ORIGINAL ARTICLE

Semi-quantitative scoring of potentially predictive markers for endocrine treatment of breast cancer: a comparison between whole sections and tissue microarrays

Katrine L Henriksen, Birgitte B Rasmussen, Anne E Lykkesfeldt, Susanne Møller, Bent Ejlertsen, Henning T Mouridsen

....

Aim: To assess whether immunohistochemically stained tissue microarrays (TMA) of 2 mm cores from paraffin embedded tumour tissue may replace whole sections in semi-quantitative evaluation of selected potential markers for endocrine treatment.
Methods: Whole sections and 2 mm cores on TMA were used for immunohistochemical staining of potential markers for endocrine treatment. The Allred scoring system was used for the markers with nuclear localisation: the oestrogen receptor, the progesterone receptor, p27 and the oestrogen receptor co-regulator amplified in breast cancer 1 (AIB1). The Allred scoring system was also used for the non-nuclear markers Bcl-2, pS2 and cyclooxygenase 2 (COX-2); the membrane receptors HER-2, insulin-like growth factor I receptor (IGF-IR) and epidermal growth factor receptor were quantified according to the guidelines for the Herceptest.
Results: The data and statistical analyses showed that the semi-quantitative evaluation of oestrogen receptor, progesterone receptor, AIB1, COX-2, HER-2 and IGF-IR on TMA blocks was comparable with analysis on whole sections.
Conclusions: This study shows that semi-quantitative scoring of 2 mm cores on TMA is feasible for several potential markers for endocrine therapy. Considering the small size of many breast tumours, the speed and cost-effectiveness of immunohistochemistry on TMA compared with whole sections, and the importance of the expression level of the proteins, semi-quantitative scoring on TMA has great potential in both retrospective and prospective studies aiming at improving the prediction of response to endocrine treatment.

See end of article for authors' affiliations
....

Correspondence to:
Dr B B Rasmussen, Department of Pathology, Nordsjaellands Hospital, Hilleroed, Helsevej 2, DK-3400 Hilleroed, Denmark; bbr@noh.regionh.dk

B reast cancer is the most common cancer among women in the western world, and much effort is put into understanding the basis of tumour development and progression, as well as into identifying markers to predict prognosis and response to treatment. Immunohistochemistry (IHC) is one of the techniques traditionally used to assess such markers. Since the availability of material is often limited, especially when it comes to small tumours and previously used material, the development of the tissue microarray (TMA) technique has helped overcome these problems,[1] and this method is now widely used to evaluate tissue specimens at DNA and protein level.[2–4]

TMAs are constructed by taking cylindrical core biopsies from "donor" paraffin blocks into a new "recipient" paraffin block. The use of cores as small as 0.6 mm has been confirmed to be adequate for analysing breast cancer specimens by IHC as positive or negative for the expression of the hormone receptors oestrogen receptor and progesterone receptor and the tyrosine kinase receptor HER-2.[5 6] A small number of studies have also confirmed the correlation between semi-quantitative analysis of oestrogen receptor, progesterone receptor and p27 expression in TMAs and analysis of whole sections.[7 8]

Response to endocrine treatment requires oestrogen-responsive tumour growth mediated via activation of the oestrogen receptor. However, whereas the presence of oestrogen receptor or progesterone receptor is a prerequisite for response, only a fraction of patients with tumours positive for oestrogen receptor or progesterone receptor benefit from endocrine treatment. Therefore, better predictive markers are urgently required. The semi-quantitative level of protein expression may be important in the search for new predictive markers for response to endocrine treatment of breast cancer, since it has

been shown that higher IHC scores for oestrogen receptor, progesterone receptor and pS2 are associated with better response to treatment with tamoxifen.[9] Thus, it is very important when using the TMA method to discover new predictive markers to know whether TMAs can be used for semi-quantitative analysis of the potential markers. A number of different proteins have been proposed as new predictive markers that could identify those patients with breast cancer who are positive for oestrogen who will benefit from different endocrine treatments, such as antioestrogen or aromatase inhibitors. These proteins include the cell cycle regulating protein p27,[10 11] the oestrogen regulated protein pS2,[12] the anti-apoptotic protein Bcl-2[13] and the oestrogen receptor coregulator protein amplified in breast cancer 1 (AIB1). AIB1 has been found to be associated with worse disease-free survival in patients receiving treatment with the antioestrogen tamoxifen.[14] Another potential predictive factor for endocrine treatment is cyclooxygenase 2 (COX-2). COX-2 is an enzyme that indirectly regulates the synthesis of the enzyme aromatase, which converts androgens to oestrogens.[15] It has been proposed that COX-2 could serve as a surrogate marker for the aromatase enzyme.[16] HER-2, the epidermal growth factor receptor (EGFR) and the insulin-like growth factor I receptor (IGF-IR), are other examples of proteins that have been described to crosstalk with the oestrogen receptor and be of importance for response to endocrine treatment.[17 18] In a neoadjuvant phase III study, patients with oestrogen receptor who also expressed EGFR or

....

Abbreviations: AIB1, oestrogen receptor co-regulator amplified in breast cancer 1; COX-2, cyclooxygenase 2; EGFR, epidermal growth factor receptor; IHC, immunohistochemistry; IGF-IR, insulin-like growth factor I receptor; TMA, tissue microarray

HER-2 had a significantly better response to treatment with the aromatase inhibitor letrozole than to antiestrogen treatment with tamoxifen.[19] IGF-IR has been shown to be an oestrogen-regulated protein,[20] associated with response to antioestrogen treatment in cell culture models.[21]

Different scoring systems exist for semi-quantitative evaluation of immunohistochemical staining of breast cancer tissue. Some of the most widely used are the Quick score, the H score and the Allred score, all described for evaluation of nuclear staining.[7 8 22 23] Some authors have also used the Allred score or an H score to evaluate pS2, which is located in the cytoplasm,[9] and for cytoplasmic staining of aromatase.[24] We have used the Allred score to determine the nuclear proteins oestrogen receptor, progesterone receptor , p27 and AIB1, and the cytoplasmic proteins Bcl-2, COX-2 and pS2. The guidelines for the Herceptest, developed for HER-2 evaluation, have been widely used for HER-2 determination[25 26] and also used to assess the membrane receptor EGFR.[19] In this study EGFR, HER-2 and IGF-IR levels were determined according to the guidelines for the Herceptest.

