *et al.* 2004 #2974]
	Microarrays	[Liu *et al.* 2007; Sotiriou *et al.* 2006; Wang *et al.* 2005; Van 't Veer *et al.* 2002]
	RNA-seq	[Shah *et al.* 2009]
	Microelectrodes	[Fang *et al.* 2009]
Tumor classification inferred from mRNAs	Microarrays	[Sorlie *et al.* 2001; Perou *et al.* 2000]
Tumor grade inferred from mRNAs	Microarrays	[Sotiriou *et al.* 2006a]
Tumor signaling pathways inferred from mRNAs	Microarrays	[Bild *et al.* 2006a; West *et al.* 2001]
Exon usage	Exon-specific micorarrays	[Andre *et al.* 2009]
Sets of miRNAs and ncRNAs	Quantitative PCR	[Zhu *et al.* 2009]
Single proteins	Immunohistochemistry	[Pertschuk *et al.* 1979]
Sets of proteins	Immunohistochemistry	[Ring *et al.* 2006]
Metabolites	Biochemical assays	[Sreekumar *et al.* 2009]
Tumor functional pathways	Image-based probes such as PET	[Contractor and Aboagye, 2009]
Clinical data	Decision support tools	[Ravdin *et al.* 2001]
Clinical and mRNA data	Integrative models	[Acharya *et al.* 2008; Cardoso *et al.* 2008]

CISH, chromogenic *in situ* hybridization; FISH, fluorescent *in situ* hybridization; miRNAs, micro-RNAs; ncRNAs, noncoding RNAs; PCR, polymerase chain reaction; PET, positron-emission tomography; SISH, silver-enhanced *in situ* hybridization; SNP, single nucleotide polymorphisms.

methods that independently sequence multiple representatives of a given DNA fragment [Thomas *et al.* 2006] allow detection of DNA sequence variations present in a minority of cells within heterogeneous tumors. These advances in profiling tumor heterogeneity are complimented by methods such as COLD Polymerase Chain Reaction (PCR) [Li and markrigiorgos, 2009]that selectively amplify minority alleles from mixtures of wild-type and mutation-containing sequences [Li *et al.* 2009], and by advances in RNA sequencing (RNA-seq) methods described below.

Defined predictive and prognostic genetic variants

The identification of DNA variants such as single nucleotide polymorphisms (SNPs) is becoming increasingly routine, including limited genotyping of tumor DNA and screening of somatic (non-tumor) DNA for mutations that predispose to cancer or alter treatment response. Some USA hospitals now undertake broad genetic testing of almost all patients with cancer. For example, Massachusetts General Hospital in Boston is screening for 110 mutations in 13 cancer-related genes [Hayden, 2009b]. Genetic variants are also used to predict toxicity, for example polymorphisms in the *UGT1A1* gene can be used to predict irinotecan toxicity [Palomaki *et al.* 2009]. Tumor genotyping is also offered privately; the Californian company CollabRx offers a service that analyses 15,000 genes (http://collabrx.com/).

Genome-wide association (GWA) studies are feeding into the clinical use of DNA variants. For example, GWA studies have identified cancer-causing mutations in breast [Easton *et al.* 2007] and colon [Tenesa and Dunlop, 2009] tumors. Somatic genetic screens are also identifying predictors of radiation sensitivity [Barnett *et al.* 2009] and the pharmacodynamics of anticancer drugs [Sawyers, 2008]. This research has been made possible by technological developments that have expanded the throughput and reduced the cost of SNP-based research, and allowed SNP-based techniques to identify DNA copy number variation. Nevertheless, SNP-based research to identify oncogenic DNA

abnormalities remains a significant challenge, due to the difficulty of separating these cancer-causing abnormalities from genetic and epigenetic 'noise' [Chin and Gray, 2008]. One of the limitations of GWA studies is that we are capturing only common SNPs, which by themselves may only contribute a small amount of risk to developing cancer. To identify those rare SNPs that contribute a large amount of risk to developing cancer for a small percentage of patients, more cost-effective technologies will be required [Shelling, 2009].

Fluorescent in situ hybridization (FISH)

FISH provides an alternative way to identify predictive or prognostically important genetic variants in cancer. The most common use of FISH has been to detect copy number changes that predict treatment response, for example changes to the c-*erb B2/neu* (*ERBB2*) gene that encodes the HER2 protein [Lebeau *et al.* 2001; Ross, 2009b]. However, the use of FISH is now expanding. A FISH assay based on the copy numbers of three genes has been successfully used to predict prognosis in breast cancer. This test uses a proprietary algorithm to integrate the information from the three genes and predict recurrence rates. Interestingly, different three-gene sets are required for hormone-receptor positive and hormone-receptor negative tumors [Davis *et al.* 2007]. This is commercialized as the eXagen BC assay. Improvements have been made to FISH in the form of chromogenic *in situ* hybridization (CISH) and silver-enhanced *in situ* hybridization (SISH). These techniques use peroxidase enzyme-labelled probes whose signals do not decay over time and allow the specimen to be viewed using bright-field microscopy. CISH and SISH have been used to assess *ERBB2* gene status [Penault-Llorca *et al.* 2009].

Gene promoter modification profiles

Epigenetics can be defined as the field of inheritable changes in gene expression that are not caused by alterations in DNA sequence. A key mechanism of epigenetics is the altered methylation of tumor-suppressor genes and of the genes encoding some micro-RNAs (miRNAs), as well as altered methylation and acetylation of the histones associated with these genes [reviewed in Esteller, 2008]. Epigenetic alterations early in tumor development may provide important predictive and prognostic tools, especially in situations where epigenetic therapies such as HDAC inhibitors are being used to reactivate gene expression [Kanai, 2007]. For example, in breast cancer, several genes associated with tumorigenesis are frequently methylated, including: *RASSF1A, HOXA5, TWIST1, CCND2, p16, BRCA1,* as well as genes encoding the estrogen receptor (*ESR1*) and the progesterone receptor (*PGR*) [Dworkin *et al.* 2009]. Epigenetic changes in several other tumor types may also provide prognostic and predictive profiles, including: ovarian cancer [Huang *et al.* 2009], prostate cancer [Bastian *et al.* 2004], glioblastoma [Nagarajan and Costello, 2009] and cutaneous tumors [Li *et al.* 2009]. This is an exciting field that is likely to grow in clinical importance. It may even be possible to identify promoter methylation of tumour cell DNA in body fluids such as sputum for early cancer detection [Belinsky, 2004].

An overview of predictive and prognostic mRNA profiles

A large number of profiles based on the abundance of mRNAs have been put forward as predictors of prognosis or treatment response in cancer. The simplest of these profiles are based on ratios of two mRNAs, such as *HOXB13 versus IL17BR*. These ratios appear to be more effective than the single mRNAs for predicting prognosis or therapeutic response [Ma *et al.* 2004]. The *HOXB13* / *IL17BR* ratio has been commercialized as the Theros H/I assay. Whole-transcriptome microarray analysis of tumors has allowed predictive and prognostic molecular profiles to be generated that utilize tens or hundreds of genes. These large mRNA profiles have been particularly successful in breast cancer [Andre and Pusztai, 2006; Chang *et al.* 2005; Weigelt *et al.* 2005; Ayers *et al.* 2004; Paik *et al.* 2004, 2006; Van De Vijver *et al.* 2002; Van 't Veer *et al.* 2002] and colorectal cancer [Anjomshoaa *et al.* 2008; Garman *et al.* 2008; Lin *et al.* 2007a, 2007b; Alvarado *et al.* 2006; Gordon *et al.* 2006; Vallbohmer *et al.* 2006; Ahmed, 2005; Shih *et al.* 2005; Arango *et al.* 2004; Li *et al.* 2004; Huerta *et al.* 2003; Mariadason *et al.* 2003; Galon *et al.* 2002; Hegde *et al.* 2001]. However, they have also been applied to several other tumor types [Mengual *et al.* 2009; Wuttig *et al.* 2009; Bloomston *et al.* 2007; Jaeger *et al.* 2007; Mandruzzato *et al.* 2006; Winnepenninckx *et al.* 2006] including tumors of unknown origin [Van Laar *et al.* 2009]. Profiles have also been generated that can predict the progression of several different tumor types [Basil *et al.* 2006].

Examples of mRNA profiles used to predict prognosis and treatment response in breast cancer

As discussed in the introduction, there are a large and rapidly growing number of publications about predictive and prognostic profiles in cancer. A high proportion of these publications focuses on mRNA-based-profiles. It is beyond the scope of this review to describe all of these mRNA-based profiles, however, an illustrative selection is described below. The 21-mRNA Onco*type* DX quantitative PCR assay was generated from a 25,000-gene microarray study and is used to calculate recurrence scores in breast cancer [Goldstein *et al.* 2008; Paik *et al.* 2004]. This assay estimates the 10-year recurrence risk in patients with ER^{+ve}, lymph $node^{-ve}$ breast tumors, and has been especially useful for identifying patients who have a low risk of recurrence. The mRNAs selected for this assay place most weight on the ER and HER2 signaling pathways. The ability of Onco*type* DX to guide treatment selection is formally being evaluated in the TAILORx clinical trial [Zujewski and Kamin, 2008].

The 70-mRNA MammaPrint profile [Van 't Veer *et al.* 2002] is, unlike Onco*type* DX, a microarray-based test. Although this test was originally criticized for the inclusion of some patients in both training and test groups, it has since been clinically validated to a high standard and has US Food and Drug Administration (FDA) approval. MammaPrint is currently used in the prospective MINDACT clinical trial, in which its molecular-based predictions are combined with predictions based on clinicopathological knowledge through the Adjuvant! Online decision support tool (see below) [Cardoso *et al.* 2008].

The Rotterdam profile [Wang *et al.* 2005] uses a 76-mRNA signature weighted towards proliferation-associated signaling pathways, and has no intersection with mRNAs used in either the Onco*type* DX or MammaPrint tests. Another example is the 186-mRNA invasiveness prognostic signature based on the mRNA expression characteristics of $CD44^{high}/CD24^{low}$ tumor cells [Liu *et al.* 2007]. The development of these and other predictive and prognostic mRNA profiles in breast cancer have been summarized in a number of excellent reviews [e.g. Bukhari and Akhtar, 2009; Geyer and Reis-Filho, 2009; Kim *et al.* 2009; Ross, 2009a; Sotiriou and Pusztai, 2009; Rakha *et al.* 2008; Ross *et al.* 2008].

mRNA profiles that provide classification and grading information

In breast cancer, sets of genes identified through mRNA profiling have been used to generate a 'molecular portrait' of tumors and have provided an informative classification into five oncogenic subtypes [Kapp *et al.* 2006; Sorlie *et al.* 2001; Perou *et al.* 2000]. For example, patients with 'basal-like' and 'HER2' breast tumor subtypes are more likely to have pathologic complete response after neoadjuvant multi-agent therapy [Rouzier *et al.* 2005]. A 50-mRNA classifier to determine these subtypes is marketed as the Breast BioClassifier (http://www.bioclassifier.com/). A set of 97 mRNAs identified through transcriptome profiling have also allowed the molecular grading of breast cancer [Sotiriou *et al.* 2006], which is currently being developed commercially as the MapQuant D× Genomic Grade assay.

Signaling pathway information derived from mRNA profiles

It is widely accepted that cell fate and function are regulated at the level of signaling networks and pathways rather than at the level of individual molecules [Thiagalingam, 2006]. For this reason, clinically useful molecular profiles often utilize information about the activity of entire signaling pathways [Bild *et al.* 2006a]. Examples include identification of *BRAF*-mutant tumors that will respond to inhibitors of MEK (an enzyme in a signaling pathway downstream of BRAF) [Solit *et al.* 2006], and the identification of *PTEN* deficient tumors likely to be resistant to trastuzumab (PTEN regulates the PI3 kinase signaling that occurs downstream of the epidermal growth factor receptor [EGFR] that is targeted by trastuzumab) [Nagata *et al.* 2004].

