

Prognostic immunohistologic markers in human tumors: why are so few used in clinical practice?

Yasodha Natkunam¹ and David Y Mason²

¹Department of Pathology, Stanford University School of Medicine, Stanford, CA, USA and ²Leukemia Research Fund Immunodiagnosics Unit, Nuffield Department of Clinical Laboratory Sciences, John Radcliffe Hospital, University of Oxford, Oxford, UK

Technological advances in gene cloning and genome-wide analyses have greatly increased the number of new tumor markers that can be detected by immunohistologic techniques. While many of these have been evaluated with respect to prognosis, there is a striking discrepancy between the number of markers reported to confer prognostic information and those that are used in clinical practice. We argue that lessons learned from epidemiological studies are applicable to studies of immunohistologic markers; in particular, advances in both fields can be vitiated by non-causal associations. We suggest that the most valuable immunohistologic markers are those that reflect genetic abnormalities, that are linked to the cell of origin, or that reflect tumor infiltrating cells or stromal reactions. It should also be appreciated that a marker that is genuinely predictive of prognosis may nevertheless not find any application in clinical practice if it becomes obsolete through the introduction of newer therapies or because there is no choice of alternative treatment strategies.

Laboratory Investigation (2006) 86, 742–747. doi:10.1038/labinvest.3700447

Keywords: immunohistochemistry; tumor; prognosis; epidemiology; therapy

‘Those who cannot learn from history are doomed to repeat it.’

George Santayana

The pathology literature contains innumerable reports of immunohistologic tumor markers claimed to predict clinical behavior. Such reports are accumulating at an even more rapid pace as a result of the explosive growth in availability of commercial antibodies raised against cloned gene products or synthetic peptides. Unfortunately, this activity has not had great impact in practical terms: pathologists continue to make diagnoses as in the past and prognostic markers are rarely used to modify treatment strategies. While important exceptions, such as the crucial relevance of analyzing estrogen receptor and Her2Neu expression in breast carcinomas, validate the approach, the overall utility of immunohistochemistry in guiding clinical treatment

decisions remains limited. Why then is there such a great disparity between the number of papers reporting novel prognostic markers and their impact on clinical practice?

One problem may be the general approach used: most immunohistologic studies of prognostic markers aim to find, in a more or less random fashion, molecular features *within* the tumor that correlate with clinical behavior. This is reminiscent of the epidemiologist’s quest, again frequently on a random basis, for factors *outside* the individual that are associated with the risk of developing disease. Epidemiological investigations have also been surprisingly unproductive in terms of advancing our understanding of the ‘causes’ of diseases and devising better strategies for their prevention and management. It is therefore instructive to see if lessons already laboriously learned from epidemiology are relevant to current attempts to find immunohistologic predictors of tumor behavior.

Two problems that bedevil epidemiological studies are associations between putative risk factors and disease frequency that are weak (eg increases in risk well under two-fold) and/or difficult to reproduce (often because of technical difficulties in gathering valid data). These two features interact: the weaker the association the greater the possibility

Correspondence: Dr Y Natkunam, MD, PhD, Department of Pathology, L235, Stanford University School of Medicine, 300 Pasteur Drive, Stanford, CA 94305, USA.

E-mail: yaso@stanford.edu

Received 9 May 2006; revised and accepted 11 May 2006

that it has been observed by chance and that it will not be found in a second study. A notorious example of the confusion that can ensue from weak associations is found in the claims that residential proximity to high-tension power lines carries an increased risk of childhood cancer, particularly acute leukemia. Extensive conflicting studies for over 25 years, including a pooled meta-analysis of 5393 leukemia cases and 10704 controls, seemed finally to have confirmed the conclusion, and logical instinct, of many scientists that there is no causal association.^{1,2} But just as it seemed, after all, safe to live near electricity pylons, a new report has reignited the controversy.³

It would be difficult to find anything comparable in terms of expenditure of effort and funds for so little benefit in studies of prognostic immunohistologic markers. However, there are many examples of correlations that are relatively weak, although apparently statistically significant, and reported in one study but not confirmed in another. For example, at least 40 molecular markers have been reported as predictors of prognosis in diffuse large B-cell lymphoma (DLBCL), but many have not been subjected to the litmus test of the epidemiologist: confirmation in a second independent study. When further studies are performed the results often lack reproducibility. For example expression of BCL6, a germinal center-associated marker, correlates with better overall survival in DLBCL in some studies⁴⁻⁶ but not in others.^{7,8} Likewise, several investigations have found that increased expression of BCL2 protein or the proliferation marker Ki-67 (MIB1) correlates with inferior survival in DLBCL,^{4,7,9-15} whereas others have found no impact on outcome.^{6,8,9,16-22} A more recent example is the transcription factor FOXP1, whose expression as detected by immunohistology was found to be associated with adverse outcome in DLBCL patients in two studies^{23,24} but not in a third.²⁵

