To bury TNM classification in breast cancer or to praise it?

Hervé Bonnefoi

Department of Medical Oncology, Université Victor Segalen, Bordeaux, and Institut Bergonié, Bordeaux, France

Introduction

Fifty years ago the International Union Against Cancer (UICC) tumour-node-metastasis (TNM) Committee defined the aims of cancer staging as: (i) to give some indication of prognosis; (ii) to aid the clinician in planning cancer treatment; (iii) to assist in evaluating the results of treatment; (iv) to facilitate the exchange of information between treatment centres; and (v) to contribute to continuing investigation of human malignancies [1]. In breast cancer, TNM classification has allowed this very heterogeneous disease to be categorised and randomised trials to be conducted in well-defined groups (for example, node-negative and node-positive early breast cancers, locally advanced or inflammatory disease, metastatic disease). TNM classifications were defined in the era of locoregional treatment of cancer. Although the role of systemic treatment has become much more clearly defined over the past two decades, such that all but the very few patients with an excellent prognosis now receive both local and systemic therapy for their breast cancer, the last edition of TNM classification did not take this change into consideration [2]. We are now in the era of combined modality treatment, with systemic treatments added to locoregional ones, since the addition of hormonal therapy, chemotherapy and anti-HER2 treatment in early breast cancer substantially decrease the risk of death [3–5]. However these treatments have side-effects and are costly. Therefore, before treating a patient with these therapies, an accurate assessment of the risk of relapse (prognostic factors) and an evaluation of the chance of response of a patient to this specific treatment (predictive factors) are necessary. As a consequence, the first two aims of the TNM classification, which have served our needs well for 50 years, need to be updated. The first aim should be 'to accurately define the prognosis' instead of 'to give some indication of prognosis' and the second aim becomes 'to aid the clinician in planning cancer treatment using factors predictive of response to both locoregional (for example, radiotherapy) and systemic treatment'. The second paradigm change is that we have moved into the era of modern molecular biology, and therefore we should make it a priority to categorise and treat breast cancer using a molecular classification that predicts clinical behaviour.

In order to avoid the next TNM classifications to becoming obsolete immediately they are published (the 7th edition is to be published in January 2009 and the 8th edition in January 2016) we need to take these important changes into consideration. This review details three proposals to be taken in consideration when defining the next TNM classifications. Recent publications and ongoing trials that we believe should have an impact on these classifications are mentioned.

First proposal: include validated and powerful prognostic factors in the classification

As a consequence of screening and increased awareness about the disease in the population, the presentation of breast cancer has changed during the last decade [6]. This has important practical implications, since the majority of tumours in western countries are now diagnosed as cT1 tumours. However, the prognosis of these T1 tumours is still very heterogeneous, even when for tumours of 1 cm or less in diameter. As an example the prognosis of node-negative tumours of 1 cm in diameter or less (without systemic therapy) varies from 25% risk of relapse at 8 years in some series [7] to less than 10% in others series [8]. It is now widely accepted that predicting the prognosis of breast cancer is much more complex than the simple rule that the smaller the tumour (or the earlier the diagnosis) the better the prognosis. Preclinical data suggest that some tumours can metastasise very early in their development and this explains why using the T stage does not capture the entire prognostic information. There is an urgent need better to understand the biology and metastatic potential of breast cancers in order to include these

findings in a modern classification. The challenge is therefore how to improve the prognostic prediction of breast cancer.