In breast cancer, mRNA profiles that are surrogate markers for the activity of a number of key signaling pathways have been developed [Chang *et al.* 2009; West *et al.* 2001]. Some of these mRNA profiles appear to feed directly into activities of tumorigenesis, such as E2F-associated pathways feeding into tumor cell proliferation [Hallstrom *et al.* 2008]. Interestingly, the activity of several specific oncogenic pathways in breast cancer appears to be associated with patient age [Anders *et al.* 2008]. Information about deregulation of these pathways has been useful both in breast tumor classification and in predicting

clinical sensitivity to therapeutic agents [Salter *et al.* 2008; Bild *et al.* 2006b; Huang *et al.* 2003]. The fact that these pathway-level molecular profiles are based on known biology is very attractive.

Technologies to generate predictive or prognostic profiles from mRNA

Profiles involving small numbers of mRNAs can utilize quantitative PCR, which has high sensitivity and dynamic range and has become a mainstay of molecular profiling in cancer. PCR is used for several commercialized mRNA signatures, for example [Ma *et al.* 2004; Paik *et al.* 2004]. However, like all molecular techniques, quantitative PCR does have its limitations, as reviewed by Bustin and colleagues, such as dependence on template quality [Bustin and Nolan, 2004]. Microarrays can capture parallel information about many more mRNAs than RT-PCR—up to whole transcriptome levels. Although microarrays are expensive, susceptible to false-positive error and require complex normalization and interpretation, they have been highly productive research tools and are also used in several current prognostic and predictive tests. The clinical value of mRNA microarrays is likely to continue to improve due to concerted efforts to improve data quality [Shi *et al.* 2006], although they are likely to be replaced in the next 10 years by the emerging technologies described below.

Profiles based on simple mRNA quantification only capture part of the available information. More than 60% of human mRNAs undergo alternative splicing [Clark *et al.* 2007] yielding hundreds of thousands of mRNA transcript variants that may have distinct functions. This alternative splicing appears to be especially relevant to cancer [Venables, 2006; Faustino and Cooper, 2003]. Recently, microarrays that can interrogate over one million human exons have become available, with approximately four measurements per exon and on average 40 measurements per gene [Abdueva *et al.* 2007]. Alternative mRNA splicing data generated using this technology may be useful for predicting tumor prognosis and treatment response. For example, exon-level microarray analysis has revealed mRNAs that are differentially spliced in breast [Andre *et al.* 2009], colorectal [Gardina *et al.* 2006] and lung [Xi *et al.* 2008] tumors, which may contribute to tumor progression.

Next-generation DNA sequencing technology has expanded into the realm of RNA, in which case it is known as 'transcriptome sequencing' or 'RNA-seq'. RNA-seq allows simultaneous analysis of all RNA molecules within a cancer cell, including alternative splice variants, mRNAs, noncoding RNAs (ncRNAs) and miRNAs [Wang *et al.* 2009b]. It is expected that in the future, RNA-seq will revolutionalize the analysis of RNAs in tumors, leading to several editorials suggesting that this is 'the beginning of the end for microarrays' [Shendure, 2008]. Due to issues with cost and analysis challenges, the replacement of microarrays with RNA-seq may still be some years away, and is unlikely to be complete. One recent example however, shows the remarkable ability of RNA-seq to analyze tumor RNA. Shah and colleagues sequenced the transcriptome of four biopsies from granulosal cell ovarian tumors, and identified a single gene, *FOXL2*, that was mutated in all tumors, but not in other ovarian tumors [Shah *et al.* 2009]. Furthermore, RNA-seq has even been used to characterize single tumor cells [Tang *et al.* 2009]. This, along with recently developed microfluidic devices capable of measuring the mRNA expression in single cells and the COLD PCR method [Li and markrigiorgos, 2009] discussed above, may represent a significant step towards quantifying tumor heterogeneity during prognostication [Toriello *et al.* 2008].

Nanotechnology is another area likely to revolutionize the use of predictive cancer markers in the next 10 years. For example, nucleic acid probes immobilized on arrays of nanostructured microelectrodes within integrated circuits are already a reality [Soleymani *et al.* 2009]. These systems have been used to successfully detect gene fusions in prostate cancer [Fang *et al.* 2009]. These technologies are in commercial development and some commentators suggest they will make PCR-free nucleic acid-based treatment response prediction available to primary care medical practitioners.

Technologies to select mRNAs for inclusion in tumor profiles

An advantage of predictive and prognostic profiles that utilize large numbers of mRNAs is their inherent redundancy—failure of one or more measurements may be compensated for by the other measured mRNAs. However, the optimal size for a clinically useful set of predictive mRNAs remains a matter for debate. It is likely that future advances will take integrative tumor

pathology even further than the pathway-level understanding described above. As Karsdal and colleagues suggest, although our current understanding has progressed to the level of molecular pathways, further benefits are likely if we can extend our knowledge to a 'systems-level' understanding of the pathophysiology of specific diseases [Karsdal *et al.* 2009]. Further mathematical algorithm developments for selecting optimal mRNA combinations for different predictive and prognostic tasks will be important. Currently, methods used range from simple classification systems to highly sophisticated machine learning methods [Sotiriou and Piccart, 2007; Saidi *et al.* 2004]. The ever-expanding systems biology databases such as the Ingenuity Pathway Analysis database (http://www.ingenuity.com/) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) [Kanehisa *et al.* 2006] may also be useful for the development of future pathway-specific molecular profiles. However, all of these methods require more formal evaluation than has been performed to date.

Limitations of mRNA-based predictive and prognostic molecular profiles

Despite the significant body of literature describing predictive or prognostic mRNA profiles for cancer, only a small number are used in current oncology practice; examples are shown in Tables 1 and 2. In fact, as few as 3% of published studies describing potential clinical applications in genomic medicine have progressed to a formal assessment of clinical utility [Khoury *et al.* 2007]. It seems that mRNA-based molecular prognostic or predictive tests for cancer have not yet achieved their full potential due to a variety of limitations including: bias [Ransohoff, 2005], poor ability to generalize [Reid *et al.* 2005], instability of gene lists [Koscielny, 2008], inadequate sample sizes [Sorlie *et al.* 2006], inadequate statistical analysis [Faratian and Bartlett, 2008] and insufficient numbers of studies testing their therapeutic utility [Tavassoli, 2009; Simon *et al.* 2003]. Other limitations include predominantly retrospective and observational study designs [Tinker *et al.* 2006], poor representation of small early-stage tumors, and in breast cancer specifically, a limited predictive spectrum beyond ER^{+ve} tumors [Geyer and Reis-Filho, 2009]. An additional criticism is that many mRNA-based predictors have grown out of academic projects rather than being designed 'from the ground up' to support key therapeutic decisions [Simon, 2005]. Kim and colleagues suggest that "... clinically useful prognostic and predictive markers are those developed with a specific clinical context in mind and tested and validated within that clinical context" [Kim *et al.* 2009]. Views about the use of molecular profiling in oncology have become polarized. As one commentator has observed, "The field of microarray expression profiling is burdened with both unrealistic hype and excessive skepticism" [Simon, 2005].

MicroRNAs and noncoding RNAs

The analysis of RNA has recently expanded into the field of ncRNA. As their name implies, ncRNA molecules do not encode protein products, nevertheless they are important regulators of gene expression, and are able to directly affect the phenotypes of cells. There are several types of ncRNAs, of which miRNAs are the most widely studied and characterized. miRNAs have become a hot topic for researchers in every field since their discovery in *Caenorhabditis elegans* in 1993 [Lee *et al.* 1993]. A link between them and cancer was sealed with the report of loss of two miRNA genes at 13q14 in chronic lymphocytic leukemia [Calin *et al.* 2002]. Since this report, the expression of many individual miRNAs has been associated with patient survival, drug treatment response and tumor metastases in a number of different cancers [Sotiropoulou *et al.* 2009; Blenkiron and Miska, 2007]. For example, in breast cancer, specific miRNAs are associated with the five mRNA-derived tumor subtypes [Blenkiron and miska, 2007]. Particular miRNAs may also be signatures of early carcinogenesis [Kalscheuer *et al.* 2008]. miRNAs are very stable in tumors, even in archival specimens, and in body fluids such as blood and serum [Chen *et al.* 2008; Mitchell *et al.* 2008]. Specific miRNAs have been sequenced from the circulating blood of patients with lung cancer and colorectal cancer [Chen *et al.* 2008], and for example, miR-155 may be differentially expressed in the serum of women with PR^{-ve} compared to women with PR^{+ve} breast cancer [Zhu *et al.* 2009].

In contrast to miRNAs, much less is known about the other types of ncRNAs, with novel ncRNAs continuing to be discovered [Guffanti *et al.* 2009]. Many ncRNAs also appear to be differentially expressed in specific tumor types (reviewed in [Mallardo *et al.* 2008]), and with the advent of RNA-seq, the possibility to discover as yet unknown but potentially important

ncRNAs is high. All of this raises the exciting prospect that ncRNAs may in the future provide clinically-useful biomarkers for cancer prognostication [Bartels and Tsongalis, 2009].

Proteins

Immunohistochemistry (IHC) of tumor tissue has been a cornerstone of protein-based tumor marker work for several decades [Pertschuk *et al.* 1979]. A full description of IHC in cancer is outside the scope of this review; however, we will briefly summarize some topical points. IHC has the advantages of providing morphological information about protein expression and is of comparatively low cost, but has the disadvantages of sensitivity to tissue processing, scoring variability [Ross *et al.* 2007] and a limited availability of validated markers. One of the most interesting developments is the use of multiple antibodies to generate synergistic data for several proteins. For example, IHC data for the proteins p53, NDRG1, CEACAM5, SLC7A5, and HTF9C appears to be superior to data from currently-used single markers for predicting outcome in ER^{+ve} breast cancer [Ring *et al.* 2006]; this is now commercialized as the Mammostrat assay.

The area of proteomics has long promised biomarkers for the early detection of cancer. For example, phosphospecific antibodies can be used to identify specific kinase activity [Cloughesy *et al.* 2008], and a large-scale survey of tyrosine kinase activity in lung cancer has identified several novel molecular features of lung cancer that are invisible at the DNA or RNA level [Rikova *et al.* 2007]. Low molecular weight circulating blood proteins—known as the 'peptidome'—may also provide potential prognostic markers for early tumors [Petricoin *et al.* 2006]. Microvesicles shed by tumors carry molecular signatures such as variant *EGFR* mRNAs, and may provide additional sources of predictive and prognostic information [Skog *et al.* 2008].

Unfortunately, several research groups have reported potential protein/proteome-based biomarkers, only to find later that they have not reached clinical utility after greater levels of scrutiny. Recently, the prostate-specific antigen (PSA) has been shown in two large clinical trials to be largely ineffective for early detection of prostate cancer, and to lead to over diagnosis and overtreatment [Barry, 2009]. Robust clinical evaluation is likely to be needed before the potential of proteomic biomarkers is fully realized.

Metabolite profiling

Metabolomics, which quantifies the metabolite content of cells or tissues, is also potentially useful for predicting treatment response, since metabolites represent the destination or endpoint of many molecular pathways. The recent discovery that sarcosine, a derivative of glycine, is elevated in the urine of men with metastatic prostate cancer [Sreekumar *et al.* 2009] raises the hope that it will be of benefit for the detection of aggressive cancer. Developments in imaging modalities (see below) that allow detection of tumor metabolites are likely to accelerate this field [Brindle, 2008].

Molecular imaging

Tumor biology and pharmacology is increasingly amenable to *in vivo* evaluation using image-based molecular probes. Positron-emission tomography (PET) offers some of the most promising examples in this area. The most commonly used PET tracer, ^{18}F-fluorodeoxyglucose (FDG), is a marker of glucose transport in gastrointestinal stromal tumors (GIST) [Contractor and Aboagye, 2009]. There are many other image-based biomarkers relevant to the biology of living tumor cells, at various stages of laboratory and clinical development. These are showing promise in several applications, including the detection of occult tumors. Although a full discussion is beyond the scope of this paper, this area is reviewed in several recent papers [Josephs *et al.* 2009; Thakur, 2009; Hargreaves, 2008].