The purpose of this study was to examine whether TMAs of 2 mm cores from paraffin-embedded tumour tissue may replace whole sections for semi-quantitative detection of oestrogen receptor, progesterone receptor , p27, AIB1, Bcl-2, COX-2, pS2, HER-2, EGFR and IGF-IR in breast cancer samples analysed by IHC. The core diameter of 2 mm was selected to ensure sufficient tumour cells for the quantitative evaluation and still obtain a significant reduction in the number of blocks to be analysed and in the costs for reagents.

MATERIALS AND METHODS
Tumour material
Archive formalin-fixed and paraffin-embedded primary tumour tissue was obtained from Danish patients with breast cancer who had participated in an international randomised clinical phase III trial comparing letrozole with tamoxifen as first-line therapy for postmenopausal women with metastatic breast cancer.[27] Altogether 89 Danish patients participated in the trial, and paraffin blocks containing primary tumour material were identified according to the patients' pathology reports. It was possible to collect primary tumour material from 69 patients. One patient had primary bilateral breast cancer, resulting in a total number of 70 tumours in this study. A total of 54 tumours were classified as invasive ductal and nine as invasive lobular carcinoma. Two tumour samples were classified as neither lobular nor ductal; four samples were not suitable for classification. One sample contained only in situ material and was not used.

Immunohistochemical staining
Sections, 3 μm, were dewaxed in coconut oil and rehydrated in a graded series of ethanol. Slides were preheated for 10 min and boiled for 15 min in a microwave oven at 600 W in TEG buffer (pH = 9, Bie & Berntsen, Denmark) for antigen retrieval and rinsed in tap water, except for slides stained for EGFR, in which a 5 min protease digestion was used as antigen retrieval. All immunostainings were performed at room temperature using the automated immunostainer Tech-mate 500 (DakoCytomation), according to the following protocol: slides were washed in TBS +0.1% BRIJ-35 detergent (AX-LAB, Denmark) and incubated with primary antibody diluted in TBS +0.1% BRIJ-35 +1% BSA +15 nM sodiumazide for 60 min. After washing, endogenous peroxidase activity was blocked using 3% H_2O_2 in TBS +0.1% BRIJ-35. The ChemMate EnVision[+] Detection Kit (Peroxidase/Dab, Rabbit/Mouse, K5007, DakoCytomation, Glostrup, Denmark) was used as detection system for the primary antibodies. After washing, slides

were counterstained with haematoxylin and dehydrated in graded series of ethanol, and finally mounted with Pertex (Histolab, Denmark). The following primary antibodies, all monoclonal mouse subtype IgG_1, were used: oestrogen receptor, clone ER1D5, 1:200 (Immunotech, Trichem Aps, Frederikssund, Denmark); pregesterone receptor, clone 16, 1:200 (Novocastra, Trichem Aps, Frederikssurd, Denmark); IGF-IR, clone 24–31, 1:200 (NeoMarkers, AH-diagnostics, Aarhus, Denmark); p27, clone F-8, 1:100 (Santa Cruz); AIB1, clone 34, 1:60 (Transduction Laboratories, BD-Bioscience, Brøndby, Denmark), COX-2, no: 611 104, 1:150 (Cayman Chemicals, AH-diagnostics, Aarhus, Denmark), pS2, clone BC04, 1:25 (DakoCytomation); Bcl-2, clone 124, 1:300 (DakoCytomation) and EGFR, clone E30, 1:50 (DakoCytomation). The specificity of the immunoreactions was verified by substituting the primary antibody with the corresponding concentration of mouse IgG_1, X 0931 (DakoCytomation). In addition, positive control slides with breast tumour tissue or other tissue known to stain positive were included in every run. HER-2 staining was performed with the ''Herceptest for the TechMate Instrument'' K 5206 (DakoCytomation) according to the manufacturer's instruction. All stainings of the individual antigens were performed in a single run to minimise inter-serial staining variation.

Preparation of tissue microarrays
TMA blocks were constructed using the TMA-builder from Histopathology Ltd (AH-diagnostics, Denmark). Targets for arraying (areas with representative invasive tumour) were identified by marking the corresponding areas on haematoxylin–eosin stained sections from each paraffin block. Two tissue cores with a diameter of 2 mm were transferred from each donor block to the recipient TMA block. Kidney tissue was placed in the first core of the upper left and right corner of the TMA block to ensure correct orientation when examining the slides.

Evaluation of immunohistochemical staining
KLH has scored all the samples and BBR has been consulted in cases of doubt. Only invasive tumour components were considered when judging the staining. One or two cores were scored. In the cases with two score values (approximately 80% of the tumours), the vast majority of the cases had identical score values or differed by one value. Fewer than 10% differed by two or more score values, and statistical evaluation of the association between the values in cases with two TMA scores showed a significant association between the scores. A preliminary investigation had shown that in the rare cases where score values differed by more than 2, the lowest score value was generally markedly below the score of the whole sections, whereas the high score value was equal to or close to the whole sections score. Therefore, the highest score values were used in the cases with two eligible scores. Semi-quantitative determination of oestrogen receptor, progesterone receptor, p27 and AIB1 was performed according to the method described by Allred et al.[22] This method was also used for semi-quantitative determination of the non-nuclear intracytoplasmic markers pS2, COX-2 and Bcl-2. In brief: the proportion of positive stained cells was rated as 0 = no cells stained positive, 1 = between 0% and 1% positive, 2 = between 1% and 10%, 3 = between 10% and 33%, 4 = between 33% and 66%, and 5 = between 66% and 100%. In addition to the proportion score, an intensity score was made on the basis of the average intensity of staining: 0 = negative, 1 = weak, 2 = intermediate and 3 = strong. The intensity score and the proportion score were added to obtain the total score; this is referred to as the Allred score, and is either 0 or between 2 and 8. Scores of 0 and 2 were interpreted as negative. Only nuclear staining was

judged when scoring oestrogen receptor, progesterone receptor, p27 and AIB1, whereas cytoplasmic staining was scored for COX-2, pS2 and Bcl-2. HER-2, EGFR and IGF-IR were scored according to the guidelines for HER-2 staining as 0, 1+, 2+ or 3+.[28] In brief: 0 = no staining or membranous staining in <10% of invasive tumour cells; 1+ = faint or barely perceptible membranous staining in >10%; 2+ = weak or complete membranous staining in >10%; 3+ = strong complete membranous staining in >10%. Scores 0 and 1+ were interpreted as negative in this study.