When discrepancies of this sort are encountered it is often argued that the immunostains were not performed satisfactorily or were not interpreted correctly. These arguments are reminiscent of epidemiologists' disagreements over deficiencies in the collection or statistical handling of data. At best, these conflicting conclusions suggest that the immunohistologic marker is unsuitable for routine use: and, at worst, that the reported correlation may be an artifact.

An even more troublesome problem in epidemiological studies which also has a clear parallel in immunohistologic studies, is posed by associations between external factors and the risk of disease that are reproducible but are non-causal. For example, it is beyond dispute that smoking causes damage to lung epithelium and increases the risk of carcinoma. However, there is also a clear statistical association between smoking and increased risk of suicide; but nobody would argue that this represents an underlying causal link (Figure 1a). The epidemiological

literature contains many examples of such associations that are difficult, or impossible, to rationalize. As an example, Table 1 lists a number of risk factors reportedly associated with heart disease (selected from among a total of 246 identified in a review 25 years ago),²⁶ most of which are highly unlikely to reflect causal links. The same risk of detecting non-causal associations exists when attempting to correlate immunohistologic tumor markers with clinical outcome. For example, Figure 1b shows how a molecule that tends to be expressed in proliferating cells might appear to be responsible for poor prognosis. However, both features are probably secondary consequences: it is unlikely that expression of the molecule itself directly drives tumor growth.

The 'success stories' of epidemiology can also be instructive. For example, many of the associations that have stood the test of time involve a plausible causal scenario beginning with a specific factor, for example a virus or a mutagenic chemical agent that triggers a cascade of intermediate events, such as DNA damage and dysplasia, leading to neoplastic transformation. These causal 'narratives', tracing a pathway from an identifiable agent to increased tumor risk, are not always immediately evident: for example, the risk of scrotal cancer in child chimney-sweeps was identified in the 18th century, many years before causal mechanisms could be proposed. Nevertheless the guiding principle is clear: one should always seek convincing explanatory models; schemes with poorly defined starting points, such as the vague description 'living near an electricity pylon', rarely open fruitful new avenues for understanding the origins of neoplasia.

The implication for immunohistologic studies is that tumor markers of genuine prognostic value often reflect acquired genetic abnormalities in the neoplastic cell that unleash downstream tumorigenic events. Epidemiological studies have uncovered a number of oncogenic microorganisms that induce a risk of tumor development, as illustrated by the link between human papilloma virus and cervical carcinoma. This finds a clear parallel in the acquired genetic abnormalities that dictate tumor behavior (eg deregulation of the cellular equivalent of viral oncogenes such as Myc or Ras). In effect these acquired oncogenic alterations can be considered the endogenous equivalent of external tumor-inducing factors. For this reason many immunohistologic markers that genuinely correlate with prognosis reflect acquired genetic lesions. Examples are listed in Table 2. It may be added that the study of prognostic markers related to genetic lesions can raise complex questions concerning the classification of tumors, and, in particular, whether a genetic lesion should be the diagnostic criterion for a disease. For this reason, it can happen on occasion that a marker is promoted, as data accumulate, from the status of prognostic or diagnostic marker to become the defining feature of a