Decision support tools

As part of clinical acumen, physicians have traditionally integrated clinical and pathological data of several different types to make treatment decisions. However, as the amount of information available grows, there comes a stage when there is too much information to manually integrate reliably; leading to the need for formal mathematical or computational models of tumors that incorporate this information [Lazebnik, 2002]. For this reason, clinicopathological data have been integrated using various types of mathematical algorithms to provide decision support tools for clinicians [Abbott and Michor, 2006]. Decision support models are available for several tumor types, for example: urological tumors [Donovan *et al.* 2009; Abbod *et al.* 2007],

HER2^{+ve} breast tumors [Lisboa *et al.* 2007], for decisions about adjuvant therapy (Adjuvant! Online, http://adjuvantonline.com/index.jsp) [Ravdin *et al.* 2001], and for selecting optimal strategies to manage colorectal liver metastases [Poston *et al.* 2005]. One group has used the term 'systems pathology' to describe their method of using pathological data with clinical data and limited molecular data for prognostication after prostate cancer surgery [Cordon-Cardo *et al.* 2007].

Combining clinical and molecular cancer predictors

A principle of the field of systems biology is to combine multiple data types in order to reduce false-positive errors [Ge *et al.* 2003]. It is attractive to apply this principle to prognostic or predictive profiles in cancer, where different types of molecular and/or clinical data may be synergistic, and where separately-generated molecular profiles may concord to add weight to predictions [Fan *et al.* 2006]. For example, in cases where quantitative PCR reports a low average abundance of a predictive mRNA in a tumor, it may be useful to know from IHC that the protein encoded by that mRNA is expressed at high levels, but only by cells in a small dysplastic region of an otherwise low-grade tumor. In another example, molecular grade signatures may assist pathologists when grading tumors that have not been optimally fixed [Sotiriou *et al.* 2006].

Shedden and colleagues have shown that molecular tumor classification can be made accurately using fewer mRNA measurements if information about pathology and tissue ontogeny is also included [Shedden *et al.* 2003]. Pittman and colleagues have combined pathway-level summaries of mRNA abundance with traditional clinical risk factors to predict breast cancer recurrence [Pittman *et al.* 2004]. Salter and colleagues have integrated ER and HER2 status, mRNA profiles and miRNA profiles for the NCI-60 cell lines to predict drug sensitivity [Salter *et al.* 2008]. Acharya and colleagues have incorporated pathway profiles from mRNA expression data with clinical risk stratification to derive prognostic 'clinicogenomic clusters' [Acharya *et al.* 2008]. In the MINDACT breast cancer clinical trial, mRNA profiles from the MammaPrint microarray assay are used alongside clincopathological knowledge contained in Adjuvant! Online, using a relatively simple voting system [Cardoso *et al.* 2008].

The problem of data volume

As tumor molecular profiling becomes more routine in clinical practice, the large amounts of data associated with each patient needs to be annotated, individual items of information appropriately linked to the other items of information, and this information stored in a secure but easily retrievable format [Wang *et al.* 2009a]. Layered on top of these challenges is the frequent involvement of patients in clinical trials, and the need for tumor molecular profiles to be available for research and regulatory purposes, with access restriction based on patient consent and ethical approval providing a further layer of complexity. In a 2003 report, the National Institutes of Health highlighted the balance between ethical use/privacy and easy access to "make data as widely and freely available as possible, while safeguarding the privacy of participants, and protecting confidential and proprietary data" [NIH, 2003]. Sharing data as freely as possible within ethical, commercial and regulatory constraints drives development of new hypotheses and fosters cross-validation, as exemplified by archives of gene expression data such as the Gene Expression Omnibus (GEO) [Edgar, 2002] and ArrayExpress [Parkinson *et al.* 2007]. Several bioinformatic initiatives such as the US National Cancer Institute's cancer Biomedical Informatics Grid (caBIG) aim to link pathological, molecular, clinical and treatment response data. In the future these large initiatives, rather than single-center research studies, are likely to become the dominant platform for generating and assessing predictors and prognosticators in cancer medicine [McCormick, 2009; Hanauer *et al.* 2007]. A current research challenge is how to integrate diverse types of clinical and molecular information from thousands of patients. An approach that appears to have promise is probabilistic Bayesian modeling, which has been used successfully to integrate microarray data from multiple studies [Sykacek *et al.* 2007] and to integrate clinical and molecular data in breast cancer [Gevaert *et al.* 2006].

The challenge of accessing tumor tissue

The uptake of molecular predictive tests in solid tumors has been slower than in leukemia, possibly due to the more difficult access to solid tumor cells [Sawyers, 2008]. In trials involving solid tumors, often the only source of tissue is from a

single diagnostic biopsy or after resection. Although not always practical, studies involving re-biopsies have been dramatically successful—for example the identification of lung cancer patients who have acquired *EGFR* kinase domain mutations during gefitinib treatment [Pao *et al.* 2005]. Studies comparing molecular profiles derived from biopsies to those from resected tumors have proved interesting—it appears that although profiles from biopsies generally overlap with those from resected tumors, they differ particularly in those molecules expressed by the tumor stroma, which may be sampled differently by each method [Symmans *et al.* 2003]. It is encouraging that even the small amounts of tumor tissue obtained by fine needle aspiration appear capable of predicting treatment response [Ayers *et al.* 2004]. This is also the attraction of noninvasive image-based testing paradigms. Due to improved diagnostic and screening procedures, some tumors are now being diagnosed as very small lesions. Many of the tumor prognostic profiles in current use originated in studies of large late-stage tumors, and therefore, even if they are clinically effective, they may not be optimal for predicting the therapeutic response of the small early-stage tumors identified in screening programs. An additional complication, even in small tumors, is tumor heterogeneity [Fey and Tobler, 1996]. To overcome this, careful selection of multiple representative tumor regions may be required [Ross *et al.* 2007; Paik *et al.* 2005; Carey *et al.* 1990].

Due to nucleic acid degradation and cross-linking, DNA or RNA extracted from formalin-fixed and paraffin-embedded (FFPE) tumor tissue tends to be of lower quality than DNA or RNA extracted from unfixed tissue [Paik *et al.* 2005]. Those assays that can analyze predictive or prognostic profiles from FFPE tumor tissue (such as Onco*type* DX) may have a practical advantage over those that cannot (such as the current version of MammaPrint). There has been a gradual development of more effective protocols for recovery of DNA and RNA from FFPE tissue, and some studies report that mRNA expression profiles from FFPE tissues and fresh frozen tissues are similar [Linton *et al.* 2009; Furusato *et al.* 2007]. However, progress is still needed towards better recovery of DNA and RNA from FFPE tissue and towards better fixation processes. These developments will not only benefit predictive clinical assays, but will also benefit cancer research by unlocking a treasure trove of FFPE archival material.

Circulating tumor cells

The identification of circulating tumor cells in peripheral blood has been recognized for many years as a promising diagnostic tool for the early detection of primary tumors and metastatic deposits [Mostert *et al.* 2009; Sleijfer *et al.* 2007]. In addition, a sufficiently sensitive assay based on circulating tumor cells could be valuable for early detection of recurrent disease. Circulating tumor cells have been found in a wide range of cancers, including breast [Ross and Slodkowska, 2009], colorectal cancer [Sergeant *et al.* 2008] and melanoma [Nezos *et al.* 2009]. Reliable detection of circulating tumor cells is difficult, since they form only a very small proportion of the heterogeneous range of cells present in the peripheral blood. However, recent studies have demonstrated clinical utility. For example, the CellSearch System (Veridex, North Raritan, New Jersey, USA) has FDA approval and has been used in several clinical settings including monitoring treatment response in prostate, colon and breast cancer [Nakamura *et al.* 2009].

Regulation, validation and cost effectiveness

Several predictive and prognostic tests have now been commercialized for decision-making in oncology [e.g. Ross *et al.* 2008; Weigelt *et al.* 2005; Paik *et al.* 2004; Van 't Veer *et al.* 2002]. The experience with cetuximab provides an example of the commercial value inherent in such testing. Cetuximab has been developed as a treatment for advanced colorectal cancer by the firms ImClone, Bristol-Myers Squibb and Merck. It was evaluated in randomized trials involving several hundreds of patients [Bokemeyer *et al.* 2009; Van Cutsem *et al.*, 2009]. During the course of these trials, Lièvre and colleagues published a study of 30 patients that found cetuximab had no effect in tumors carrying mutant forms of the proto-oncogene K-ras [Lievre *et al.* 2006]. This finding was confirmed in the larger studies, and as a consequence the market for cetuximab shrank by around 35%, the frequency of K-ras mutant tumors. This finding clearly means that anyone contemplating purchase of cetuximab to treat bowel cancer will place a high value on K-ras mutational analysis. Such predictive tests will increasingly come to represent marketable intellectual property.

The regulation of predictive and prognostic molecular tests for cancer is complex. In the USA, there is a dual path system where the majority of molecular prognostic tests (which

are developed in-house by testing laboratories as 'home brew' tests) are not examined by the FDA, provided technical aspects of the test are certified. In contrast, those tests marketed as 'kits' must be fully reviewed by the FDA as medical devices [Schmidt, 2008]. Several commentators have argued that this produces a situation where many molecular tests are not supported by adequate evidence before they are marketed [Katsanis *et al.* 2008]. For example, a recent study of HER2 testing in breast cancer patients to predict trastuzumab response suggested that there may be concerning variations in practice and a significant incidence of inadequate documentation, as well as potentially inaccurate results [Phillips *et al.* 2009]. However, this view is balanced by the fact that verifying the analytical and clinical validity of cancer biomarkers is a complicated process [Schmidt, 2008; Gutman and Kessler, 2006], and that regulatory processes can greatly slow the translation of important medical advances into patient care, as well as greatly increase costs [Rogowski *et al.* 2009; Pitts, 2008]. It is also interesting to consider the ethical issues of developing predictive tests that, due to cost, may not be available to the majority of the world's population, and may not even be cost effective for affluent populations [Ioannidis, 2009].

Another issue is that the calibration of a test will influence its cost effectiveness. To the extent that a test is amenable to receiver-operator analysis (analysis of the balance between sensitivity and specificity), designation of the threshold for positive *versus* negative results is important. Lenient thresholds will accept lower likelihoods as predicting positive benefit from a treatment, and tend to increase access. Stricter thresholds will demand a higher likelihood before predicting positive benefit, and tend to reduce access [Laking *et al.* 2006 #126]. Inevitably, proponents of treatments, such as the pharmaceutical industry, will have an interest in the design of tests that predict benefit from their agent. Tests designed for use in wealthier settings may not be well-calibrated for poorer settings, in that they may support less-cost-effective allocations of treatment.

Overall, it appears that the development of appropriate regulatory structures, a robust evidence base [Faratian and Bartlett, 2008] and physician preparedness for molecular prognostic tests in cancer have lagged behind the development of the tests themselves [Avard and Knopper, 2009; GenomeWeb, 2009; Scheuner *et al.* 2008]. It is encouraging to see the emergence of large clinical trials for some breast cancer predictive tests; the Onco*type* DX test (TAILORx, [Zujewski and Kamin, 2008]) and the MammaPrint test (MINDACT, [Cardoso *et al.* 2008]). However, the design of these trials may not allow direct conclusions to be drawn about the clinical usefulness of these tests [Koscielny, 2008]. As more cancer biomarkers are commercialized, there needs to be increasing focus on providing an evidence base for improved clinical outcome using prospective randomized clinical trials, or if possible, using clinical trials in which randomization for specific treatment modalities is based on a biomarker. In addition, to allow for the optimal use of these assays, a wide range of clinical settings (such as locally advanced or metastatic disease) needs to be evaluated. To facilitate optimally cost-effective use of such tests, the user needs to retain the freedom to define thresholds for positive versus negative results. The cost of providing this necessary level of reassurance to oncologists may well limit the number of tests that can be developed in the future.