Statistical analysis

Agreement between the two ordinal variables: scores for whole sections and TMA were determined by calculating the kappa score, and the 95% confidence intervals (CIs) are given. Kappa statistics were also used to test agreement between scores on two TMA cores from the same tumour. The hypothesis of no systematic difference between the scores was analysed with Wilcoxon's Signed Rank Sum test on the difference (TMA–whole sections). The concordance in determination of a positive or negative response is given on the basis of the score of whole sections versus TMA. In all tests, a p value of <0.05 was used as the level of significance.

RESULTS

Paraffin blocks were available from 70 tumours, and 66 blocks had invasive tumour components adequate for TMA blocks. On microscopic examination of the TMA slides some cores had floated off, reducing the total number of TMA to be compared with whole sections to between 44 and 64.

Immunohistochemical staining and scoring of results of marker proteins

Figure 1 shows representative examples of immuno-histochemical staining of oestrogen receptor, progesterone receptor, p27, AIB1, COX-2, Bcl-2 and pS2, all scored by the Allred scoring system. An example of a low or intermediate score as well as a high score is shown for all seven markers. Staining for oestrogen receptor (fig 1A, 1B) and progesterone receptor (fig 1C, 1D) was observed as a distinct nuclear staining of varying intensity, as described previously in other studies.[23] [29] The patients enrolled in this study had all been previously classified as receptor positive (oestrogen receptor and/or progesterone receptor positive) or unknown after surgery for primary breast cancer,[27] and the results from this analysis confirm these previous tests, since the only tumour that was oestrogen receptor negative was positive for progesterone receptor. From table 1, it can be seen that the majority of the tumours had high Allred scores for oestrogen receptor and progesterone receptor. Staining of p27 (fig 1E, 1F) was observed in the nucleus, with weak or intermediate staining intensity in most cases. Nuclear staining was also seen in benign/normal epithelia, and in lymphocytes. Both low, intermediate and high p27 score values were observed (table 1). Examples of AIB1 staining are seen in fig 1G and 1H. AIB1 immunoreactivity was observed in the nucleus, in most cases with an intermediate or high Allred score value (table 1). Figure 1I and 1J show staining of COX-2, mostly seen as a granular staining in the cytoplasm, as also described by others.[30] [31] Surprisingly, some samples also showed staining in the nucleus and plasma membrane. As seen in table 1, low, intermediate and high COX-2 Allred score values were observed. Figure 1K and 1l illustrate staining of the anti-apoptotic protein Bcl-2, which appears as cytoplasmic staining. In most cases, a high Bcl-2 score value was observed (table 1). Figure 1M and 1N shows staining of pS2. Generally, the pS2 staining was granular and heterogenic with a tendency

to perinuclear localisation as seen in figure 1N. In all, 18 tumours (28%) were classified as pS2 negative (table 1).

Figure 2 shows staining of the membrane proteins HER-2, EGFR and IGF-IR, all scored by the guidelines for the Herceptest. An example of a 1+ and a 3+ score is shown for all three markers. Staining of HER-2 (fig 2A, 2B) was observed in the plasma membrane. Herceptest scoring values of 1+, 2+ and 3+ were observed for HER-2, but most of the samples scored 0 (table 1). Examples of membranous EGFR staining are seen in figure 2C and 2D. Since all samples in this study were scored as 0 (table 1), figure 2C and 2D shows examples of EGFR-positive breast cancer control samples obtained from breast tumours negative for oestrogen receptor. Staining of IGF-IR (fig 2E, 2F) had the same appearance as the HER-2 staining, except that IGF-IR staining was also seen often in the cytoplasm and in normal epithelia as well as in benign lesions. Only membrane staining was considered for the scoring. IGF-IR scoring values of 2+ and 3+ were observed for most of the samples (table 1).

Comparison of marker scores from whole sections and TMA

For each marker, the scores from TMA analysis and whole sections were compared, and the result from the oestrogen receptor scoring is shown as an example in table 2.

A total of 55 tumours were scored in both whole sections and TMA, and identical scores were observed in 35 samples; 14 samples varied by no more than one score. Four samples varied by two scores, whereas only two tumours varied by three scores or more.

Table 3 shows the distribution of the difference between TMA and whole sections scored for all markers. The best correspondence between scores was observed for the membrane markers HER-2, EGFR and IGF-IR, in which an equal score was found in 93%, 100% and 69%, respectively. For EGFR, all samples were scored as 0 in both whole sections and TMA. The largest variation in scores was observed for p27, but pS2, Bcl-2 and progesterone receptor also had examples of samples with a difference of up to 7 or 8 in Allred score value. Nevertheless, in the total analysis only 12.2% of the scores have a difference of more than ±1 between TMA and whole sections.

From table 3, it appears that the scores from whole sections were more often higher than lower than the score from TMA. Therefore a statistical analysis of the direction of the difference between whole sections and TMA was performed by Wilcoxon's signed rank sum test; the p values are given in table 3. As can be seen, no systematic discrepancy was present between scores obtained from whole sections and TMA, except for p27, Bcl-2 and pS2, where there was a significant systematic discrepancy towards a higher score in whole sections than in TMA. Table 3 also shows the kappa scores with 95% CIs and the p values for test of no association between the score values for TMA and whole section. EGFR was excluded from this analysis because all samples scored 0. Only 17 of the 44 samples scored for HER-2 had score values above 0, and thus the effective sample size for testing the agreement for this marker is smaller for the others. The p values for all markers except HER2 and EGFR was <0.01%, indicating a statistically significant agreement between the scoring values obtained by whole sections and TMA.