A first approach would be to integrate classic independent prognostic factors that are not part of the TNM classification, such as the grade or the age. The Nottingham Prognostic Index (NPI), first described in 1982, is compiled from grade, size and lymph node status of the primary tumour [9]. The NPI has been prospectively validated both intra- and inter-centre [10–12]. More recently, a numerical tool called Adjuvant! provides an averaged estimate of the risk for an individual [13]. This estimate is based on information that is part of the TNM system, with the addition of the following parameters: grade, age, oestrogen receptor status and co-morbidities. HER2 status will be included soon in the new version of this program. The risk estimate provided by Adjuvant! has been independently validated by others [14,15]. In addition to prognostication, the program provides an averaged estimate of the benefit from different treatment modalities, to give an estimate of risk of breast cancer-related death and risk of relapse at a 10-year follow-up. Estimates of mortality are derived from the Surveillance, Epidemiology, and End-Results (SEER) registry estimates of outcome for breast cancer patients in the general population in the United States of America (USA). There is still a lot of controversy about the addition of grade as a prognostic factor because of the poor reproducibility of this data, and the use of this factor as part of a prognostic index (NPI) or of a programme (Adjuvant!) will not overcome this problem. For this reason, a Breast Task Force representing the American Joint Committee on Cancer decided recently not to add the grade to the new TNM classification [16]. Microarray methods may overcome the reproducibility issue of histological grade and replace the histological grade assessment in the near future with a genomic grade. The Bordet group has recently identified a genomic grade signature by comparing expression profiles of histological grade 1 and grade 3 tumours [17]. This signature was used to calculate a genomic grade index (GGI). When this index was used in a validation dataset of 570 tumours, the GGI was able to separate the histological grade 2 group ($n = 216$) into two categories with distinctive risks of recurrence (hazard ratio (HR) = 3.61, 95% confidence interval (CI) 2.25–5.78; $P < 0.001$). Moreover, in a multivariate analysis, only GGI, lymph node status and tumour size remained statistically significant, the GGI being the strongest prognostic factor (HR = 1.99).

A second approach in order to improve the prognostic prediction of the TNM classification would be to integrate modern prognostic factors. In 2001 Issacs and colleagues reviewed more than 100 new prognostic factors reported these last 10–15 years [18]. Many were 'statistically significant' but very few (if any) qualify for a clinically useful tool, which can be defined as a marker predicting for a relapse associated with a HR of 2 or more [19]. However, since 2001 four prognostic factors have been reported with a potential clinical impact. These markers are reviewed briefly below.

Bone marrow metastasis

The prognostic value of bone marrow metastasis was reported recently in a pooled analysis of 4703 patients with node-negative and node-positive tumours [20]. The presence of tumour cells in the bone marrow was an independent factor associated with the risk of death during the first 5 years of follow-up (HR 1.81) and after 5 years (HR 1.58). However, in the subgroup of patients who did not receive systemic adjuvant therapy this difference does not appear as important.

UPA/PA1 overexpression

UPA/PA-1 overexpression assessed by enzyme-linked immunosorbent assay (ELISA) is a strong independent prognostic factor and qualify for level of evidence 1 in node-negative tumours [21,22]. However this ELISA test requires >300 mg of frozen tissue, which might be problematic with small tumours. Trials using smaller quantities are ongoing.

Multiple gene signatures using cDNA microarrays

The prognostic value of several multiple gene expression patterns, or 'signatures', using cDNA arrays were reported recently. In a case control study (78 node-negative patients) the Amsterdam group identified a 70-gene signature as an independent predicting factor for the risk of early distant metastasis (HR = 18) [23]. The prognosis value of the 70-gene signature was confirmed in a larger series of tumours that included 295 patients (151 node-negative tumours including 61 already included in the case control trial and 144 node-positive tumours) [24]. At 10 years the distant metastasis-free survival for node-negative tumours was 44.1% in patients with the bad profile and 90% in patients with the good profile. In the multivariate analysis the 70-gene signature was the

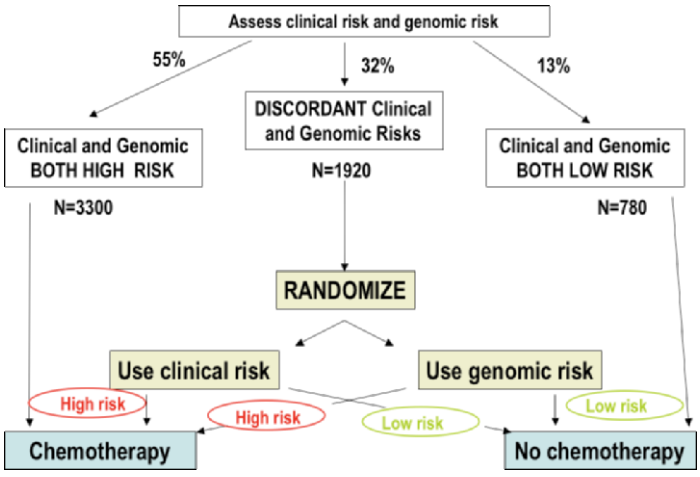

Fig. 1. Design of MINDACT trial.