Prognostic and predictive molecular tests as part of new drug development

Aside from the issues discussed immediately above, cancer biomarkers have the potential to play a valuable role in future drug development, as part of the FDA's Critical Path Initiative [Karsdal *et al.* 2009]. This is an enormous area, and is largely outside the scope of this review. However, a small number of examples especially relevant to cancer will be discussed. Tamoxifen activity is dependent on its conversion by the CYP2D6 enzyme to its active metabolite endoxifen, and several studies show that poor metabolizers of tamoxifen do not get as much benefit from the drug as other patients [reviewed by Hoskins *et al.* 2009]. The Roche Amplichip (Roche, Basel, Switzerland), which measures the genetic variation in CYP2D6 and CYP2C19, is being applied to tamoxifen treatment in an attempt to improve efficacy for individual patients. However, well-designed clinical trials will be required to confirm the Amplichip's clinical utility in this setting [Hoskins *et al.* 2009]. In addition to the Amplichip, there are numerous other commercial pharmacogenetic initiatives, such as the Belgian company DNAVision (Charleroi, Belgium), that market genetic tests

to predict adverse reactions to 5-fluorouracil and irinotecan in colorectal cancer (http://www.dnavision.be/).

Molecular profiling at the DNA or RNA level may also be used to enrich clinical trial participants for those patients most likely to benefit from a drug [Adjei *et al.* 2009]. While this may narrow the eventual market for the drug, it may also reduce the likelihood of expensive failures at late developmental phases [Gutierrez *et al.* 2009]. The molecular profile of likely drug responders may be known before the trial begins, or may rely on comparing tumor DNA/mRNA profiles from phase II 'responders' with the profiles of phase II 'nonresponders [Maitournam and Simon, 2005]. Molecular profiles may also be used to predict pharmacodynamics. For example, in target engagement studies, instead of selecting phase II doses based on the maximum tolerated dose from a phase I dose-escalation study, phase II doses are determined by assessing the effect of a drug on the molecular profile of the drug's targeted pathway [Sawyers, 2008].

Conclusion

It is an exciting time to be an oncologist or cancer biologist. Significant advances have been made in technologies for DNA and RNA sequencing, SNP genotyping, mRNA and ncRNA quantification, molecular pathway analysis and integrative bioinformatics. These advances have greatly expanded our understanding of the molecular basis of tumor progression and individual treatment response. Several predictive and prognostic tests based on this new knowledge have revolutionized cancer medicine. However, this field still has a long way to go before it reaches its full potential. To accelerate their benefit to patients, prognostic and predictive markers need to be developed 'from the ground up' with specific clinical contexts in mind, and subjected to robust clinical evaluation. Methods developed for use with small amounts of FFPE tumor material or circulating tumor cells, as well as noninvasive molecular imaging, are likely to be especially useful. We may need to pay particular attention to optimal tumor sampling and to predictive methods that formally combine clinical and molecular data into a synergistic understanding of tumor progression. Bioinformatic initiatives that combine large amounts of cancer data such as caBIG are likely to play increasingly important roles. Overarching all of these points is the need to train translational physician-scientists who have a deep appreciation of medicine, cancer biology and the business of clinical implementation. As predictors and prognosticators continue their exciting expansion into cancer medicine, we must ensure they retain a strong evidence base.

Acknowledgements

We wish to acknowledge Drs. Michael Black, Parry Guilford, Ben Lawrence, John McCall, Arend Merrie and Li Wang for reading the manuscript. Our research in this area is supported by grants from the New Zealand Foundation of Research Science and Technology, the New Zealand Breast Cancer Research Trust, the Cancer Society of New Zealand, and the Health Research Council of New Zealand's International Investment Opportunities Fund.

Conflict of interest statement

None declared.

References

Abbod, M.F., Catto, J.W., Linkens, D.A. and Hamdy, F.C. (2007) Application of artificial intelligence to the management of urological cancer. *J Urol* 178: 1150–1156.

Abbott, L.H. and Michor, F. (2006) Mathematical Models of Targeted Cancer Therapy. *Br J Cancer* 95: 1136–1141.

Abdueva, D., Wing, M.R., Schaub, B. and Triche, T.J. (2007) Experimental comparison and evaluation of the affymetrix exon and U133plus2 genechip arrays. *PLoS One* 2: e913.

Acharya, C.R., Hsu, D.S., Anders, C.K., Anguiano, A., Salter, K.H., Walters, K.S. *et al.* (2008) Gene expression signatures, clinicopathological features, and individualized therapy in breast cancer. *JAMA* 299: 1574–1587.

Adjei, A.A., Christian, M. and Ivy, P. (2009) Novel designs and end points for phase II clinical trials. *Clin Cancer Res* 15: 1866–1872.

Ahmed, F.E. (2005) Molecular markers that predict response to colon cancer therapy. *Expert Rev Mol Diagn* 5: 353–375.

Alvarado, M.D., Jensen, E.H. and Yeatman, T.J. (2006) The potential role of gene expression in the management of primary and metastatic colorectal cancer. *Cancer Control* 13: 27–31.

Amado, R.G., Wolf, M., Peeters, M., Van Cutsem, E., Siena, S., Freeman, D.J. *et al.* (2008) Wild-type Kras is required for panitumumab efficacy in patients with metastatic colorectal cancer. *J Clin Oncol* 26: 1626–1634.

Anders, C.K., Acharya, C.R., Hsu, D.S., Broadwater, G., Garman, K., Foekens, J.A. *et al.* (2008)

Age-specific differences in oncogenic pathway deregulation seen in human breast tumors. *PLoS One* 3: e1373.

Andre, F. and Pusztai, L. (2006) Molecular classification of breast cancer: implications for selection of adjuvant chemotherapy. *Nat Clin Pract Oncol* 3: 621–632.

Andre, F., Michiels, S., Dessen, P., Scott, V., Suciu, V., Uzan, C. *et al.* (2009) Exonic expression profiling of breast cancer and benign lesions: a retrospective analysis. *Lancet Oncol* 10: 381–390.

Anjomshoaa, A., Lin, Y.H., Black, M.A., McCall, J.L., Humar, B., Song, S. *et al.* (2008) Reduced expression of a gene proliferation signature is associated with enhanced malignancy in colon cancer. *Br J Cancer* 99: 966–973.

Arango, D., Wilson, A.J., Shi, Q., Corner, G.A., Aranes, M.J., Nicholas, C. *et al.* (2004) Molecular mechanisms of action and prediction of response to oxaliplatin in colorectal cancer cells. *Br J Cancer* 91: 1931–1946.

Arriagada, R., Bergman, B., Dunant, A., Le Chervalier, T., Pignon, J.P. and Vansteenkiste, J. (2004) Cisplatin-based adjuvant chemotherapy in patients with completely resected non-small-cell lung cancer. *N Engl J Med* 350: 351–360.

Avard, D. and Knopper, B.M. (2009) Genomic medicine: considerations for health professionals and the public. *Genome Med* 1: 25–26.

Ayers, M., Symmans, W.F., Stec, J., Damokosh, A.I., Clark, E., Hess, K. *et al.* (2004) Gene expression profiles predict complete pathologic response to neoadjuvant paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide chemotherapy in breast cancer. *J Clin Oncol* 22: 2284–2293.

Barnett, G.C., West, C.M., Dunning, A.M., Elliott, R.M., Coles, C.E., Pharoah, P.D. *et al.* (2009) Normal tissue reactions to radiotherapy: towards tailoring treatment dose by genotype. *Nat Rev Cancer* 9: 134–142.

Barry, M.J. (2009) Screening for prostate cancer—the controversy that refuses to die. *N Engl J Med* 360: 1351–1354.

Bartels, C.L. and Tsongalis, G.J. (2009) MicroRNAs: novel biomarkers for human cancer. *Clin Chem* 55: 623–631.

Basil, C.F., Zhao, Y., Zavaglia, K., Jin, P., Panelli, M.C., Voiculescu, S. *et al.* (2006) Common cancer biomarkers. *Cancer Res* 66: 2953–2961.

Bastian, P.J., Yegnasubramanian, S., Palapattu, G.S., Rogers, C.G., Lin, X., De Marzo, A.M. *et al.* (2004) Molecular biomarker in prostate cancer: the role of CpG island hypermethylation. *Eur Urol* 46: 698–708.

Belinsky, S.A. (2004) Gene promoter hypermethylation as a biomarker in lung cancer. *Nat Rev Cancer* 4: 1–11.

Bild, A.H., Potti, A. and Nevins, J.R. (2006a) Linking oncogenic pathways with therapeutic opportunities. *Nat Rev Cancer* 6: 735–741.

Bild, A.H., Yao, G., Chang, J.T., Wang, Q., Potti, A., Chasse, D. *et al.* (2006b) Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature* 439: 353–357.

Blenkiron, C. and Miska, E.A. (2007) miRNAs in cancer: approaches, aetiology, diagnostics and therapy. *Hum Mol Genet* 16(Spec No 1): R106–113.

Blenkiron, C., Goldstein, L.D., Thorne, N.P., Spiteri, I., Chin, S.F., Dunning, M.J. *et al.* (2007) MicroRNA expression profiling of human breast cancer identifies new markers of tumor subtype. *Genome Biol* 8: R214.

Bloomston, M., Frankel, W.L., Petrocca, F., Volinia, S., Alder, H., Hagan, J.P. *et al.* (2007) MicroRNA expression patterns to differentiate pancreatic adenocarcinoma from normal pancreas and chronic pancreatitis. *JAMA* 297: 1901–1908.

Bokemeyer, C., Bondarenko, I., Makhson, A., Hartmann, J.T., Aparicio, J., De Braud, F. *et al.* (2009) Fluorouracil, leucovorin, and oxaliplatin with and without cetuximab in the first-line treatment of metastatic colorectal cancer. *J Clin Oncol* 27: 663–671.

Bowers, J., Mitchell, J., Beer, E., Buzby, P.R., Causey, M., Efcavitch, J.W. *et al.* (2009) Virtual terminator nucleotides for next-generation dna sequencing. *Nat Methods* 6: 593–595.

Brindle, K. (2008) New approaches for imaging tumour responses to treatment. *Nat Rev Cancer* 8: 94–107.

Bringuier, P.P., Umbas, R., Schaafsma, H.E., Karthaus, H.F., Debruyne, F.M. and Schalken, J.A. (1993) Decreased E-cadherin immunoreactivity correlates with poor survival in patients with bladder tumors. *Cancer Res* 53: 3241–3245.

Bui, M.H., Seligson, D., Han, K.R., Pantuck, A.J., Dorey, F.J., Huang, Y. *et al.* (2003) Carbonic anhydrase Ix is an independent predictor of survival in advanced renal clear cell carcinoma: implications for prognosis and therapy. *Clin Cancer Res* 9: 802–811.

Bukhari, M.H. and Akhtar, Z.M. (2009) Comparison of accuracy of diagnostic modalities for evaluation of breast cancer with review of literature. *Diagn Cytopathol* 37: 416–424.

Bustin, S.A. and Nolan, T. (2004) Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. *J Biomol Tech* 15: 155–166.

Calin, G.A., Dumitru, C.D., Shimizu, M., Bichi, R., Zupo, S., Noch, E. *et al.* (2002) Frequent deletions and down-regulation of micro-RNA genes Mir15 and Mir16 at 13q14 in chronic lymphocytic leukemia. *Proc Natl Acad Sci U S A* 99: 15524–15529.

Cardoso, F., Van 't Veer, L., Rutgers, E., Loi, S., Mook, S. and Piccart-Gebhart, M.J. (2008) Clinical application of the 70-gene profile: the Mindact trial. *J Clin Oncol* 26: 729–735.

Carey, F.A., Lamb, D. and Bird, C.C. (1990) Importance of sampling method in DNA analysis of lung cancer. *J Clin Pathol* 43: 820–823.

Chang, H.Y., Nuyten, D.S., Sneddon, J.B., Hastie, T., Tibshirani, R., Sorlie, T. *et al.* (2005) Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. *Proc Natl Acad Sci U S A* 102: 3738–3743.