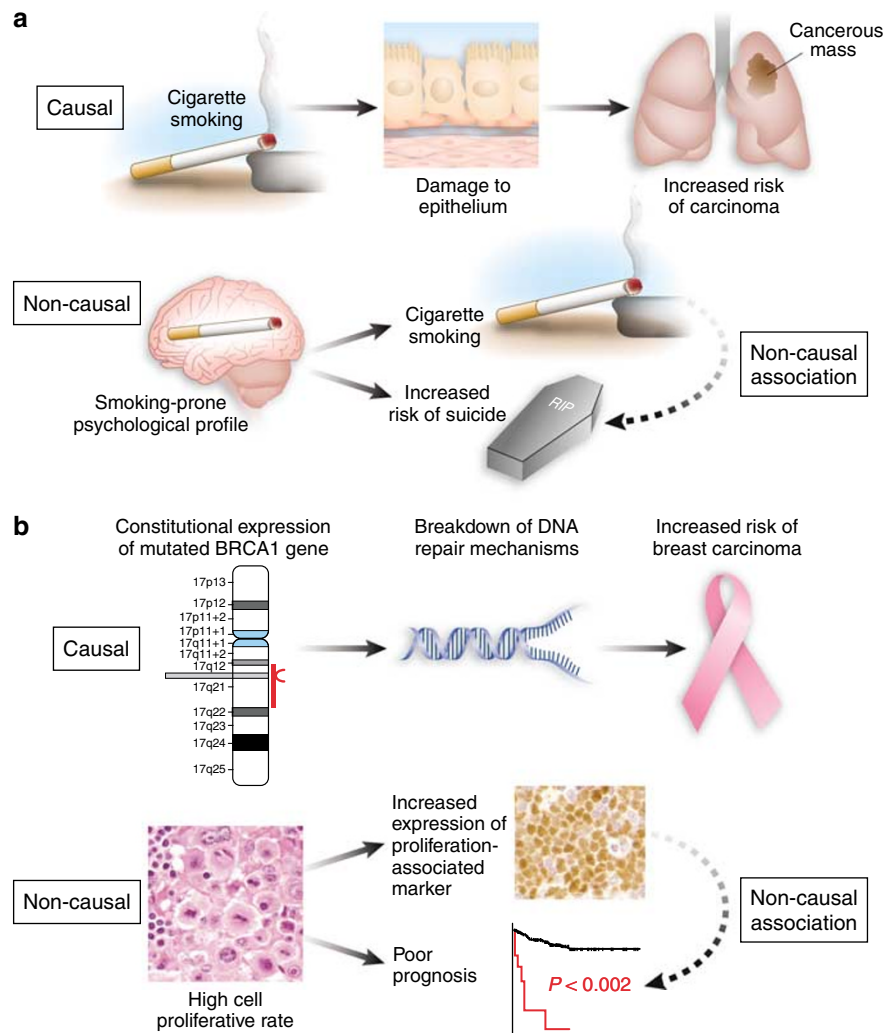


Figure 1 This scheme illustrates how non-causal associations between risk factors and tumor development can be confused with causal links (a). This is paralleled by studies of prognostic tumor markers (b). The distinction between causal and non-causal associations becomes evident in the example shown in (a) if the consequences of stopping smoking are considered: the risk of carcinoma will diminish, but the risk of suicide will stay the same (or might well rise!).

Table 1 Risk factors for heart disease

| | | |
|-----------------------|--------------------------|-------------------|
| Age | Jewish religion | Short stature |
| Alcoholism | Low socioeconomic status | Slow beard growth |
| Baldness | Male sex | Snoring |
| Coffee | No garlic | Too much milk |
| Divorced parents | Noise | Too little milk |
| Extramarital sex | Nonsupportive boss | Total abstinence |
| Good financial status | Not eating mackerel | Un-loving wife |
| Intelligent wife | Not being a Mormon | Vasectomy |

This table is adapted from a review by Hopkins *et al.*²⁶

neoplasm. For example, it is now uncommon to make the diagnosis of a gastrointestinal stromal tumor if it lacks c-KIT (CD117) expression.

Immunocytochemical abnormalities that reflect the presence of acquired genetic alterations are not

the only category of valid prognostic markers, and here data from microarray-based profiling of tumors provide some useful clues. These studies have defined abnormal expression patterns secondary to genetic alterations that correlate with clinical behavior, but they have also drawn attention to three other prognostically relevant categories of markers.

Firstly, several gene expression studies have identified patterns that are of prognostic significance because they reveal the cell of origin of the tumor. For example, a poor prognosis subtype of breast carcinoma arising from basal cells can be defined on the basis of gene expression.²⁷ Similarly, molecular profiling in DLBCL has provided evidence for prognostically diverse subtypes that appear to derive from different stages of B-cell differentiation.²⁸ In fact, this is no surprise: the entire process of categorizing lymphomas is largely based on immunohistologic markers of maturation and lineage, and the very different clinical behavior