strongest prognostic factor with a HR of 4.6 (95% CI 2.3–9.2; $P < 0.001$). A multicentre validation trial with 307 node-negative patients was then performed by the TRANSBIG consortium and confirmed the independent prognostic value of the 70-gene signature [25]. Another signature using cDNA arrays has been reported with similar prognosis value by the Rotterdam group [26] and validated using independent data from four European centres [27]. These trials are level 3 of evidence. In an attempt to get to level 1 of evidence the European Organisation on Research and Treatment of Cancer (EORTC) and the Breast International Group (BIG) have recently activated a large prospective randomised trial (Microarray In Node-negative Disease may Avoid ChemoTherapy; MINDACT; Fig. 1). An amendment is in preparation to allow for inclusion of patients with node-positive tumours in this trial.

Multiple gene signatures using RT-PCR

Another multiple gene approach has been developed in the USA using multiple reverse-transcriptase polymerase chain reaction (RT-PCR) assays to quantify expression of several genes in formalin-fixed paraffin-embedded tissue. First a gene signature was developed. These genes were selected from 250 candidate genes. Their correlation with the likelihood of distant recurrence was studied in a series of 447 patients with node-negative oestrogen-positive breast cancer treated with tamoxifen (most were included in NSABP trial B-20). A panel of 16 cancer-related genes and 5 reference genes were selected to create a signature, design an algorithm and create a 'recurrence score' (low, intermediate and high scores). This signature was then tested (some would say validated) in a series of 668 node-negative patients treated with tamoxifen alone in NSABP trial B-14 [28]. The rate of distant recurrence was 6.8% in the low recurrence score group, 14.3% in the intermediate score group and 30.5% in the high recurrence score group ($P < 0.001$). Interestingly, when the recurrence score was added to age and tumour size in a multivariate Cox model, the recurrence score remains the only independent prognostic factor with a hazard ratio of 3.21 (95% CI: 2.23–4.61; $P < 0.001$). The National Surgical Adjuvant Breast and Bowel Project (NSABP) updated this trial, adding patients included in the other arms of trial B-14 (no systemic treatment) and trial B-20 (MF (methotrexate, fluorouracil) or CMF (methotrexate, fluorouracil, cyclophosphamide) chemotherapy) [29]. Thus the 21-gene signature was analysed as a prognostic factor and a predictive factor. The results suggest that the high recurrence score group is also predictive of response to chemotherapy (with little or no benefit of chemotherapy in the low and intermediate groups). On the other hand the benefit of tamoxifen in trial B-14 was almost entirely restricted to the low and intermediate risk categories. A prospective trial (Trial Assigning IndividuaLized Options for Treatment (Rx); TAILORX; Fig. 2) in the context of the Program for the Assessment of Cancer Clinical Tests (PACTT) is now ongoing in the USA conducted by The Breast Cancer Intergroup (TBCI). Patients with node-negative oestrogen-positive tumours are eligible for this trial. Those with a low recurrence score are treated with endocrine therapy alone. Those with high recurrence score and an intermediate recurrence score are randomised to receive either endocrine therapy alone or endocrine therapy plus chemotherapy. Those with a high-risk score are randomised in different trials comparing chemotherapy regimens.