Chang, J.T., Carvalho, C., Mori, S., Bild, A.H., Gatza, M.L., Wang, Q. *et al.* (2009) A genomic strategy to elucidate modules of oncogenic pathway signaling networks. *Mol Cell* 34: 104–114.

Chen, X., Ba, Y., Ma, L., Cai, X., Yin, Y., Wang, K. *et al.* (2008) Characterization of microrna s in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. *Cell Res* 18: 997–1006.

Chin, L. and Gray, J.W. (2008) Translating insights from the cancer genome into clinical practice. *Nature* 452: 553–563.

Clark, T.A., Schweitzer, A.C., Chen, T.X., Staples, M.K., Lu, G., Wang, H. *et al.* (2007) Discovery of tissue-specific exons using comprehensive human exon microarrays. *Genome Biol* 8: R64.

Cloughesy, T.F., Yoshimoto, K., Nghiemphu, P., Brown, K., Dang, J., Zhu, S. *et al.* (2008) Antitumor activity of rapamycin in a phase I trial for patients with recurrent PTEN-deficient glioblastoma. *PLoS Med* 5: e8.

Coate, L.E., John, T., Tsao, M. and Shepherd, F.A. (2009) Molecular predictive and prognostic markers in non-small-cell lung cancer. *Lancet* 10: 1001–1010.

Cobo, M.S., Massuti, B., Moran, T., Chaib, L., Pérez-Roca, U., Jiménez, D. *et al.* (2008) Spanish Customized Adjuvant Trial (SCAT) based on BRCA1 mRNA levels. *Proc Am Soc Clin Oncol* 26: abstr 7533.

Cohen, S.J., Punt, C.J., Iannotti, N., Saidman, B.H., Sabbath, K.D., Gabrail, N.Y. *et al.* (2009) Prognostic significance of circulating tumor cells in patients with metastatic colorectal cancer. *Ann Oncol* 20: 1223–1229.

Contractor, K.B. and Aboagye, E.O. (2009) Monitoring predominantly cytostatic treatment response with 18f-FDG PET. *J Nucl Med* 50(Suppl 1): 97S–105S.

Cordon-Cardo, C., Kotsianti, A., Verbel, D.A., Teverovskiy, M., Capodieci, P., Hamann, S. *et al.* (2007) Improved prediction of prostate cancer recurrence through systems pathology. *J Clin Invest* 117: 1876–1883.

Cristofanilli, M., Budd, G.T., Ellis, M.J., Stopeck, A., Matera, J., Miller, M.C. *et al.* (2004) Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *N Engl J Med* 351: 781–791.

Davis, L.M., Harris, C., Tang, L., Doherty, P., Hraber, P., Sakai, Y. *et al.* (2007) Amplification patterns of three genomic regions predict distant recurrence in breast carcinoma. *J Mol Diagn* 9: 327–336.

Dematteo, R.P., Ballman, K.V., Antonescu, C.R., Maki, R.G., Pisters, P.W., Demetri, G.D. *et al.* (2009) Adjuvant imatinib mesylate after resection of localised, primary gastrointestinal stromal tumour: a randomised, double-blind, placebo-controlled trial. *Lancet* 373: 1097–1104.

Donovan, M.J., Khan, F.M., Fernandez, G., Mesa-Tejada, R., Sapir, M., Zubek, V.B. *et al.* (2009) Personalized prediction of tumor response and cancer progression on prostate needle biopsy. *J Urol* 182: 125–132.

Dowsett, M., Houghton, J., Iden, C., Salter, J., Farndon, J., A'hern, R. *et al.* (2006) Benefit from adjuvant tamoxifen therapy in primary breast cancer patients according oestrogen receptor, progesterone receptor, Egf receptor and Her2 status. *Ann Oncol* 17: 818–826.

Dowsett, M., Smith, I.E., Ebbs, S.R., Dixon, J.M., Skene, A., A'Hern, R. *et al.* (2007) Prognostic value of Ki67 expression after short-term presurgical endocrine therapy for primary breast cancer. *J Natl Cancer Inst* 99: 167–170.

Dunn, J., Baborie, A., Alam, F., Joyce, K., Moxham, M., Sibson, R. *et al.* (2009) Extent of MGMT promoter methylation correlates with outcome in glioblastomas given temozolomide and radiotherapy. *Br J Cancer* 101: 124–131.

Dworkin, A.M., Huang, T.H. and Toland, A.E. (2009) Epigenetic alterations in the breast: implications for breast cancer detection, prognosis and treatment. *Semin Cancer Biol* 19: 165–171.

Early Breast Cancer Trialists' Collaborative Group. (1998) Tamoxifen for early breast cancer: an overview of the randomised trials. *Lancet* 351: 1451–1467.

Easton, D.F., Pooley, K.A., Dunning, A.M., Pharoah, P.D., Thompson, D., Ballinger, D.G. *et al.* (2007) Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 447: 1087–1093.

Edgar, R., Domrachev, M. and Lash, A.E. (2002) Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 30: 207–210.

Elledge, R.M., Green, S., Pugh, R., Allred, D.C., Clark, G.M., Hill, J. *et al.* (2000) Estrogen receptor (Er) and progesterone receptor (Pgr), by ligand-binding assay compared with Er, Pgr and Ps2, by immunohistochemistry in predicting response to tamoxifen in metastatic breast cancer: a southwest oncology group study. *Int J Cancer* 89: 111–117.

Esteller, M. (2008) Epigenetics in cancer. *N Engl J Med* 358: 1148–1159.

Fan, C., Oh, D.S., Wessels, L., Weigelt, B., Nuyten, D.S., Nobel, A.B. *et al.* (2006) Concordance among gene-expression-based predictors for breast cancer. *N Engl J Med* 355: 560–569.

Fang, Z., Soleymani, L., Pampalakis, G., Yoshimoto, M., Squire, J.A., Sargent, E.H. *et al.* (2009) Direct

profiling of cancer biomarkers in tumor tissue using a multiplexed nanostructured microelectrode integrated circuit. *ACS Nano* 3: 3207–3213.

Faratian, D. and Bartlett, J. (2008) Predictive markers in breast cancer—the future. *Histopathology* 52: 91–98.

Faustino, N.A. and Cooper, T.A. (2003) Pre-mRNA splicing and human disease. *Genes Dev* 17: 419–437.

Ferrone, C.R., Finkelstein, D.M., Thayer, S.P., Muzikansky, A., Fernandez-Delcastillo, C. and Warshaw, A.L. (2006) Perioperative Ca19-9 levels can predict stage and survival in patients with resectable pancreatic adenocarcinoma. *J Clin Oncol* 24: 2897–2902.

Fey, M.F. and Tobler, A. (1996) Tumour heterogeneity and clonality—an old theme revisited. *Ann Oncol* 7: 121–128.

Furusato, B., Shaheduzzaman, S., Petrovics, G., Dobi, A., Seifert, M., Ravindranath, L. *et al.* (2007) Transcriptome analyses of benign and malignant prostate epithelial cells in formalin-fixed paraffin-embedded whole-mounted radical prostatectomy specimens. *Prostate Cancer Prostatic Dis* 11: 194–197.

Galon, J., Franchimont, D., Hiroi, N., Frey, G., Boettner, A., Ehrhart-Bornstein, M. *et al.* (2002) Gene profiling reveals unknown enhancing and suppressive actions of glucocorticoids on immune cells. *FASEB J* 16: 61–71.

Gardina, P.J., Clark, T.A., Shimada, B., Staples, M.K., Yang, Q., Veitch, J. *et al.* (2006) Alternative splicing and differential gene expression in colon cancer detected by a whole genome exon array. *BMC Genomics* 7: 325.

Garman, K.S., Acharya, C.R., Edelman, E., Grade, M., Gaedcke, J., Sud, S. *et al.* (2008) A genomic approach to colon cancer risk stratification yields biologic insights into therapeutic opportunities. *Proc Natl Acad Sci U S A* 105: 19432–19437.

Ge, H., Walhout, A.J. and Vidal, M. (2003) Integrating 'omic' information: a bridge between genomics and systems biology. *Trends Genet* 19: 551–560.

GenomeWeb (2009) Pharmacogenetic-based prescribing slow to take hold. GenomeWeb Daily News; 16 February 2009.

Gerdes, J., Lemke, H., Baisch, H., Wacker, H.H., Schwab, U. and Stein, H. (1984) Cell cycle analysis of a cell proliferation-associated human nuclear antigen defined by the monoclonal antibody Ki-67. *J Immunol* 133: 1710–1715.

Gevaert, O., De Smet, F., Timmerman, D., Moreau, Y. and De Moor, B. (2006) Predicting the prognosis of breast cancer by integrating clinical and microarray data with bayesian networks. *Bioinformatics* 22: e184–190.

Geyer, F.C. and Reis-Filho, J.S. (2009) Microarray-based gene expression profiling as a clinical tool for breast cancer management: are we there yet? *Int J Surg Path* 17: 285–302.

Glas, A., Roepman, P., Salazar, R., Capella, G., Moreno, V., Westerga, J. *et al.* (2009) Development and validation of a robust prognostic and predictive signature for colorectal cancer (CRC) patients. *J Clin Oncol* 27: (Supplement abstract 4036).

Goldstein, L.J., Gray, R., Badve, S., Childs, B.H., Yoshizawa, C., Rowley, S. *et al.* (2008) Prognostic utility of the 21-gene assay in hormone receptor-positive operable breast cancer compared with classical clinicopathologic features. *J Clin Oncol* 26: 4063–4071.

Gordon, M.A., Gil, J., Lu, B., Zhang, W., Yang, D., Yun, J. *et al.* (2006) Genomic profiling associated with recurrence in patients with rectal cancer treated with chemoradiation. *Pharmacogenomics* 7: 67–88.

Guffanti, A., Iacono, M., Pelucchi, P., Kim, N., Solda, G., Croft, L.J. *et al.* (2009) A transcriptional sketch of a primary human breast cancer by 454 deep sequencing. *BMC Genomics* 10: 163.

Gutierrez, M.E., Kummar, S. and Giaccone, G. (2009) Next generation oncology drug development: opportunities and challenges. *Nat Rev Clin Oncol* 6: 259–265.

Gutman, S. and Kessler, L.G. (2006) The US Food and Drug Administration perspective on cancer biomarker development. *Nat Rev Cancer* 6: 565–571.

Hallstrom, T.C., Mori, S. and Nevins, J.R. (2008) An E2f1-dependent gene expression program that determines the balance between proliferation and cell death. *Cancer Cell* 13: 11–22.

Hanauer, D.A., Rhodes, D.R., Sinha-Kumar, C. and Chinnaiyan, A.M. (2007) Bioinformatics approaches in the study of cancer. *Curr Mol Med* 7: 133–141.

Hargreaves, R.J. (2008) The role of molecular imaging in drug discovery and development. *Clin Pharmacol Ther* 83: 349–353.

Hayden, E.C. (2009a) Genomics shifts focus to rare diseases. *Nature* 461: 24.

Hayden, E.C. (2009b) Personalized cancer therapy gets closer. *Nature* 458: 131–132.

Hegde, P., Qi, R., Gaspard, R., Abernathy, K., Dharap, S., Earle-Hughes, J. *et al.* (2001) Identification of tumor markers in models of human colorectal cancer using a 19,200-element complementary DNA microarray. *Cancer Res* 61: 7792–7797.

Hoskins, J.M., Carey, L.A. and McLeod, H.L. (2009) CYP2D6 and tamoxifen: DNA matters in breast cancer. *Nat Rev Cancer* 9: 576–586.

Huang, E., West, M. and Nevins, J.R. (2003) Gene expression profiling for prediction of clinical characteristics of breast cancer. *Recent Prog Horm Res* 58: 55–73.

Huang, Y.W., Jansen, R.A., Fabbri, E., Potter, D., Liyanarachchi, S., Chan, M.W. *et al.* (2009)

Identification of candidate epigenetic biomarkers for ovarian cancer detection. *Oncol Rep* 22: 853–861.