Comparison of positive and negative results on whole sections and TMA

From a clinical point of view the most important information may be whether positive and negative results from TMA blocks are comparable with the results from whole sections. According to the cut-off levels used in this study, the number and
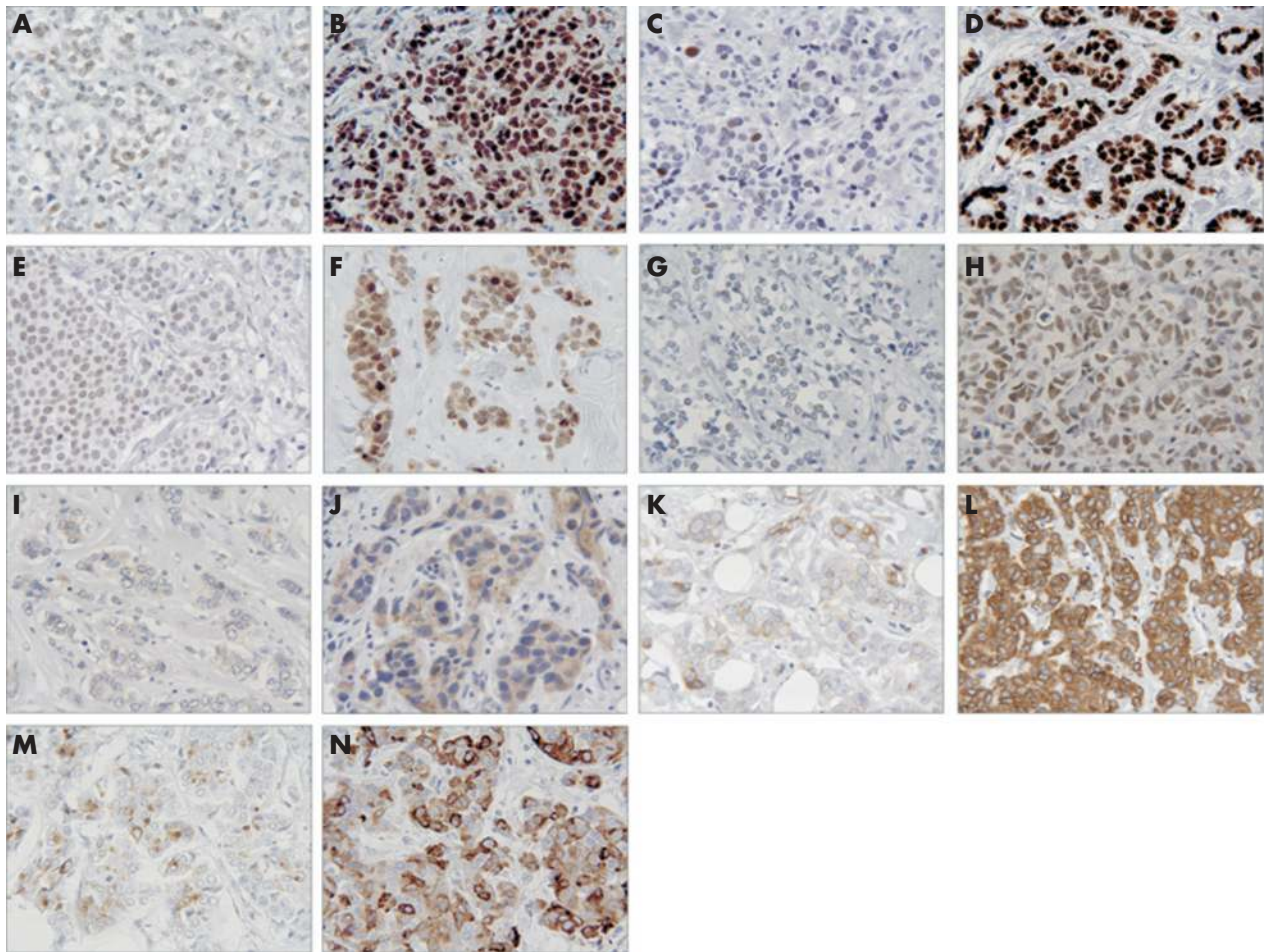
Figure 1 Examples of immunohistochemical staining of markers scored by the Allred scoring system: nuclear markers are shown in A–H and cytoplasmic markers in I–N. (A) Weak oestrogen receptor staining in most cells, Allred score 6; (B) strong oestrogen receptor staining, Allred score 8; (C) strong progesterone receptor staining in few cells, Allred score 4; (D) strong progesterone receptor staining in most cells, Allred score 8; (E) weak p27 staining in many cells, Allred score 5; (F) intermediate p27 staining in most cells, Allred score 7; (G) weak AIB1 staining in few cells, Allred score 3; (H) intermediate AIB1 staining in most cells, Allred score 7; (I) weak COX-2 staining in few cells, Allred score 3; (J) intermediate COX-2 staining in most cells, Allred score 7; (K) intermediate Bcl-2 staining in few cells, Allred score 5; (L) strong Bcl-2 staining in most cells, Allred score 8; (M) intermediate pS2 staining in few cells, Allred score 5; (N) strong pS2 staining in most cells, Allred score 8.

percentage of samples with concordance and discordance between positive and negative results were calculated (table 4). Apart from p27 and pS2, for which the discordance between positive and negative result was 13% and 14%, respectively, the concordance for the remaining markers exceeded 90%.

Table 1 Number of scores on whole sections

Allred score	Nuclear localisation				Cytoplasmic localisation			Herceptest score	Membrane localisation		
	Oestrogen receptor	Progesterone receptor	p27	AIB1	COX2	Bcl2	pS2		HER2	EGFR	IGF-IR
0	1	8	8	4	7	5	16	0	27	58	2
2	0	0	0	2	1	3	2				
3	1	4	8	1	4	3	6	1	7	0	6
4	1	7	9	4	3	2	13				
5	4	3	10	12	12	1	6	2	7	0	19
6	8	6	4	11	9	5	9				
7	22	6	10	11	7	24	6	3	3	0	27
8	18	20	6	7	3	15	6				
Total	55	54	55	52	46	58	64	Total	44	58	54