Table 2 Examples of prognostically significant immunocytochemical markers that are indicative of an underlying genetic abnormality

| <i>Immunocytochemical marker</i> | <i>Underlying genetic abnormality</i> | <i>Comment</i> |
|---|--|--|
| BCL2 in germinal centre cells | t(14;18) | Suppression of apoptosis leading to prolonged survival of neoplastic cells |
| Cytoplasmic NPM in AML | C terminal mutations in gene | Associated with relatively good prognosis. The mechanism is unknown but may involve disrupt NPM interactions with nuclear molecules such as p53 |
| Cyclin D1 in small cell lymphoma | t(11;14) | Disruption of cell cycle leading to a lymphoma of worse prognosis than other small cell lymphomas |
| Strong cytoplasmic phosphotyrosine in lymphoma | Lymphomas with this feature carry a translocated <i>ALK</i> gene | Activation of this kinase induces abnormal intracellular cell signaling, and the resulting tumor is more chemosensitive compared with T-cell lymphomas that arise through different mechanisms |
| High ZAP70 expression in CLL | Not certain | Tends to correlate with unmutated Ig V region genes, and such cases for unknown reasons, have a worse prognosis |
| Altered distribution of PML in acute promyelocytic leukemia | Indicative of PML:RAR α fusion | This anomaly defines a subtype of leukemia that has special clinical features. However, the resulting disruption of PML from nuclear bodies can be reversed upon treatment with retinoic acid |
| Immunohistologic pattern of nuclear localization of BCL10 in extranodal MALT lymphoma | t(11;18) or t(1;14) | Leads to refractory disease resistant to Helicobacter pylori eradication treatment |
| Strong staining for ErbB2 (Her2Neu) | Amplification of gene | Disruption of signal transduction and normal cell cycle control leading to aggressive tumor growth |

of individual lymphoid neoplasms is believed to reflect their derivation from distinct lymphoid cell subpopulations. This also raises (as for genetic lesions) questions of disease definition: for example the expression of terminal deoxynucleotidyl transferase (TdT) could be considered a marker of prognosis within the spectrum of lymphoid neoplasms, but in practice it is recognized as defining the clinically distinctive lymphomas that arise from precursor B- and T cells. It is possible that a variety of prognostically valid markers reflect a tumor's cellular origin but are not at present recognized as such for want of evidence.

Secondly, gene expression profiling studies have provided evidence for the impact of stromal cell reactions on tumor behavior. For example, prognosis in breast cancer correlates with the expression of genes comprising a 'wound-response signature',²⁹ and with the expression of extracellular matrix genes and growth factors.³⁰ Tumor-infiltrating cellular elements involved in innate and adaptive immunity (eg macrophages and cytotoxic lymphocytes) have also been implicated in the modulation of tumorigenesis.^{31–33} These findings raise the possibility that immunohistologic analysis of background cellular infiltrate/stromal reactions might be used to predict prognosis. For example, increased numbers of non-neoplastic macrophages and a combination of low numbers of regulatory T cells and increased cytotoxic T cells appear to predict for

inferior survival in follicular and Hodgkin lymphomas.^{33,34} It may seem paradoxical that cytotoxic T cells, which should combat the tumor, correlate with poor prognosis, but this is consistent with studies of diffuse large B cell and Hodgkin lymphomas,^{32,33} in which tumors that contain higher numbers of cytotoxic T cells fare worse. It therefore seems likely that increasing attention will focus in the future on the prognostic implications of the background non-neoplastic component in tumors, and these clearly lend themselves well to immunohistologic analysis. It must be added, however, that robust guidelines for reproducibly evaluating and quantifying these non-neoplastic cell populations will need to be established.

Thirdly, microarray-based gene expression studies can occasionally identify genes whose overexpression, even if it is not clearly associated with either a genetic alteration or with cellular origin, is empirically shown to correlate with survival. Immunohistologic detection of the corresponding protein might then allow these findings to be introduced in the diagnostic laboratory. ZAP70 in chronic lymphocytic leukemia is one example, albeit a somewhat controversial one, of a marker that has emerged in this way. However, such examples are rare and the chances of finding one by random screening for markers whose expression correlates with survival is low. It can be argued that immunohistologic studies should follow the lead of gene expression

studies and screen large numbers of markers, but this is to ignore the practical reality that even the most ambitious immunohistologic project can only cope with tens of antibodies. Thus, the study of numerous randomly chosen immunohistologic markers increases the workload without a commensurate increase in the chance of success.