Huerta, S., Harris, D.M., Jazirehi, A., Bonavida, B., Elashoff, D., Livingston, E.H. *et al.* (2003) Gene expression profile of metastatic colon cancer cells resistant to cisplatin-induced apoptosis. *Int J Oncol* 22: 663–670.

Ioannidis, J.P. (2009) Personalized genetic prediction: too limited, too expensive, or too soon? *Ann Int Med* 150: 139–141.

Jaeger, J., Koczan, D., Thiesen, H.J., Ibrahim, S.M., Gross, G., Spang, R. *et al.* (2007) Gene expression signatures for tumor progression, tumor subtype, and tumor thickness in laser-microdissected melanoma tissues. *Clin Cancer Res* 13: 806–815.

James, C.R., Quinn, J.E., Mullan, P.B., Johnston, P.G. and Harkin, D.P. (2007) BRCA1, a potential predictive biomarker in the treatment of breast cancer. *Oncologist* 12: 142–150.

Josephs, D., Spicer, J. and O'Doherty, M. (2009) Molecular imaging in clinical trials. *Target Oncol.* Epub 21September 2009.

Kalscheuer, S., Zhang, X., Zeng, Y. and Upadhyaya, P. (2008) Differential expression of microRNAs in early-stage neoplastic transformation in the lungs of F344 rats chronically treated with the tobacco carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone. *Carcinogenesis* 29: 2394–2399.

Kanai, Y. (2007) Alteration of DNA methylation associated with abnormalities of DNA methyltransferases in human cancers during transition from a precancerous to a malignant state. *Carcinogenesis* 28: 2434 – 2442.

Kanehisa, M., Goto, S., Hattori, M., Aoki-Kinoshita, K.F., Itoh, M., Kawashima, S. *et al.* (2006) From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res* 34: D354–357.

Kapp, A.V., Jeffrey, S.S., Langerod, A., Borresen-Dale, A.L., Han, W., Noh, D.Y. *et al.* (2006) Discovery and validation of breast cancer subtypes. *BMC Genomics* 7: 231.

Karsdal, M.A., Henriksen, K., Leeming, D.J., Mitchell, P., Duffin, K., Barascuk, N. *et al.* (2009) Biochemical markers and the FDA critical path: how biomarkers may contribute to the understanding of pathophysiology and provide unique and necessary tools for drug development. *Biomarkers* 14: 181–202.

Katsanis, S.H., Javitt, G. and Hudson, K. (2008) Public health. A case study of personalized medicine. *Science* 320: 53–54.

Khoury, M.J., Gwinn, M., Yoon, P.W., Dowling, N., Moore, C.A. and Bradley, L. (2007) The continuum of translation research in genomic medicine: how can we accelerate the appropriate integration of human genome discoveries into health care and disease prevention? *Genet Med* 9: 665–674.

Kim, G.P., Colangelo, L.H., Wieand, H.S., Paik, S., Kirsch, I.R., Wolmark, N. *et al.* (2007) Prognostic and predictive roles of high-degree microsatellite instability in colon cancer: a National Cancer Institute-National Surgical Adjuvant Breast and Bowel Project Collaborative Study. *J Clin Oncol* 25: 767–772.

Kim, C., Taniyama, Y. and Paik, S. (2009) Gene expression-based prognostic and predictive markers for breast cancer: a primer for practicing pathologists. *Arch Pathol Lab Med* 133: 855–859.

Koscielny, S. (2008) Critical review of microarray-based prognostic tests and trials in breast cancer. *Curr Opin Obstet Gynecol* 20: 47–50.

Kuremsky, J.G., Tepper, J.E. and McLeod, H.L. (2009) Biomarkers for response to neoadjuvant chemoradiation for rectal cancer. *Int J Radiat Oncol Biol Phys* 74: 673–688.

Laking, G., Lord, J. and Fischer, A. (2006) The economics of diagnosis. *Health Econ* 15: 1109–1120.

Lazebnik, Y. (2002) Can a biologist fix a radio? Or, what I learned while studying apoptosis. *Cancer Cell* 2: 179–182.

Lebeau, A., Deimling, D., Kaltz, C., Sendelhofert, A., Iff, A., Luthardt, B. *et al.* (2001) Her-2/Neu analysis in archival tissue samples of human breast cancer: comparison of immunohistochemistry and fluorescence in situ hybridization. *J Clin Oncol* 19: 354–363.

Lee, R.C., Feinbaum, R.L. and Ambros, V. (1993) The *C. elegans* heterochronic gene Lin-4 encodes small RNAs with antisense complementarity to Lin-14. *Cell* 75: 843–854.

Ley, T.J., Mardis, E.R., Ding, L., Fulton, B., McLellan, M.D., Chen, K. *et al.* (2008) DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. *Nature* 456: 66–72.

Li, J. and Makrigiorgos, G.M. (2009) Cold-PCR: a new platform for highly improved mutation detection in cancer and genetic testing. *Biochem Soc Trans* 37: 427–432.

Li, M., Lin, Y.M., Hasegawa, S., Shimokawa, T., Murata, K., Kameyama, M. *et al.* (2004) Genes associated with liver metastasis of colon cancer, identified by genome-wide cDNA microarray. *Int J Oncol* 24: 305–312.

Li, J., Martinka, M. and Li, G. (2008) Role of Ing4 in human melanoma cell migration, invasion and patient survival. *Carcinogenesis* 29: 1373–1379.

Li, Y., Sawalha, A.H. and Lu, Q. (2009) Aberrant DNA methylation in skin diseases. *J Dermatol Sci* 54: 143–149.

Lievre, A., Bachet, J.B., Le Corre, D., Boige, V., Landi, B., Emile, J.F. *et al.* (2006) Kras mutation status is predictive of response to cetuximab therapy in colorectal cancer. *Cancer Res* 66: 3992–3995.

Lin, H.M., Chatterjee, A., Lin, Y.H., Anjomshoaa, A., Fukuzawa, R., McCall, J.L. *et al.* (2007a) Genome

wide expression profiling identifies genes associated with colorectal liver metastasis. *Oncol Rep* 17: 1541–1549.

Lin, Y.H., Friederichs, J., Black, M.A., Mages, J., Rosenberg, R., Guilford, P.J. *et al.* (2007b) Multiple gene expression classifiers from different array platforms predict poor prognosis of colorectal cancer. *Clin Cancer Res* 13: 498–507.

Linton, K., Hey, Y., Dibben, S., Miller, C., Freemont, A., Radford, J. *et al.* (2009) Methods comparison for high-resolution transcriptional analysis of archival material on Affymetrix Plus 2.0 and Exon 1.0 microarrays. *Biotechniques* 47: 587–596.

Lipponen, P., Eskelinen, M., Syrjanen, S., Tervahauta, A. and Syrjanen, K. (1991) Use of immunohistochemically demonstrated C-Erb B-2 oncoprotein expression as a prognostic factor in transitional cell carcinoma of the urinary bladder. *Eur Urol* 20: 238–242.

Lisboa, P.J., Etchells, T.A., Jarman, I.H., Aung, M.S., Chabaud, S., Bachelot, T. *et al.* (2007) Development of a rule based prognostic tool for her 2 positive breast cancer patients. *Conf Proc IEEE Eng Med Biol Soc* 1: 5416–5419.

Liu, R., Wang, X., Chen, G.Y., Dalerba, P., Gurney, A., Hoey, T. *et al.* (2007) The Prognostic role of a gene signature from tumorigenic breast-cancer cells. *N Engl J Med* 356: 217–226.

Liukkonen, T., Lipponen, P., Raitanen, M., Kaasinen, E., Ala-Opas, M., Rajala, P. *et al.* (2000) Evaluation of P21waf1/Cip1 and cyclin D1 expression in the progression of superficial bladder cancer. Finbladder Group. *Urol Res* 28: 285–292.

Logothetis, C.J., Xu, H.J., Ro, J.Y., Hu, S.X., Sahin, A., Ordonez, N. *et al.* (1992) Altered expression of retinoblastoma protein and known prognostic variables in locally advanced bladder cancer. *J Natl Cancer Inst* 84: 1256–1261.

Ma, X.J., Wang, Z., Ryan, P.D., Isakoff, S.J., Barmettler, A., Fuller, A. *et al.* (2004) A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. *Cancer Cell* 5: 607–616.

Maitournam, A. and Simon, R. (2005) On the efficiency of targeted clinical trials. *Stat Med* 24: 329–339.

Mallardo, M., Poltronieri, P. and D'urso, O.F. (2008) Non-protein coding RNA biomarkers and differential expression in cancers: a review. *J Exp Clin Cancer Res* 27: 19.

Mandruzzato, S., Callegaro, A., Turcatel, G., Francescato, S., Montesco, M.C., Chiarion-Sileni, V. *et al.* (2006) A gene expression signature associated with survival in metastatic melanoma. *J Transl Med* 4: 50.

Mariadason, J.M., Arango, D., Shi, Q., Wilson, A.J., Corner, G.A., Nicholas, C. *et al.* (2003) Gene expression profiling-based prediction of response of colon carcinoma cells to 5-fluorouracil and camptothecin. *Cancer Res* 63: 8791–8812.

Mascaux, C., Iannino, N., Martin, B., Paesmans, M., Berghmans, T., Dusart, M. *et al.* (2005) The role of ras oncogene in survival of patients with lung cancer: a systematic review of the literature with meta-analysis. *Br J Cancer* 92: 131–139.

Mass, R.D., Press, M.F., Anderson, S., Cobleigh, M.A., Vogel, C.L., Dybdal, N. *et al.* (2005) Evaluation of clinical outcomes according to Her2 detection by fluorescence in situ hybridization in women with metastatic breast cancer treated with trastuzumab. *Clin Breast Cancer* 6: 240–246.

McCormick, K.A. (2009) Individualizing cancer care with interoperable information systems. *Stud Health Technol Inform* 146: 383–390.

Mengual, L., Burset, M., Ars, E., Lozano, J.J., Villavicencio, H., Ribal, M.J. *et al.* (2009) DNA microarray expression profiling of bladder cancer allows identification of noninvasive diagnostic markers. *J Urol* 182: 741–748.

Meropol, N.J., Gold, P.J., Diasio, R.B., Andria, M., Dhami, M., Godfrey, T. *et al.* (2006) Thymidine phosphorylase expression is associated with response to capecitabine plus irinotecan in patients with metastatic colorectal cancer. *J Clin Oncol* 24: 4069–4077.

Metzger, R., Danenberg, K., Leichman, C.G., Salonga, D., Schwartz, E.L., Wadler, S. *et al.* (1998) High basal level gene expression of thymidine phosphorylase (platelet-derived endothelial cell growth factor) in colorectal tumors is associated with nonresponse to 5-fluorouracil. *Clin Cancer Res* 4: 2371–2376.

Mitchell, P.S., Parkin, R.K., Kroh, E.M., Fritz, B.R., Wyman, S.K., Pogosova-Agadjanyan, E.L. *et al.* (2008) Circulating microRNAs as stable blood-based markers for cancer detection. *Proc Natl Acad Sci U S A* 105: 10513–10518.

Miyake, H., Eto, H., Arakawa, S., Kamidono, S. and Hara, I. (2002) Over Expression of Cd44v8-10 in urinary exfoliated cells as an independent prognostic predictor in patients with urothelial cancer. *J Urol* 167: 1282–1287.

Mostert, B., Sleijfer, S., Foekens, J.A. and Gratama, J.W. (2009) Circulating tumor cells (CTCs): detection methods and their clinical relevance in breast cancer. *Cancer Treat Rev* 35: 463–474.

Nagarajan, R.P. and Costello, J.F. (2009) Epigenetic mechanisms in glioblastoma multiforme. *Semin Cancer Biol* 19: 188–197.

Nagata, Y., Lan, K., Zhou, X., Tan, M., Esteva, F., Sahin, A. *et al.* (2004) PTEN activation contributes to tumor inhibition by trastuzumab, and loss of PTEN predicts trastuzumab resistance in patients. *Cancer Cell* 6: 117–127.