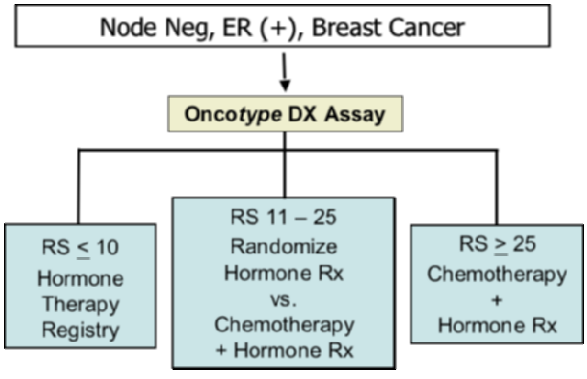

Fig. 2. Design of TAILORX trial.

Second proposal: include predictive factors of response to locoregional and systemic treatment

Predictive factors of endocrine responsiveness

In January 2005 the panellists during the St Gallen consensus meeting made a fundamental change in the pragmatic algorithm for selection of adjuvant systemic therapy when they agreed to first consider the (potential) endocrine responsiveness of the tumour rather than the assessment of the overall risk of relapse [30]. This change was maintained during the 2007 meeting. Three categories were defined: the endocrine responsive group; the endocrine non-responsive group; and tumours of uncertain endocrine responsiveness. The second step of the algorithm was to classify these three groups according to risk into low-, intermediate- and high-risk categories. As far as planning systemic treatment is concerned, this 'classification' led to the following recommendations: avoid chemotherapy in the endocrine responsive group (with the exception of the high risk category) and choose longer or modern chemotherapy regimens in the endocrine non-responsive group. The main weakness of this approach lies in the definition of the uncertain endocrine responsiveness group. New predictive factors including multigene signatures may lead in the near future to improved definition of this group and better identification of patients who need the addition of adjuvant chemotherapy to endocrine therapy. This question is currently being addressed in the subgroup of patients with intermediate recurrence score included in the TAILORX trial (patients with intermediate RS are randomised to receive either endocrine therapy alone or endocrine therapy plus chemotherapy).

Predictive factors of response to anti-HER2 therapy

Considering the striking benefit of trastuzumab demonstrated in four adjuvant trials [4,5,31–33], there is no doubt that HER2 determination (either protein expression or gene amplification) should be part of the next TNM classification.

Predictive factors of response to specific chemotherapy regimens

Patients with T2 or T3 tumours who achieve a pathological complete response (pCR) after neoadjuvant chemotherapy have a 10% risk of relapse at 5 years, which is very different from the 'classic' 30–60% risk of relapse of the whole population of patients with T2–T3 tumours even when treated with optimal neo-adjuvant or adjuvant chemotherapy. Thus, pCR is considered a surrogate for chemosensitivity. However, accurate predictive biological markers of response to chemotherapy are lacking. Recently, several groups have reported promising results with the use of multiple gene signatures [34–37], but these results need to be confirmed. Interestingly, Rouzier and colleagues have addressed the question of the chemosensitivity of breast cancer according to different molecular subtypes identified by the Stanford group [38]. They found that the basal-like and erbB2 subgroups are more sensitive to chemotherapy (45% pCR) than the luminal and normal-like subgroups (6% and 0% pCR, respectively), confirming some earlier studies that have also suggested that oestrogen receptor (ER)-negative tumours have a higher pathological response rate to chemotherapy [39,40]. These data suggest that a breast cancer classification based on molecular subtypes identifies different subgroups of breast cancer with potential clinical consequences (at least as far response to chemotherapy is concerned). These perspectives are discussed in the next section.

Third proposal: categorise breast cancer using a molecular classification

DNA microarrays have made a significant contribution to classifying tumour samples into groups that can predict clinical behaviour [41]. Unsupervised statistical analysis of tumour samples has led to expression-based classifications for many types of cancer, including breast cancer. Ductal invasive carcinomas can be divided into at least five molecular subtypes (luminal A-B-C, basal-like, normal breast-like and ERBB2 types) first identified by Perou and colleagues in 2000 [42] and confirmed repetitively

by the same group [43,44] and others [45]. In addition, these molecular subtypes were shown to predict distant metastasis-free survival and/or overall survival. However, the number of cases whose survival was analysed according to molecular subtypes was small: 49 cases in the first publication by Sorlie [43], 97 and 72 in the second publication [44] and 99 in the Sotiriou publication [45], such that we cannot be entirely confident that the classification is optimal. Two recent publications have identified a new subtype of oestrogen-negative breast cancers which are androgen-receptor positive [46,47]. These two publications are a good example of the utility of microarrays to identify new targets that could be used for new therapies. These data suggest that a molecular classification should replace the TNM classification to give better prediction of clinical behaviour and of response to therapy. However, before we can bury the TNM classification, research programmes need to concentrate on the following objectives:

(1) to confirm the relationship between these different molecular subtypes and survival. The data available so far are based on small series of patients. A combined analysis of the data which are in the public domain would be a first step in the era of 'in silico oncology'. The second step would be to confirm these results using array data generated in the context of ongoing large prospective trials (for example MINDACT trial); (2) to better define molecular subgroups. There is nearly a consensus regarding two subtypes identified in different series, the basal-like group and the luminal A group. However, there is still a lot of debate about other groups. The definition of the luminal B and C groups is unclear. The definition of the HER2 amplified group is for many unsatisfactory. Data from Perou suggest that this group is part of the oestrogen-negative category with an amplicon located on chromosome 17 including the oncogene HER-2. However, the cell type of origin of this subgroup is undefined, and it is clear that amplification of HER-2 gene can also occur in cancers otherwise classified as belonging to the luminal group; (3) to develop methods, whether based on RT-PCR or immunohistochemistry assays that can be used in paraffin-embedded tumour tissue in diagnostic laboratories around the world, in order to provide tumour classifications by molecular subtypes for all patients. For example, Perou's group have analysed the basal-like subtype by immunohistochemistry using antibodies against the breast basal cell keratins 5/6 and 17 [42], and confirmed that all molecular 'basal-like' tumours showed staining for either keratins 5/6 or 17 or both and were negative for oestrogen receptors. Without the ability to reproduce a consistent molecular classification in routine practice, we will have the same problem as we currently have with grade, where there is clear recognition of its prognostic power, but an inability to depend on it in practice.

Conclusion

The landscape of breast cancer has changed in the last 10 years: breast cancer mortality is decreasing in the western world, patients present with smaller tumours, systemic treatment is standard for almost all patients, but the important risk of over-treatment remains, and new molecular biology techniques have created a new taxonomy. Despite these changes, the aims of the TNM classification defined 50 years ago are still valid, but need to be updated or rephrased to include powerful prognostic and predictive factors of response to systemic therapies. We can expect that the 7th edition of the TNM classification to be published in January 2009 to include HER2 status. The 8th edition to be published in January 2016 will be much more challenging. Experts will have to choose between 'evolution' and 'revolution'. The 'evolution' approach will integrate, for example, the 70-gene signature (MammaPrint^{TM}) or the 21-gene signature (OncotypeDX^{TM}) if validated in ongoing large prospective trials. The 'revolution' approach will take in consideration new subtypes of breast cancer with validated survival data in order to move to a molecular classification. To be able to make these changes, clinicians, pathologists, molecular biologists and biostatisticians will have to work hard in very close collaboration in the coming years.

Conflict of interest statement

None declared.

References

1 Gospodarowicz M, Benedet L, Hutter RV, Fleming I, Henson DE, Sobin LH. History and international developments in cancer staging. *Cancer Prev Control* 1998, **2**, 262–268.
2 Sobin LH, Wittekind C. *TNM: classification of malignant tumours, 6th edn.* New York, Wiley-Liss, 2002.
3 Early Breast Cancer Trialists' Collaborative G. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet* 2005, **365**, 1687–1717.
4 Romond EH, Perez EA, Bryant J, *et al.* Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. [see comment]. *New Engl J Med* 2005, **353**(16), 1673–1684.

5 Smith I, Procter M, Gelber RD, *et al.* Two-year follow-up of trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer: a randomised controlled trial. *Lancet* 2007, **369**, 29–36.