Finally, even the most powerful markers of prognosis can fall at one of two final hurdles, both relating to the practical realities of treatment. Firstly, advances in management can obviate the need for stratifying patients. For example, Hodgkin lymphoma currently has such a good outcome that it could be difficult to define prognostic markers that would have wide utility. Similarly, it has been reported that survival differences associated with BCL2 and BCL6 protein expression in patients with DLBCL are no longer seen when the anti-CD20 antibody rituximab is added to anthracyclin-based chemotherapy.^{6,35} Secondly, the practical relevance of a marker of probable tumor behavior is related to the strength of its predictive power and the available treatment. A marker associated with only a 25% difference in long-term survival in a disease for which there is essentially only a single treatment option will probably not find any practical application. Conversely, the clinical relevance is obvious for a marker that defines a subgroup of patients with a very good prognosis in a disease for which the conventional treatment carries a high risk of mortality/morbidity. The same would also be true if a very poor prognosis patient subgroup for which more heroic treatment might be justifiable can be identified on the basis of marker expression.

These factors explain further why so many reported prognostic markers have failed to appear in the clinic. One positive aspect is that the number of valuable markers can only increase, as it becomes possible to detect more cellular molecules in tissue biopsies. However, in designing strategies for identifying prognostic markers it may be beneficial to learn from the lessons of the many epidemiological studies of the past. Ultimately the search for prognostic tumor markers should be no more immune to cost-benefit analysis than any other field of research.

...the guiding principle is clear:
one should always seek convincing
explanatory models...

References

- 1 Exposure to power-frequency magnetic fields and the risk of childhood cancer. UK Childhood Cancer Study Investigators. *Lancet* 1999;354:1925–1931.
- 2 Ahlbom A, Day N, Feychting M, *et al.* A pooled analysis of magnetic fields and childhood leukaemia. *Br J Cancer* 2000;83:692–698.

- 3 Draper G, Vincent T, Kroll ME, *et al.* Childhood cancer in relation to distance from high voltage power lines in England and Wales: a case-control study. *BMJ* 2005; 330:1290.
- 4 Hans CP, Weisenburger DD, Greiner TC, *et al.* Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. *Blood* 2004;103:275–282.
- 5 Lossos IS, Jones CD, Warnke R, *et al.* Expression of a single gene, BCL-6, strongly predicts survival in patients with diffuse large B-cell lymphoma. *Blood* 2001;98:945–951.
- 6 Winter JN, Weller EA, Horning SJ, *et al.* Prognostic significance of Bcl-6 protein expression in DLBCL treated with CHOP or R-CHOP: a prospective correlative study. *Blood* 2006;107:4207–4213.
- 7 Barrans SL, Carter I, Owen RG, *et al.* Germinal center phenotype and bcl-2 expression combined with the International Prognostic Index improves patient risk stratification in diffuse large B-cell lymphoma. *Blood* 2002;99:1136–1143.
- 8 Takeshita M, Iwashita A, Kurihara K, *et al.* Histologic and immunohistologic findings and prognosis of 40 cases of gastric large B-cell lymphoma. *Am J Surg Pathol* 2000;24:1641–1649.
- 9 Colomo L, Lopez-Guillermo A, Perales M, *et al.* Clinical impact of the differentiation profile assessed by immunophenotyping in patients with diffuse large B-cell lymphoma. *Blood* 2003;101:78–84.
- 10 Gascoyne RD, Adomat SA, Krajewski S, *et al.* Prognostic significance of Bcl-2 protein expression and Bcl-2 gene rearrangement in diffuse aggressive non-Hodgkin's lymphoma. *Blood* 1997;90:244–251.
- 11 Hermine O, Haioun C, Lepage E, *et al.* Prognostic significance of bcl-2 protein expression in aggressive non-Hodgkin's lymphoma. Groupe d'Etude des Lymphomes de l'Adulte (GELA). *Blood* 1996;87:265–272.
- 12 Kramer MH, Hermans J, Wijburg E, *et al.* Clinical relevance of BCL2, BCL6, and MYC rearrangements in diffuse large B-cell lymphoma. *Blood* 1998;92: 3152–3162.
- 13 Muris JJ, Meijer CJ, Vos W, *et al.* Immunohistochemical profiling based on Bcl-2, CD10 and MUM1 expression improves risk stratification in patients with primary nodal diffuse large B cell lymphoma. *J Pathol* 2006; 208:714–723.
- 14 Grogan TM, Lippman SM, Spier CM, *et al.* Independent prognostic significance of a nuclear proliferation antigen in diffuse large cell lymphomas as determined by the monoclonal antibody Ki-67. *Blood* 1988;71: 1157–1160.
- 15 Miller TP, Grogan TM, Dahlberg S, *et al.* Prognostic significance of the Ki-67-associated proliferative antigen in aggressive non-Hodgkin's lymphomas: a prospective Southwest Oncology Group trial. *Blood* 1994; 83:1460–1466.
- 16 Harada S, Suzuki R, Uehira K, *et al.* Molecular and immunological dissection of diffuse large B cell lymphoma: CD5+, and CD5– with CD10+ groups may constitute clinically relevant subtypes. *Leukemia* 1999;13:1441–1447.
- 17 Hill ME, MacLennan KA, Cunningham DC, *et al.* Prognostic significance of BCL-2 expression and bcl-2 major breakpoint region rearrangement in diffuse large cell non-Hodgkin's lymphoma: a British National Lymphoma Investigation Study. *Blood* 1996;88: 1046–1051.