AIB1, oestrogen receptor co-regulator amplified in breast cancer 1; COX2 cyclooxygenase 2; EGFR epidermal growth factor receptor; IGF-IR, insulin-like growth factor I receptor.
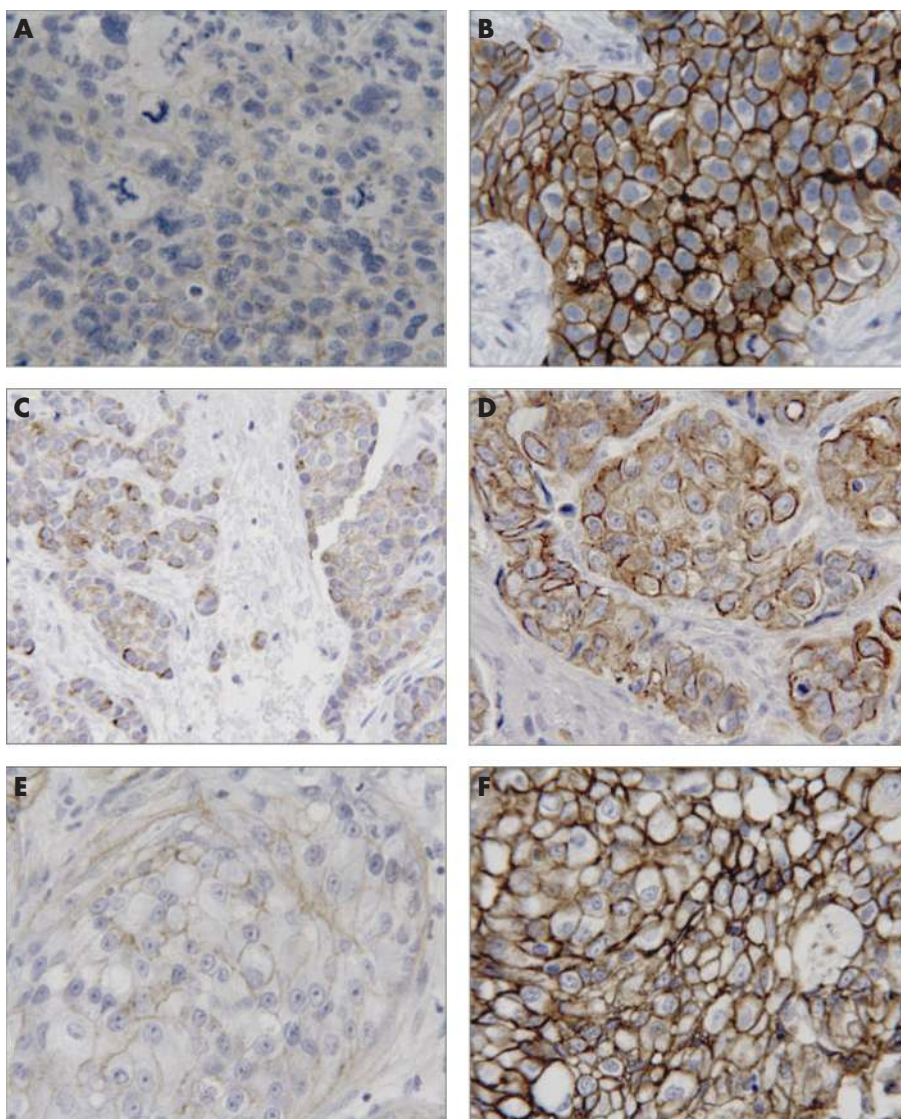
Figure 2 Examples of immunohistochemical staining of membrane markers scored by the Herceptest scoring system. (A) Faint HER-2 staining, Herceptest score 1+; (B) strong HER-2 staining, Herceptest score 3+; (C) faint epidermal growth factor receptor (EGFR) staining, Herceptest score 1+; (D) strong EGFR staining, Herceptest score 3+. (E) faint insulin-like growth factor I receptor (IGF-IR) staining, Herceptest score 1+; (F) strong IGF-IR staining, Herceptest score 3+. The samples for EGFR staining do not belong to this series of samples, but were selected from EGFR-positive control breast cancer samples.

DISCUSSION

Many resources and much effort are put into understanding the basis of breast cancer development and progression, as well as into identifying markers for prognosis and prediction of response to treatment. In the search for new markers for response to endocrine treatment the semi-quantitative level of specific markers may be important. For oestrogen receptor, progesterone receptor and pS2, high IHC score levels were associated with better response to tamoxifen treatment.[9] Furthermore, the presence of as few as 1–10% of cells weakly

Table 2 Oestrogen receptor Allred score from whole sections and tissue microarrays

	TMA 0	TMA 2	TMA 3	TMA 4	TMA 5	TMA 6	TMA 7	TMA 8	Total
Whole section 0	1	0	0	0	0	0	0	0	1
Whole section 2	0	0	0	0	0	0	0	0	0
Whole section 3	1	0	0	0	0	0	0	0	1
Whole section 4	0	0	0	0	0	1	0	0	1
Whole section 5	0	0	1	1	1	1	0	0	4
Whole section 6	1	0	0	0	1	3	1	2	8
Whole section 7	0	0	0	0	0	2	13	7	22
Whole section 8	0	0	0	0	0	0	1	17	18
Total	3	0	1	1	2	7	15	26	55

TMA, tissue microarray.

Table 3 Distribution of the difference between scores (tissue microarray score–whole section score); test of the hypothesis of no systematic difference. The estimated kappa score quantifies the agreement

Difference (TMA-whole section)	Nuclear localisation				Cytoplasmic localisation			Membrane localisation		
	Oestrogen receptor	Progesterone receptor	p27	AIB1	COX2	Bcl2	pS2	HER-2*	EGFR†	IGF-IR
−8	—	1	1	—	—	—	—	—	—	—
−7	—	—	1	—	—	—	1	—	—	—
−6	1	—	—	—	—	1	1	—	—	—
−5	—	—	1	—	—	1	1	—	—	—
−4	—	1	1	—	—	1	3	—	—	—
−3	1	3	4	—	4	1	1	—	—	—
−2	1	2	5	5	2	4	5	—	—	—
−1	5	8	11	10	8	7	11	2	—	10
0	**35**	**31**	**24**	**28**	**24**	**39**	**34**	**41**	**58**	**37**
1	9	3	5	9	7	3	7	1	—	7
2	3	4	1	—	1	1	—	—	—	—
3	—	1	—	—	—	—	—	—	—	—
>4	—	—	1	—	—	—	—	—	—	—
Test of agreement										
p Value‡	0.67	0.20	0.001	0.07	0.07	0.007	0.001	NA	NA	0.63
Kappa score	0.49	0.48	0.35	0.46	0.43	0.58	0.43	0.88	NA	0.49
95% CI	0.33% to 0.65%	0.34% to 0.62%	0.20% to 0.50%	0.29% to 0.62%	0.26% to 0.60%	0.42% to 0.73%	0.3% to 0.57%	0.74% to 1.00%		0.29 to 0.70
p Value§	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01		<0.01

AIB1, oestrogen receptor co-regulator amplified in breast cancer 1; COX2, cyclooxygenase 2; EGFR, epidermal growth factor receptor; IGF-IR, insulin-like growth factor I receptor; NA, not applicable; TMA, tissue microarrays.
*For the marker HER-2, only 17 of 44 samples have values above zero. Thus the effective sample size for testing the agreement is smaller than for the other markers.
†It is not possible to test the agreement for the marker EGFR, because all samples have the value 0.
‡Wilcoxon's rank sum test of no systematic difference.
§Test of no association.

positive for oestrogen receptor (corresponding to an Allred score greater than 2) predicts a higher probability of disease-free survival for patients receiving endocrine treatment than with tumours negative for oestrogen receptor. [23] TMA is a fast and cost-effective technique for IHC analysis that requires only a small amount of paraffin-embedded tumour tissue. We conducted this study to analyse whether the TMA technique can be used for semi-quantitative IHC analyses of a panel of well-established as well as new potentially predictive markers for endocrine treatment.