Nakamura, S., Yagata, H., Ohno, S., Yamaguchi, H., Iwata, H., Tsunoda, N. *et al.* (2009) Multi-center

study evaluating circulating tumor cells as a surrogate for response to treatment and overall survival in metastatic breast cancer. *Breast Cancer.* Epub 1 August 2009.

Nezos, A., Lembessis, P., Sourla, A., Pissimissi, N., Gogas, H. and Koutsilieris, M. (2009) Molecular markers detecting circulating melanoma cells by reverse transcription polymerase chain reaction: methodological pitfalls and clinical relevance. *Clin Chem Lab Med* 47: 1–11.

Ng, S.B., Turner, E.H., Robertson, P.D., Flygare, S.D., Bigham, A.W., Lee, C. *et al.* (2009) Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* 461: 272–276.

Nguyen, P.L., Swanson, P.E., Jaszcz, W., Aeppli, D.M., Zhang, G., Singleton, T.P. *et al.* (1994) Expression of epidermal growth factor receptor in invasive transitional cell carcinoma of the urinary bladder. A multivariate survival analysis. *Am J Clin Pathol* 101: 166–176.

NIH (2003) NIH data sharing policy and implementation guidance. http://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm.

Olaussen, K.A., Dunant, A., Fouret, P., Brambilla, E., André, F., Haddad, V. *et al.* (2006) DNA repair by Ercc1 in non-small-cell lung cancer and cisplatin-based adjuvant chemotherapy. *N Engl J Med* 355: 983–991.

Oldenhuis, C.N.A.M., Oosting, S.F., Gietema, J.A. and De Vries, E.G.E. (2008) Prognostic versus predictive value of biomarkers in oncology. *Eur J Cancer* 44: 946–953.

Paik, S., Shak, S., Tang, G., Kim, C., Baker, J., Cronin, M. *et al.* (2004) A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 351: 2817–2826.

Paik, S., Kim, C.Y., Song, Y.K. and Kim, W.S. (2005) Technology insight: application of molecular techniques to formalin-fixed paraffin-embedded tissues from breast cancer. *Nat Clin Pract Oncol* 2: 246–254.

Paik, S., Tang, G., Shak, S., Kim, C., Baker, J., Kim, W. *et al.* (2006) Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J Clin Oncol* 24: 3726–3734.

Palomaki, G.E., Bradley, L.A., Douglas, M.P., Kolor, K. and Dotson, W.D. (2009) Can Ugt1a1 genotyping reduce morbidity and mortality in patients with metastatic colorectal cancer treated with irinotecan? An evidence-based review. *Genet Med* 11: 21–34.

Pao, W., Miller, V.A., Politi, K.A., Riely, G.J., Somwar, R., Zakowski, M.F. *et al.* (2005) Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the egfr kinase domain. *PLoS Med* 2: e73.

Parkinson, H., Kapushesky, M., Shojatalab, M., Abeygunawardena, N., Coulson, R., Farne, A. *et al.* (2007) Arrayexpress—a public database of microarray experiments and gene expression profiles. *Nucleic Acids Res* 35: D747–750.

Penault-Llorca, F., Bilous, M., Dowsett, M., Hanna, W., Osamura, R.Y., Ruschoff, J. *et al.* (2009) Emerging technologies for assessing Her2 amplification. *Am J Clin Pathol* 132: 539–548.

Perou, C.M., Sorlie, T., Eisen, M.B., Van De Rijn, M., Jeffrey, S.S., Rees, C.A. *et al.* (2000) Molecular portraits of human breast tumours. *Nature* 406: 747–752.

Pertschuk, L.P., Gaetjens, E., Carter, A.C., Brigati, D.J., Kim, D.S. and Tobin, E.H. (1979) Histochemistry of steroid receptors in breast cancer: an overview. *Ann Clin Lab Sci* 9: 219–224.

Petricoin, E.F., Belluco, C., Araujo, R.P. and Liotta, L.A. (2006) The blood peptidome: a higher dimension of information content for cancer biomarker discovery. *Nat Rev Cancer* 6: 961–967.

Phillips, K.A., Marshall, D.A., Haas, J.S., Elkin, E.B., Liang, S.Y., Hassett, M.J. *et al.* (2009) Clinical practice patterns and cost effectiveness of human epidermal growth receptor 2 testing strategies in breast cancer patients. *Cancer* 115: 5166–5174.

Pittman, J., Huang, E., Dressman, H., Horng, C.F., Cheng, S.H., Tsou, M.H. *et al.* (2004) Integrated modeling of clinical and gene expression information for personalized prediction of disease outcomes. *Proc Natl Acad Sci U S A* 101: 8431–8436.

Pitts, P.J. (2008) FDA and the critical path to twenty-first-century medicine. *J Med Philos* 3: 515–523.

Poston, G.J., Adam, R., Alberts, S., Curley, S., Figueras, J., Haller, D. *et al.* (2005) Oncosurge: a strategy for improving resectability with curative intent in metastatic colorectal cancer. *J Clin Oncol* 23: 7125–7134.

Rakha, E.A., El-Sayed, M.E., Reis-Filho, J.S. and Ellis, I.O. (2008) Expression profiling technology: its contribution to our understanding of breast cancer. *Histopathology* 52: 67–81.

Rangel, J., Torabian, S., Shaikh, L., Nosrati, M., Baehner, F.L., Haqq, C. *et al.* (2006) Prognostic significance of nuclear receptor coactivator-3 overexpression in primary cutaneous melanoma. *J Clin Oncol* 24: 4565–4569.

Ransohoff, D.F. (2005) Bias as a threat to the validity of cancer molecular-marker research. *Nat Rev Cancer* 5: 142–149.

Ravdin, P.M., Siminoff, L.A., Davis, G.J., Mercer, M.B., Hewlett, J., Gerson, N. *et al.* (2001) Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. *J Clin Oncol* 19: 980–991.

Reid, J.F., Lusa, L., De Cecco, L., Coradini, D., Veneroni, S., Daidone, M.G. *et al.* (2005) Limits of predictive models using microarray data for breast cancer clinical treatment outcome. *J Natl Cancer Inst* 97: 927–930.

Reschke, M., Mihic-Probst, D., Van Der Horst, E.H., Knyazev, P., Wild, P.J., Hutterer, M. *et al.* (2008) Her3 is a determinant for poor prognosis in melanoma. *Clin Cancer Res* 14: 5188–5197.

Richter, J., Wagner, U., Kononen, J., Fijan, A., Bruderer, J., Schmid, U. *et al.* (2000) High-throughput tissue microarray analysis of cyclin e gene amplification and overexpression in urinary bladder cancer. *Am J Pathol* 157: 787–794.

Rikova, K., Guo, A., Zeng, Q., Possemato, A., Yu, J., Haack, H. *et al.* (2007) Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. *Cell* 131: 1190–1203.

Ring, B.Z., Seitz, R.S., Beck, R., Shasteen, W.J., Tarr, S.M., Cheang, M.C. *et al.* (2006) Novel prognostic immunohistochemical biomarker panel for estrogen receptor-positive breast cancer. *J Clin Oncol* 24: 3039–3047.

Rogowski, W.H., Grosse, S.D. and Khoury, M.J. (2009) Challenges of translating genetic tests into clinical and public health practice. *Nat Rev Genet* 10: 489–495.

Rosell, R., Danenberg, K.D., Alberola, V., Bepler, G., Sanchez, J.J., Comps, C. *et al.* (2004) Ribonucleotide reductase messenger RNA expression and survival in gemcitabine/cisplatin-treated advanced non-small cell lung cancer patients. *Clin Cancer Res* 10: 1318–1325.

Rosell, R., Skrzypski, M., Jassem, E., Taron, M., Bartolucci, R. *et al.* (2007) BRCA1: a novel prognostic factor in resected non-small-cell lung cancer. *PLos One* 2: e1129.

Ross, J.S. (2009a) Multigene classifiers, prognostic factors, and predictors of breast cancer clinical outcome. *Adv Anat Pathol* 16: 204–215.

Ross, J.S. (2009b) Breast cancer biomarkers and Her2 testing after 10 years of anti-Her2 therapy. *Drug News Perspect* 22: 93–106.

Ross, J.S. and Slodkowska, E.A. (2009) Circulating and disseminated tumor cells in the management of breast cancer. *Am J Clin Pathol* 132: 237–245.

Ross, J.S., Symmans, W.F., Pusztai, L. and Hortobagyi, G.N. (2007) Standardizing slide-based assays in breast cancer: hormone receptors, Her2, and sentinel lymph nodes. *Clin Cancer Res* 13: 2831–2835.

Ross, J.S., Hatzis, C., Symmans, W.F., Pusztai, L. and Hortobagyi, G.N. (2008) Commercialized multigene predictors of clinical outcome for breast cancer. *Oncologist* 13: 477–493.

Rouzier, R., Perou, C.M., Symmans, W.F., Ibrahim, N., Cristofanilli, M., Anderson, K. *et al.* (2005) Breast cancer molecular subtypes respond differently to preoperative chemotherapy. *Clin Cancer Res* 11: 5678–5685.

Rouzier, R., Pusztai, L., Delaloge, S., Gonzalez-Angulo, A.M., Andre, F., Hess, K.R. *et al.* (2005) Nomograms to predict pathologic complete response and metastasis-free survival after preoperative chemotherapy for breast cancer. *J Clin Oncol* 23: 8331–8339.

Saidi, S.A., Holland, C.M., Kreil, D.P., Mackay, D.J., Charnock-Jones, D.S., Print, C.G. *et al.* (2004) Independent component analysis of microarray data in the study of endometrial cancer. *Oncogene* 23: 6677–6683.

Salter, K.H., Acharya, C.R., Walters, K.S., Redman, R., Anguiano, A., Garman, K.S. *et al.* (2008) An integrated approach to the prediction of chemotherapeutic response in patients with breast cancer. *PLoS One* 3: e1908.

Sawyers, C.L. (2008) The cancer biomarker problem. *Nature* 452: 548–552.

Scartozzi, M., Bittoni, A., Pistelli, M., Galizia, E., Berardi, R., Giampieri, R. *et al.* (2009) Toward molecularly selected chemotherapy for advanced gastric cancer: state of the art and future perspectives. *Cancer Treat Rev* 35: 451–462.

Scheuner, M.T., Sieverding, P. and Shekelle, P.G. (2008) Delivery of genomic medicine for common chronic adult diseases: a systematic review. *JAMA* 299: 1320–1334.

Schmidt, C. (2008) Regulators weigh risks of consumer genetic tests. *Nat Biotechnol* 26: 145–146.

Schmitz-Drager, B.J., Goebell, P.J., Ebert, T. and Fradet, Y. (2000) P53 Immunohistochemistry as a prognostic marker in bladder cancer Playground for urology scientists? *Eur Urol* 38: 691–699; discussion 700.

Sequist, L.V., Bell, D.W., Lynch, T.J. and Haber, D.A. (2007) Molecular predictors of response to epidermal growth factor receptor antagonists in non-small-cell lung cancer. *J Clin Oncol* 25: 587–595.

Sergeant, G., Pennickx, F. and Topal, B. (2008) Quantitative RT-PCR detection of colorectal tumour cells in peripheral blood—a systematic review. *J Surg Res* 150: 144–152.

Seve, P., Lai, R., Ding, K., Winton, K., Butts, C., Mackey, J. *et al.* (2007) Class III beta-tubulin expression and benefit from adjuvant cisplatin/vinorelbine chemotherapy in operable non-small cell lung cancer: analysis of NCIC JBR.10. *Clin Cancer Res* 13: 994–999.

Shah, S.P., Kobel, M., Senz, J., Morin, R.D., Clarke, B.A., Wiegand, K.C. *et al.* (2009) Mutation of Foxl2 in granulosa-cell tumors of the ovary. *N Engl J Med* 360: 2719–2729.