- 18 Sanchez E, Chacon I, Plaza MM, *et al.* Clinical outcome in diffuse large B-cell lymphoma is dependent on the relationship between different cell-cycle regulator proteins. *J Clin Oncol* 1998;16:1931–1939.
- 19 Jerkeman M, Aman P, Cavallin-Stahl E, *et al.* Prognostic implications of BCL6 rearrangement in uniformly treated patients with diffuse large B-cell lymphoma—a Nordic Lymphoma Group study. *Int J Oncol* 2002;20:161–165.
- 20 Saez AI, Saez AJ, Artiga MJ, *et al.* Building an outcome predictor model for diffuse large B-cell lymphoma. *Am J Pathol* 2004;164:613–622.
- 21 Seki R, Okamura T, Koga H, *et al.* Prognostic significance of the F-box protein Skp2 expression in diffuse large B-cell lymphoma. *Am J Hematol* 2003;73:230–235.
- 22 Zhang A, Ohshima K, Sato K, *et al.* Prognostic clinicopathologic factors, including immunologic expression in diffuse large B-cell lymphomas. *Pathol Int* 1999;49:1043–1052.
- 23 Banham AH, Connors JM, Brown PJ, *et al.* Expression of the FOXP1 transcription factor is strongly associated with inferior survival in patients with diffuse large B-cell lymphoma. *Clin Cancer Res* 2005;11:1065–1072.
- 24 Barrans SL, Fenton JA, Banham A, *et al.* Strong expression of FOXP1 identifies a distinct subset of diffuse large B-cell lymphoma (DLBCL) patients with poor outcome. *Blood* 2004;104:2933–2935.
- 25 Hans CP, Weisenburger DD, Greiner TC, *et al.* Expression of PKC-beta or cyclin D2 predicts for inferior survival in diffuse large B-cell lymphoma. *Mod Pathol* 2005;18:1377–1384.
- 26 Hopkins PN, Williams RR. A survey of 246 suggested coronary risk factors. *Atherosclerosis* 1981;40:1–52.
- 27 Sorlie T. Molecular portraits of breast cancer: tumour subtypes as distinct disease entities. *Eur J Cancer* 2004;40:2667–2675.
- 28 Alizadeh AA, Eisen MB, Davis RE, *et al.* Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 2000;403:503–511.
- 29 Chang HY, Nuyten DS, Sneddon JB, *et al.* Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. *Proc Natl Acad Sci USA* 2005;102:3738–3743.
- 30 West RB, Nuyten DS, Subramanian S, *et al.* Determination of stromal signatures in breast carcinoma. *PLoS Biol* 2005;3:e187.
- 31 de Visser KE, Eichten A, Coussens LM. Paradoxical roles of the immune system during cancer development. *Nat Rev Cancer* 2006;6:24–37.
- 32 Muris JJ, Meijer CJ, Cillessen SA, *et al.* Prognostic significance of activated cytotoxic T-lymphocytes in primary nodal diffuse large B-cell lymphomas. *Leukemia* 2004;18:589–596.
- 33 Alvaro T, Lejeune M, Salvado MT, *et al.* Outcome in Hodgkin's lymphoma can be predicted from the presence of accompanying cytotoxic and regulatory T cells. *Clin Cancer Res* 2005;11:1467–1473.
- 34 Farinha P, Masoudi H, Skinnider BF, *et al.* Analysis of multiple biomarkers shows that lymphoma-associated macrophage (LAM) content is an independent predictor of survival in follicular lymphoma (FL). *Blood* 2005;106:2169–2174.
- 35 Mounier N, Briere J, Gisselbrecht C, *et al.* Rituximab plus CHOP (R-CHOP) overcomes bcl-2-associated resistance to chemotherapy in elderly patients with diffuse large B-cell lymphoma (DLBCL). *Blood* 2003;101:4279–4284.