Ten types of breast cancer?

After years of study of almost 2000 breast cancers, a collaborative research group from Canada and the UK recently reported that they could identify ten distinct types of breast cancer, each associated with a different prognosis. Two broad categories of prognostic factors currently exist: those indicating the level of tumour advancement and those indicating tumour aggressiveness. Established prognostic factors—tumour size and number of affected nodes—which were identified in the 1960s, belong to both categories. These factors underpin the concept of early diagnosis and provide the impetus for screening. Other prognostic features tend to be used to subclassify tumours by their metastatic potential, are generally stable throughout the natural history of the cancer (ie, tumours rarely switch classes), and are the basis of personalised treatment.

From 1960 to 1990, the abundance in breast tumours of three receptor proteins (oestrogen receptor [ER], progesterone receptor [PR], and HER2), when measured by immunohistochemistry and combined with tumour size and number of nodes, predicted prognosis reliably, and helped define response to treatment. In 2000, Charles Perou and colleagues proposed that breast cancers could be better classified than previously by measurement of RNA expression of a large number of genes. Five or more categories were proposed, including luminal A and basal-type cancers. Although cancer researchers widely accept expression-based categories, such as 50-gene set (PAM50) intrinsic subtypes, these categories have not replaced the traditional (and more conveniently measured) receptor-based categories in the cancer clinic.

12 years after the Perou paper, Christina Curtis and colleagues now propose a fourth-generation classification scheme. Genetically, tumours vary substantially because of inherited (germline) and acquired (somatic) variation. Germline variation consists of small point mutations, such as single nucleotide polymorphisms and copy number variants; a copy number variant is an increase or decrease from the modal copy number of two for a given chromosomal segment. Tumours recapitulate the germline genetic fingerprint, but also exhibit acquired genetic variation. Tumour-specific mutations likewise include single nucleotide polymorphisms and copy number aberrations. Gene amplifications (eg, of ERBB2) and deletions are specific classes of aberration. A homozygous deletion involves the complete loss of a genetic segment within the tumour. Curtis and colleagues assumed that genomic variation promoted carcinogenesis by induction of abnormal gene expression (eg, because amplification of ERBB2 leads to overexpression of HER2 in HER2-positive cancers) and despite substantial random variation, if many tumours were studied, discernible patterns would be noted between recurrent genetic anomalies and outliers in gene expression (either up or down). Tumour variability could be identified by incorporation of both the genetic mutation status (tumour genome) and RNA expression levels (transcriptome) of the tumour.

The investigators tested the model by comparing the entire genetic landscape of 1000 cancers with the gene expression profile from the same tumours, and validated their findings in a second group of patients. Several groups of genetic alterations reproducibly resulted in predictable patterns of RNA expression. Of the different types of mutation (inherited and acquired), copy number aberrations accounted for the greatest amount of variability in gene expression between cancers. These aberrations modified the expression of nearby (cis-acting) or more distant genes, either on the same or on a different (trans-acting) chromosome in roughly equal numbers, but the effects of the cis-acting variants were generally more profound. On the basis of data
from cis-acting variants, ten categories emerged; each accounted for 4–17% of cancers. The 10-year survival rates varied from 40% to 90%, dependent on the category. The categories added independent predictive value beyond size, grade, and nodal status (ER, PR, and HER2 status were not included in the model).

Results of this study are of wide scientific interest and are potentially clinically useful. Patients in cluster 3 had a much lower risk of death (13% died) than did those in other clusters (29% died), and some women in this cluster might potentially forego chemotherapy. Ideally, translational studies will allow us to align the type of chemotherapy and the clinical outcome cluster by cluster. Tests based on gene expression, such as that based on the PAM50 categories, provide a probability that the cancer will progress that goes beyond standard immunohistochemistry; however, advice is not given as to the best choice of chemotherapy. Unknown is whether cluster-based categorisation will supplement or supersedes present prognostic models based on the six canonical variables (tumour size, positive lymph nodes, tumour grade, ER, PR, and HER2) or whether the new information will prove useful for treatment choice. In some cases, the new clusters recapitulate (or subdivide) more traditional categories, but in other cases they are substantially different. For example, cluster 4 contains both ER-positive and ER-negative breast cancers and various intrinsic subtypes. Within each cluster, several genetic alterations have been identified and within or near these chromosomal regions might reside one or more driver genes, which could become the natural target for novel treatments. Furthermore, within a given cluster, a rare genomic change might be noted in only a few patients. For example, a small number of patients exhibited amplification of the epidermal growth factor receptor gene, and some had rare homozygous deletions in BRCA2. These rare alterations in the genomic landscape, which are present in fewer than 1% of patients, are prime targets for treatment, although they subdivide the large cluster in which they are contained. If so, the number of relevant categories of breast cancer could exceed ten.

The situation is complicated by tumour heterogeneity. In Nik-Zainal and colleagues’ study, the investigators had to sequence the genome of a breast cancer 188 times to identify the full diversity of mutations, because most of the 70 690 mutations detected were present in only a few cells. 26 762 different mutations, including trisomy 1q and a mutation in TP53, were present in all cancer cells, which is consistent with the assumption that these mutations are among the dominant driver mutations. Three of the extended panel of 21 cancers (which were sequenced to a lesser degree) had gene amplifications of ERBB2; reassuringly, these were among the early ubiquitous mutations (many somatic mutations arose later on within the amplified segment). To avoid selective outgrowth of resistant subclones, therapeutic targets should be a driver mutation that is present in all cancer cells.

Collection of sufficient data to correlate specific clusters, or rare copy number aberrations, with the optimum chemotherapy, and potentially targeted therapy, will be challenging—thousands of patients will have to be followed-up closely for 10 years or more.

Steven A Narod
Women’s College Research Institute, University of Toronto, Toronto, ON M5G 1N8, Canada
steven.narod@wchospital.ca
I declare that I have no conflicts of interest.