Shedden, K.A., Taylor, J.M., Giordano, T.J., Kuick, R., Misek, D.E., Rennert, G. *et al.* (2003) Accurate molecular classification of human cancers based on gene expression using a simple classifier with a pathological tree-based framework. *Am J Pathol* 163: 1985–1995.

Shelling, A. (2009) Progress in the study of genetic disease: bringing new light to complex problems. *Postgrad Med J* 85.

Shendure, J. (2008) The beginning of the end for microarrays? *Nat Methods* 5: 585–587.

Shi, L., Reid, L.H., Jones, W.D., Shippy, R., Warrington, J.A., Baker, S.C. *et al.* (2006) The Microarray Quality Control (MAQC) Project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat Biotechnol* 24: 1151–1161.

Shih, W., Chetty, R. and Tsao, M.S. (2005) Expression profiling by microarrays in colorectal cancer (review). *Oncol Rep* 13: 517–524.

Simon, R. (2005) Roadmap for developing and validating therapeutically relevant genomic classifiers. *J Clin Oncol* 23: 7332–7341.

Simon, R., Radmacher, M.D., Dobbin, K. and McShane, L.M. (2003) Pitfalls in the use of DNA microarray data for diagnostic and prognostic classification. *J Natl Cancer Inst* 95: 14–18.

Singer, S., Rubin, B.P., Lux, M.L., Chen, C.J., Demetri, G.D., Fletcher, C.D. *et al.* (2002) Prognostic value of kit mutation type, mitotic activity, and histologic subtype in gastrointestinal stromal tumors. *J Clin Oncol* 20: 3898–3905.

Skog, J., Wurdinger, T., Van Rijn, S., Meijer, D.H., Gainche, L., Sena-Esteves, M. *et al.* (2008) Glioblastoma microvesicles transport RNA and proteins that promote tumour growth and provide diagnostic biomarkers. *Nat Cell Biol* 10: 1470–1476.

Sleijfer, S., Gratama, J.W., Sieuwerts, A.M., Kraan, J., Marens, J.W. and Foekens, J.A. (2007) Circulating tumour cell detection on its way to routine diagnostic implementation? *Eur J Cancer* 43: 2645–2650.

Soleymani, L., Fang, Z., Sargent, E.H. and Kelley, S.O. (2009) Programming the detection limits of biosensors through controlled nanostructuring. *Nat Nanotechnol* 4: 844–848.

Solit, D.B., Garraway, L.A., Pratilas, C.A., Sawai, A., Getz, G., Basso, A. *et al.* (2006) Braf mutation predicts sensitivity to Mek inhibition. *Nature* 439: 358–362.

Sorlie, T., Perou, C.M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H. *et al.* (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA* 98: 10869–10874.

Sorlie, T., Perou, C.M., Fan, C., Geisler, S., Aas, T., Nobel, A. *et al.* (2006) Gene expression profiles do not consistently predict the clinical treatment response in locally advanced breast cancer. *Mol Cancer Ther* 5: 2914–2918.

Sotiriou, C. and Piccart, M.J. (2007) Taking gene-expression profiling to the clinic: when will molecular signatures become relevant to patient care? *Nat Rev Cancer* 7(7): 545–553.

Sotiriou, C. and Pusztai, L. (2009) Gene-expression signatures in breast cancer. *N Engl J Med* 360: 790–800.

Sotiriou, C., Wirapati, P., Loi, S., Harris, A., Fox, S., Smeds, J. *et al.* (2006) Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. *J Natl Cancer Inst* 98: 262–272.

Sotiropoulou, G., Pampalakis, G., Lianidou, E. and Mourelatos, Z. (2009) Emerging roles of microRNAs as molecular switches in the integrated circuit of the cancer cell. *RNA* 15: 1443–1461.

Sreekumar, A., Poisson, L.M., Rajendiran, T.M., Khan, A.P., Cao, Q., Yu, J. *et al.* (2009) Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. *Nature* 457: 910–914.

Stein, J.P., Ginsberg, D.A., Grossfeld, G.D., Chatterjee, S.J., Esrig, D., Dickinson, M.G. *et al.* (1998) Effect of P21waf1/Cip1 expression on tumor progression in bladder cancer. *J Natl Cancer Inst* 90: 1072–1079.

Sykacek, P., Clarkson, R., Print, C., Furlong, R. and Micklem, G. (2007) Bayesian modelling of shared gene function. *Bioinformatics* 23: 1936–1944.

Symmans, W.F., Ayers, M., Clark, E.A., Stec, J., Hess, K.R., Sneige, N. *et al.* (2003) Total RNA yield and microarray gene expression profiles from fine-needle aspiration biopsy and core-needle biopsy samples of breast carcinoma. *Cancer* 97: 2960–2971.

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N. *et al.* (2009) mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods* 6: 377–382.

Tavassoli, F.A. (2009) Challenges in breast pathology: new twists on old problems. *Arch Pathol Lab Med* 133: 852–854.

Tenesa, A. and Dunlop, M.G. (2009) New insights into the aetiology of colorectal cancer from genome-wide association studies. *Nat Rev Genet.* Epub 12 May 2009.

Thakur, M.L. (2009) Genomic biomarkers for molecular imaging: predicting the future. *Semin Nucl Med* 39: 236–246.

Thiagalingam, S. (2006) A cascade of modules of a network defines cancer progression. *Cancer Res* 66: 7379–7385.

Thomas, R.K., Nickerson, E., Simons, J.F., Janne, P.A., Tengs, T., Yuza, Y. *et al.* (2006) Sensitive mutation detection in heterogeneous cancer specimens by massively parallel picoliter reactor sequencing. *Nat Med* 12: 852–855.

Tinker, A.V., Boussioutas, A. and Bowtell, D.D. (2006) The challenges of gene expression microarrays for the study of human cancer. *Cancer Cell* 9: 333–339.

Toriello, N.M., Douglas, E.S., Thaitrong, N., Hsiao, S.C., Francis, M.B., Bertozzi, C.R. *et al.* (2008) Integrated microfluidic bioprocessor for single-cell

gene expression analysis. *Proc Natl Acad Sci U S A* 105: 20173–20178.

Tsao, M., Aviel-Ronen, S., Ding, K., Lau, D., Liu, N., Sakurada, A. *et al.* (2007) Prognostic and predictive importance of p53 and Ras for adjuvant chemotherapy in non small-cell lung cancer. *J Clin Oncol* 25: 5240–5247.

Tunuguntla, H.S. and Jorda, M. (2008) Diagnostic and prognostic molecular markers in renal cell carcinoma. *J Urol* 179: 2096–2102.

Vallbohmer, D., Iqbal, S., Yang, D.Y., Rhodes, K.E., Zhang, W., Gordon, M. *et al.* (2006) Molecular determinants of irinotecan efficacy. *Int J Cancer* 119: 2435–2442.

Van Cutsem, E., Kohne, C.H., Hitre, E., Zaluski, J., Chang Chien, C.R., Makhson, A. *et al.* (2009) Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer. *N Engl J Med* 360: 1408–1417.

Van De Vijver, M.J., He, Y.D., Van 't Veer, L.J., Dai, H., Hart, A.A., Voskuil, D.W. *et al.* (2002) A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 347: 1999–2009.

Van Laar, R.K., Ma, X.J., De Jong, D., Wehkamp, D., Floore, A.N., Warmoes, M.O. *et al.* (2009) Implementation of a novel microarray-based diagnostic test for cancer of unknown primary. *Int J Cancer* 125: 1390–1397.

Van 't Veer, L.J., Dai, H., Van De Vijver, M.J., He, Y.D., Hart, A.A., Mao, M. *et al.* (2002) Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415: 530–536.

Venables, J.P. (2006) Unbalanced alternative splicing and its significance in cancer. *Bioessays* 28: 378–386.

Wanebo, H.J., Rao, B., Pinsky, C.M., Hoffman, R.G., Stearns, M., Schwartz, M.K. *et al.* (1978) Preoperative carcinoembryonic antigen level as a prognostic indicator in colorectal cancer. *N Engl J Med* 299: 448–451.

Wang, Y., Klijn, J.G., Zhang, Y., Sieuwerts, A.M., Look, M.P., Yang, F. *et al.* (2005) Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* 365: 671–679.

Wang, Y., Dai, D.L., Martinka, M. and Li, G. (2007) Prognostic significance of nuclear Ing3 expression in human cutaneous melanoma. *Clin Cancer Res* 13: 4111–4116.

Wang, X., Liu, L., Fackenthal, J., Cummings, S., Olopade, O.I., Hope, K. *et al.* (2009a) Translational integrity and continuity: personalized biomedical data integration. *J Biomed Inform* 42: 100–112.

Wang, Z., Gerstein, M. and Snyder, M. (2009b) RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 10: 57–63.

Watanabe, T., Wu, T.T., Catalano, P.J., Ueki, T., Satriano, R., Haller, D.G. *et al.* (2001) Molecular predictors of survival after adjuvant chemotherapy for colon cancer. *N Engl J Med* 344: 1196–1206.

Weigelt, B., Hu, Z., He, X., Livasy, C., Carey, L.A., Ewend, M.G. *et al.* (2005) Molecular portraits and 70-gene prognosis signature are preserved throughout the metastatic process of breast cancer. *Cancer Res* 65: 9155–9158.

West, M., Blanchette, C., Dressman, H., Huang, E., Ishida, S., Spang, R. *et al.* (2001) Predicting the clinical status of human breast cancer by using gene expression profiles. *Proc Natl Acad Sci U S A* 98: 11462–11467.

Winnepenninckx, V., Lazar, V., Michiels, S., Dessen, P., Stas, M., Alonso, S.R. *et al.* (2006) Gene expression profiling of primary cutaneous melanoma and clinical outcome. *J Natl Cancer Inst* 98: 472–482.

Winton, T., Livingston, R., Johnson, D., Rigas, J., Johnston, M., Butts, C. *et al.* (2005) Vinorelbine plus cisplatin vs observation in resected non-small-cell lung cancer. *N Engl J Med* 352: 2589–2597.

Wolmark, N., Fisher, B., Wieand, H.S., Henry, R.S., Lerner, H., Legault-Poisson, S. *et al.* (1984) The prognostic significance of preoperative carcinoembryonic antigen levels in colorectal cancer. Results from NSABP (National Surgical Adjuvant Breast and Bowel Project) Clinical Trials. *Ann Surg* 199: 375–382.

Wuttig, D., Baier, B., Fuessel, S., Meinhardt, M., Herr, A., Hoefling, C. *et al.* (2009) Gene signatures of pulmonary metastases of renal cell carcinoma reflect the disease-free interval and the number of metastases per patient. *Int J Cancer* 125: 474–482.

Xi, L., Feber, A., Gupta, V., Wu, M., Bergemann, A.D., Landreneau, R.J. *et al.* (2008) Whole genome exon arrays identify differential expression of alternatively spliced, cancer-related genes in lung cancer. *Nucleic Acids Res* 36: 6535–6547.

Xu, K., Hou, S. and Du, Z. (2002) Prognostic value of matrix metalloproteinase-2 and tissue inhibitor of metalloproteinase-2 in bladder carcinoma. *Chin Med J* 115: 743–745.

Zheng, Z., Chen, T., Li, X., Haura, E., Sharma, A. and Bepler, G. (2007) DNA synthesis and repair genes Rrmi and Ercc1 in lung cancer. *N Engl J Med* 356: 800–808.

Zhu, C.Q., Da Cunha Santos, G., Ding, K., Sakurada, A., Cutz, J.C., Liu, N. *et al.* (2008) Role of Kras and Egfr as Biomarkers of response to erlotinib in National Cancer Institute of Canada Clinical Trials Group Study Br.21. *J Clin Oncol* 26: 4268–4275.

Zhu, W., Qin, W., Atasoy, U. and Sauter, E.R. (2009) Circulating microRNAs in breast cancer and healthy subjects. *BMC Res Notes* 2: 89.

Zujewski, J.A. and Kamin, L. (2008) Trial assessing individualized options for treatment for breast cancer: the Tailorx trial. *Future Oncol* 4: 603–610.