Generally, the nuclear markers, for which the Allred scoring system was originally introduced,[22] were easier to evaluate by this scoring system than the cytoplasmic markers. Both oestrogen receptor and progesterone receptor have for a long time been known to predict response to endocrine treatment,[32] and staining methods for both markers are well established. Furthermore, semi-quantitative evaluation of oestrogen receptor and progesterone receptor in TMA samples has been examined previously, and although the scores were not exactly the same in whole sections and TMA, the association between oestrogen receptor and progesterone receptor expression in the whole sections and TMA was statistically significant.[7] In our study, we also observed a statistically significant association between oestrogen receptor and progesterone receptor values scored in whole sections and TMA, p<0.01. Exactly the same score was found in 63.6% and 57.4% of the oestrogen receptor and progesterone receptor samples, respectively. Both sampling

and observer variability influence the score, and it is noteworthy that with the complex Allred score, as many as 89.1% and 77.8% of the oestrogen receptor and progesterone receptor scores, respectively, are within only 1 score difference. In only 2% of the tumour samples did a score difference of 1 in oestrogen receptor or progesterone receptor value result in altering the categorisation of the tumour to negative or positive. Thus, this analysis clearly supports using semi-quantitative IHC analysis of oestrogen receptor and progesterone receptor in TMA in future studies as a convenient substitute for semi-quantitation in whole sections.

All antibodies were recommended by the manufacturers for IHC light microscopy analysis of formalin-fixed and paraffin-embedded samples, except for the AIB1-antibody, which was recommended for immunofluorescence and western analysis. In spite of this, the AIB1 antibody has been proven to work well in both immunocytochemistry and IHC with MCF-7 breast cancer cells and breast tumour samples.[33] In this study, only 5 out of 46 samples analysed for AIB1 deviated more than 1 score in whole sections and TMA (table 3), and a significant association, with no systematic discrepancy, was found between whole sections and TMA scores.

In a few samples stained for p27, both cytoplasmic as well as nuclear staining were observed. This has been described as a sign of bad prognosis,[34] but as it was observed in no more than a few cases, only nuclear p27 staining was considered for the Allred score. Both p27 and pS2 staining displayed a very

Table 4 Comparison of positive and negative results on whole sections and tissue microarrays

	Nuclear localisation				Cytoplasmic localisation			Membrane localisation		
	Oestrogen receptor	Progesterone receptor	p27	AIB1	COX2	Bcl2	pS2	HER-2	EGFR	IGF-IR
Concordance (%)	96	93	87	96	91	91	86	100	100	91
Discordance (%)	4	7	13	4	9	9	14	0	0	9

AIB1, oestrogen receptor co-regulator amplified in breast cancer 1; COX2, cyclooxygenase 2; EGFR, epidermal growth factor receptor; IGF-IR, insulin-like growth factor I receptor.

heterogeneous tissue distribution in the same whole sections sample. This may explain why p27 and pS2 had a pronounced systematic discrepancy towards a higher score in whole sections compared with TMA (table 3) and why both of these markers had a concordance of positive and negative results of <90%. COX-2 staining, on the other hand, appeared relatively heterogeneous between samples. Some samples showed not only perinuclear and cytoplasmic staining, but also staining of the nucleus and plasma membrane, which made COX-2 difficult and time consuming to analyse. In spite of these difficulties in the evaluation of COX-2, a significant association was observed when comparing whole sections and TMA (table 3). For Bcl-2, no difficulties in the scoring procedure could explain the observed statistical discrepancy between higher scores in whole sections than in TMA. Even though a higher score value for Bcl-2 was found in whole sections, a significant association was found between whole sections and TMA analysis. Because of the systematic tendency towards higher Bcl-2 scores in whole sections, further examination of Bcl-2 with the TMA method would be required to justify semi-quantitative analysis of Bcl-2 with TMA.

Most samples scored 0 for the membrane markers HER-2 and EGFR (table 1). The observation that only seven samples scored 2+ and three samples 3+ for HER-2 was not surprising, as high HER-2 levels are correlated with oestrogen receptor negativity.[35] Only 3 out of 44 samples did not receive identical HER-2 scoring results in whole sections and TMA (table 3). This number was not adequate to examine whether the scoring results obtained from whole sections were higher than scores from TMA, but the significant association between whole sections and TMA scores confirms other studies in which analysis of HER-2 by the Herceptest was found suitable for TMA.[5][36] Several papers have described absent or low EGFR levels in oestrogen receptor positive tumours,[37–40] and a review paper reports that oestrogen receptor positive tumours were EGFR positive in 4–51% of the samples, with a mean of 29%.[41] None of the samples analysed in this study expressed EGFR, neither on whole sections nor on TMA. Therefore, it is obvious that the proportion of equal scores is 100%. However, no adequate conclusions can be drawn on the basis of this result. In contrast to HER-2 and EGFR, many of the samples were scored as 2+ or 3+ for IGF-IR. Other studies have found comparable results; for example, 87.5% of ductal and lobular breast cancer types were found to score 6–8 when evaluated by the Allred scoring system.[42] Compared with HER-2, more cytoplasmatic IGF-IR immunoreactivity was observed, as also described by others.[43][44] In 31.5% of samples, IGF-IR scoring did not display identical results in whole sections and TMA. But no systematic discrepancy was shown between whole sections and TMA, and a significant association was found between whole sections and TMA. Of the analysed membrane markers, we find TMA as good as whole sections for semi-quantitative IHC analysis of both HER-2 and IGF-IR.