6 Cady B. New era in breast cancer. Impact of screening on disease presentation. *Surg Oncol Clin N Am* 1997, **6**, 195–202.

7 Fisher B, Dignam J, Tan-Chiu E, *et al.* Prognosis and treatment of patients with breast tumors of one centimeter or less and negative axillary lymph nodes. *J Natl Cancer Inst* 2001, **93**, 112–120.

8 Lippman ME, Hayes DF. Adjuvant therapy for all patients with breast cancer? *J Natl Cancer Inst* 2001, **93**, 80–82.

9 Haybittle JL, Blamey RW, Elston CW, *et al.* A prognostic index in primary breast cancer. *Br J Cancer* 1982, **45**, 361–366.

10 Todd JH, Dowle C, Williams MR, *et al.* Confirmation of a prognostic index in primary breast cancer. *Br J Cancer* 1987, **56**, 489–492.

11 Sundquist M, Thorstenson S, Brudin L, Wingren S, Nordenskjold B. Incidence and prognosis in early onset breast cancer. *Breast* 2002, **11**, 30–35.

12 D'Eredita G, Giardina C, Martellotta M, Natale T, Ferrarese F. Prognostic factors in breast cancer: the predictive value of the Nottingham Prognostic Index in patients with a long-term follow-up that were treated in a single institution. *Eur J Cancer* 2001, **37**, 591–596.

13 Ravdin PM, Siminoff LA, Davis GJ, *et al.* Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. *J Clin Oncol* 2001, **19**, 980–991.

14 Olivotto IA, Bajdik CD, Ravdin PM, *et al.* Population-based validation of the prognostic model ADJUVANT! for early breast cancer. *J Clin Oncol* 2005, **23**, 2716–2725.

15 Lundin J, Lehtimaki T, Lundin M, *et al.* Generalisability of survival estimates for patients with breast cancer – a comparison across two population-based series. *Eur J Cancer* 2006, **42**, 3228–3235.

16 Singletary SE, Allred C, Ashley P, *et al.* Revision of the American Joint Committee on Cancer staging system for breast cancer. *J Clin Oncol* 2002, **20**, 3628–3636.

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

18 Isaacs C, Stearns V, Hayes DF. New prognostic factors for breast cancer recurrence. *Semin Oncol* 2001, **28**, 53–67.

19 Hayes DF, Trock BJ, Harris AL. Assessing the clinical impact of prognostic factors: when is 'statistically significant' clinically useful? *Breast Cancer Res Treat* 1998, **52**, 305–319.

20 Braun S, Vogl FD, Naume B, *et al.* A pooled analysis of bone marrow micrometastasis in breast cancer. *N Engl J Med* 2005, **353**, 793–802.

21 Janicke F, Prechtl A, Thomssen C, *et al.* Randomized adjuvant chemotherapy trial in high-risk, lymph node-negative breast cancer patients identified by urokinase-type plasminogen activator and plasminogen activator inhibitor type 1. *J Natl Cancer Inst* 2001, **93**, 913–920.

22 Harbeck N, Kates RE, Schmitt M, *et al.* Urokinase-type plasminogen activator and its inhibitor type 1 predict disease outcome and therapy response in primary breast cancer. *Clin Breast Cancer* 2004, **5**, 348–352.

23 van't Veer L, Dai H, van de Vijver M. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002, **415**, 530–535.

24 van de Vijver MJ, He YD, van't Veer L, *et al.* A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 2002, **347**, 1999–2009.

25 Buyse M, Loi S, van't Veer L, *et al.* Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *J Natl Cancer Inst* 2006, **98**, 1183–1192.

26 Wang Y, Klijn JG, Zhang Y, *et al.* Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. [see comment]. *Lancet* 2005, **365**, 671–679.

27 Foekens JA, Atkins D, Zhang Y, *et al.* Multicenter validation of a gene expression-based prognostic signature in lymph node-negative primary breast cancer. *J Clin Oncol* 2006, **24**, 1665–1671.