A previous finding[8] has disclosed a good agreement between p27 analysis from whole sections and TMA with both the Quick score and the H score. The comparison was performed with only 20 samples, in contrast to this study, with 55 whole sections and TMA samples. We also found a statistically significant association between the semi-quantitative p27 scores in whole sections and TMA. However, we also observed a significant systematic tendency towards a higher score in whole sections than in TMA for p27; and the discordance between positive and negative results exceeds 10% (table 4). Thus, we conclude that this optimised p27 IHC is not suited for scoring with the TMA technique.

In summary, our results showed a statistically significant association between the scoring values for oestrogen receptor, progesterone receptor, p27, AIB1, COX-2, Bcl-2, pS2, HER-2 and IGF-IR obtained in TMA with 2 mm cores and whole sections, but with a systematic tendency towards a higher score in whole sections than in TMA for p27, Bcl-2 and pS2. Furthermore, the concordance between positive and negative scoring result exceeded 90% for oestrogen receptor, progesterone receptor, AIB1, COX-2, Bcl-2, HER-2 and IGF-IR. Considering that the goal of this study is to include several markers in the selection of treatment for the patient, a discordance of <10% for the individual markers may be considered acceptable. Thus, we conclude that semi-quantitative analysis of breast cancer TMAs for oestrogen receptor, progesterone receptor, AIB1, COX-2, HER-2 and IGF-IR can replace analysis on whole sections. When only a positive or negative scoring result is required, Bcl-2 can be added to this list of markers.

. . . .

Authors' affiliations

Katrine L Henriksen, Anne E Lykkesfeldt, Department of Tumor Endocrinology, Institute of Cancer Biology, Danish Cancer Society, Copenhagen Ø, Denmark
Birgitte B Rasmussen, Department of Pathology, Roskilde County Hospital, Roskilde, Denmark
Susann Møller, Henning T Mouridsen, Danish Breast Cancer Cooperative Group, Rigshospitalet, Copenhagen Ø, Denmark
Bent Ejlertsen, Henning T Mouridsen, Department of Oncology, Rigshospitalet, Copenhagen Ø, Denmark

Competing interests: None.

Ethics approval: The local science ethics committees for Copenhagen and Frederiksberg counties and the Danish Data Protection Agency have approved the research protocol.

REFERENCES

1 **Kononen J**, Bubendorf L, Kallioniemi A, *et al*. Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat Med* 1998;**4**:844–7.
2 **Kallioniemi OP**, Wagner U, Kononen J, *et al*. Tissue microarray technology for high-throughput molecular profiling of cancer. *Hum Mol Genet* 2001;**10**:657–62.
3 **Hao X**, Sun B, Hu L, *et al*. Differential gene and protein expression in primary breast malignancies and their lymph node metastases as revealed by combined cDNA microarray and tissue microarray analysis. *Cancer* 2004;**100**:1110–22.

Take-home messages

- Semi-quantitative evaluation of the oestrogen receptor, the progesterone receptor, the oestrogen receptor co-regulator amplified in breast cancer 1 (AIB1), cyclooxygenase-2 (COX-2), HER-2 and insulin-like growth factor I receptor (IGF-IR) on 2 mm cores in tissue microarray (TMA) blocks can replace immunohistochemical analysis of whole sections.
- Semi-quantitative scoring of these proteins on TMA has great potential in both retrospective and prospective studies that aim to improve the prediction of response to endocrine treatment.

4 **Kay E**, O'Grady A, Morgan JM, *et al*. Use of tissue microarray for interlaboratory validation of HER2 immunocytochemical and FISH testing. *J Clin Pathol* 2004;**57**:1140–4.

5 **Van den Eynden GG**, Van dA, I, Van Laere S, et al. Validation of a tissue microarray to study differential protein expression in inflammatory and non-inflammatory breast cancer. *Breast Cancer Res Treat* 2004;**85**:13–22.

6 **Zhang D**, Salto-Tellez M, Putti TC, *et al*. Reliability of tissue microarrays in detecting protein expression and gene amplification in breast cancer. *Mod Pathol* 2003;**16**:79–84.

7 **Gillett CE**, Springall RJ, Barnes DM, *et al*. Multiple tissue core arrays in histopathology research: a validation study. *J Pathol* 2000;**192**:549–53.

8 **Barnes A**, Pinder SE, Bell JA, *et al*. Expression of p27kip1 in breast cancer and its prognostic significance. *J Pathol* 2003;**201**:451–9.

9 **Elledge RM**, Green S, Pugh R, *et al*. Estrogen receptor (ER) and progesterone receptor (PgR), by ligand-binding assay compared with ER, PgR and pS2, by immuno-histochemistry in predicting response to tamoxifen in metastatic breast cancer: a Southwest Oncology Group Study. *Int J Cancer* 2000;**89**:111–17.

10 **Alkarain A**, Slingerland J. Deregulation of p27 by oncogenic signaling and its prognostic significance in breast cancer. *Breast Cancer Res* 2004;**6**:13–21.

11 **Arteaga CL**. Cdk inhibitor p27Kip1 and hormone dependence in breast cancer. *Clin Cancer Res* 2004;**10**:368S–71S.

12 **Fitzgibbons PL**, Page DL, Weaver D, *et al*. Prognostic factors in breast cancer. College of American Pathologists Consensus Statement 1999. *Arch Pathol Lab Med* 2000;**124**:966–78.

13 **Daidone MG**, Luisi A, Martelli G, *et al*. Biomarkers and outcome after tamoxifen treatment in node-positive breast cancers from elderly women. *Br J Cancer* 2000;**82**:270–7.

14 **Osborne CK**, Bardou V, Hopp TA, *et al*. Role of the estrogen receptor coactivator AIB1 (SRC-3) and HER-2/neu in tamoxifen resistance in breast cancer. *J Natl Cancer Inst* 2003;**95**:353–61.

15 **Richards JA**, Petrel TA, Brueggemeier RW. Signaling pathways regulating aromatase and cyclooxygenases in normal and malignant breast cells. *J Steroid Biochem Mol Biol* 2002;**80**:203–12.

16 **Brodie AM**, Lu Q, Long BJ, *et al*. Aromatase and COX-2 expression in human breast cancers. *J Steroid Biochem Mol Biol* 2001;**79**:41–7.

17 **Schiff R**, Massarweh SA, Shou J, *et al*. Cross-talk between estrogen receptor and growth factor pathways as a molecular target for overcoming endocrine resistance. *Clin Cancer Res* 2004;**10**:331S–6S.

18 **Yee D**, Lee AV. Crosstalk between the insulin-like growth factors and estrogens in breast cancer. *J Mammary Gland Biol Neoplasia* 2000;**5**:107–15.

19 **Ellis MJ**, Coop A, Singh B, *et al*. Letrozole is more effective neoadjuvant endocrine therapy than tamoxifen for ErbB-1- and/or ErbB-2-positive, estrogen receptor-positive primary breast cancer: evidence from a phase III randomized trial. *J Clin Oncol* 2001;**19**:3808–16.

20 **Stewart AJ**, Johnson MD, May FE, *et al*. Role of insulin-like growth factors and the type I insulin-like growth factor receptor in the estrogen-stimulated proliferation of human breast cancer cells. *J Biol Chem* 1990;**265**:21172–8.

21 **Brockdorff BL**, Heiberg I, Lykkesfeldt AE. Resistance to different antiestrogens is caused by different multi-factorial changes and is associated with reduced expression of IGF receptor Ialpha. *Endocr Relat Cancer* 2003;**10**:579–90.

22 **Allred DC**, Harvey JM, Berardo M, *et al*. Prognostic and predictive factors in breast cancer by immunohistochemical analysis. *Mod Pathol* 1998;**11**:155–68.

23 **Harvey JM**, Clark GM, Osborne CK, *et al*. Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer. *J Clin Oncol* 1999;**17**:1474–81.

24 **Sasano H**, Edwards DP, Anderson TJ, *et al*. Validation of new aromatase monoclonal antibodies for immunohistochemistry: progress report. *J Steroid Biochem Mol Biol* 2003;**86**:239–44.

25 **Olsen KE**, Knudsen H, Rasmussen BB, *et al*. Amplification of HER2 and TOP2A and deletion of TOP2A genes in breast cancer investigated by new FISH probes. *Acta Oncol* 2004;**43**:35–42.

26 **Dowsett M**, Bartlett J, Ellis IO, *et al*. Correlation between immunohistochemistry (HercepTest) and fluorescence in situ hybridization (FISH) for HER-2 in 426 breast carcinomas from 37 centres. *J Pathol* 2003;**199**:418–23.

27 **Mouridsen H**, Gershanovich M, Sun Y, *et al*. Superior efficacy of letrozole versus tamoxifen as first-line therapy for postmenopausal women with advanced breast cancer: results of a phase III study of the International Letrozole Breast Cancer Group. *J Clin Oncol* 2001;**19**:2596–606.

28 **Jacobs TW**, Gown AM, Yaziji H, *et al*. Specificity of HercepTest in determining HER-2/neu status of breast cancers using the United States Food and Drug Administration-approved scoring system. *J Clin Oncol* 1999;**17**:1983–7.

29 **Johnston SR**, Saccani-Jotti G, Smith IE, *et al*. Changes in estrogen receptor, progesterone receptor, and pS2 expression in tamoxifen-resistant human breast cancer. *Cancer Res* 1995;**55**:3331–8.

30 **Tan KB**, Yong WP, Putti TC. Cyclooxygenase-2 expression: a potential prognostic and predictive marker for high-grade ductal carcinoma in situ of the breast. *Histopathology* 2004;**44**:24–8.

31 **Half E**, Tang XM, Gwyn K, *et al*. Cyclooxygenase-2 expression in human breast cancers and adjacent ductal carcinoma in situ. *Cancer Res* 2002;**62**:1676–81.

32 **Osborne CK**, Yochmowitz MG, Knight WA III, *et al*. The value of estrogen and progesterone receptors in the treatment of breast cancer. *Cancer* 1980;**46**:2884–8.

33 **List HJ**, Reiter R, Singh B, *et al*. Expression of the nuclear coactivator AIB1 in normal and malignant breast tissue. *Breast Cancer Res Treat* 2001;**68**:21–8.

34 **Liang J**, Zubovitz J, Petrocelli T, *et al*. PKB/Akt phosphorylates p27, impairs nuclear import of p27 and opposes p27-mediated G1 arrest. *Nat Med* 2002;**8**:1153–60.

35 **Ciocca DR**, Fujimura FK, Tandon AK, *et al*. Correlation of HER-2/neu amplification with expression and with other prognostic factors in 1103 breast cancers. *J Natl Cancer Inst* 1992;**84**:1279–82.

36 **Camp RL**, Charette LA, Rimm DL. Validation of tissue microarray technology in breast carcinoma. *Lab Invest* 2000;**80**:1943–9.

37 **Ellis MJ**, Coop A, Singh B, *et al*. Letrozole inhibits tumor proliferation more effectively than tamoxifen independent of HER1/2 expression status. *Cancer Res* 2003;**63**:6523–31.

38 **Abd El-Rehim DM**, Pinder SE, Paish CE, *et al*. Expression and co-expression of the members of the epidermal growth factor receptor (EGFR) family in invasive breast carcinoma. *Br J Cancer* 2004;**91**:1532–42.

39 **Tsutsui S**, Kataoka A, Ohno S, *et al*. Prognostic and predictive value of epidermal growth factor receptor in recurrent breast cancer. *Clin Cancer Res* 2002;**8**:3454–60.

40 **Walker RA**, Dearing SJ. Expression of epidermal growth factor receptor mRNA and protein in primary breast carcinomas. *Breast Cancer Res Treat* 1999;**53**:167–76.

41 **Klijn JG**, Berns PM, Schmitz PI, *et al*. The clinical significance of epidermal growth factor receptor (EGF-R) in human breast cancer: a review on 5232 patients. *Endocr Rev* 1992;**13**:3–17.

42 **Ouban A**, Muraca P, Yeatman T, *et al*. Expression and distribution of insulin-like growth factor-1 receptor in human carcinomas. *Hum Pathol* 2003;**34**:803–8.

43 **Happerfield LC**, Miles DW, Barnes DM, *et al*. The localization of the insulin-like growth factor receptor 1 (IGFR-1) in benign and malignant breast tissue. *J Pathol* 1997;**183**:412–17.

44 **Turner BC**, Haffty BG, Narayanan L, *et al*. Insulin-like growth factor-I receptor overexpression mediates cellular radioresistance and local breast cancer recurrence after lumpectomy and radiation. *Cancer Res* 1997;**57**:3079–83.