28 Paik S, Shak S, Tang G, *et al.* A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 2004, **351**, 2817–2826.

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

30 Goldhirsch A, Glick JH, Gelber RD, Coates AS, Thurlimann B, Senn HJ. Meeting highlights: international expert consensus on the primary therapy of early breast cancer 2005. *Ann Oncol* 2005, **16**, 1569–1583.

31 Piccart-Gebhart MJ, Procter M, Leyland-Jones B, *et al.* Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. [see comment]. *N Engl J Med* 2005, **353**, 1659–1672.

32 Joensuu H, Kellokumpu-Lehtinen PL, Bono P, *et al.* Adjuvant docetaxel or vinorelbine with or without trastuzumab for breast cancer. [see comment]. *N Engl J Med* 2006, **354**, 809–820.

33 Slamon DJ, Eiermann W, Robert N, *et al.* Phase III randomized trial comparing doxorubicin and cyclophosphamide followed by docetaxel (AC $\rightarrow$ T) with doxorubicin and cyclophosphamide followed by docetaxel and trastuzumab (AC $\rightrightarrows$ H) with docetaxel, carboplatin and trastuzumab (TCH) in HER2 positive early breast cancer patients: BCIRG 006 study. *Breast Cancer Res Treat* 2005, **94**(suppl 1), S5 (abstract 1).

34 Chang JC, Wooten EC, Tsimelzon A, *et al.* Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. [see comment]. *Lancet* 2003, **362**, 362–369.

35 Ayers M, Symmans WF, Stec J, *et al.* Gene expression profiles predict complete pathologic response to neoadjuvant paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide chemotherapy in breast cancer. *J Clin Oncol* 2004, **22**, 2284–2293.

36 Gianni L, Zambetti M, Clark K, *et al.* Gene expression profiles in paraffin-embedded core biopsy tissue predict response to chemotherapy in women with locally advanced breast cancer. *J Clin Oncol* 2005, **23**, 7265–7277.

37 Hess KR, Anderson K, Symmans WF, *et al.* Pharmacogenomic predictor of sensitivity to preoperative chemotherapy with paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide in breast cancer. *J Clin Oncol* 2006, **24**, 4236–4244.

38 Rouzier R, Perou CM, Symmans WF, *et al.* Breast cancer molecular subtypes respond differently to preoperative chemotherapy. *Clin Cancer Res* 2005, **11**, 5678–5685.

39 Colleoni M, Viale G, Zahrieh D, *et al.* Chemotherapy is more effective in patients with breast cancer not expressing steroid hormone receptors: a study of preoperative treatment. *Clinical Cancer Res* 2004, **10**, 6622–6628.

40 Gonzalez-Angulo AM, McGuire SE, Buchholz TA, *et al.* Factors predictive of distant metastases in patients with breast cancer

who have a pathologic complete response after neoadjuvant chemotherapy. *J Clin Oncol* 2005, **23**, 7098–7104.

41 Chung CH, Bernard PS, Perou CM. Molecular portraits and the family tree of cancer. *Nat Genet* 2002, **32**(suppl), 533–540.

42 Perou CM, Sorlie T, Eisen MB, *et al.* Molecular portraits of human breast tumours. *Nature* 2000, **406**, 747–752.

43 Sorlie T, Perou CM, Tibshirani R, *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA* 2001, **98**, 10869–10874.

44 Sorlie T, Tibshirani R, Parker J, *et al.* Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci USA* 2003, **100**, 8418–8423.

45 Sotiriou C, Neo SY, McShane LM, *et al.* Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *Proc Natl Acad Sci USA* 2003, **100**, 10393–10398.

46 Farmer P, Bonnefoi H, Becette V, *et al.* Identification of molecular apocrine breast tumours by microarray analysis. *Oncogene* 2005, **24**, 4660–4671.

47 Doane AS DM, Lal P, Donaton M, Zhang L, Hudis C, Gerald WL. An estrogen receptor-negative breast cancer subset characterised by a hormonally regulated transcriptional program and response to androgen. *Oncogene* 2006, **25**, 1–15.