

Keywords: breast cancer; classification; immunohistochemistry; stratification

Identification of key clinical phenotypes of breast cancer using a reduced panel of protein biomarkers

A R Green^{*,1}, D G Powe², E A Rakha^{1,2}, D Soria^{3,4}, C Lemetre⁵, C C Nolan¹, F F T Barros¹, R D Macmillan⁶, J M Garibaldi^{3,4}, G R Ball⁵ and I O Ellis^{1,2}

¹Breast Cancer Pathology Research Group, Division of Oncology, School of Medicine, Academic Unit of Clinical Oncology, University of Nottingham, Nottingham City Hospital, Hucknall Road, Nottingham NG5 1PB, UK; ²Cellular Pathology, Nottingham University Hospitals NHS Trust, Nottingham, UK; ³School of Computer Science, University of Nottingham, Nottingham, UK; ⁴Advanced Data Analysis Centre, University of Nottingham, Nottingham, UK; ⁵School of Biomedical and Natural Sciences, Nottingham Trent University, Nottingham, UK and ⁶The Breast Institute, Nottingham University Hospitals NHS Trust, Nottingham, UK

Background: Breast cancer is a heterogeneous disease characterised by complex molecular alterations underlying the varied behaviour and response to therapy. However, translation of cancer genetic profiling for use in routine clinical practice remains elusive or prohibitively expensive. As an alternative, immunohistochemical analysis applied to routinely processed tissue samples could be used to identify distinct biological classes of breast cancer.

Methods: In this study, 1073 archival breast tumours previously assessed for 25 key breast cancer biomarkers using immunohistochemistry and classified using clustering algorithms were further refined using naïve Bayes classification performance. Criteria for class membership were defined using the expression of a reduced panel of 10 proteins able to identify key molecular classes. We examined the association between these breast cancer classes with clinicopathological factors and patient outcome.

Results: We confirm patient classification similar to established genotypic biological classes of breast cancer in addition to novel sub-divisions of luminal and basal tumours. Correlations between classes and clinicopathological parameters were in line with expectations and showed highly significant association with patient outcome. Furthermore, our novel biological class stratification provides additional prognostic information to the Nottingham Prognostic Index.

Conclusion: This study confirms that distinct molecular phenotypes of breast cancer can be identified using robust and routinely available techniques and both the luminal and basal breast cancer phenotypes are heterogeneous and contain distinct subgroups.

Breast cancer, which is the most common cancer in women (Parkin *et al*, 2001; Kamangar *et al*, 2006) is a complex disease characterised by multiple molecular alterations. The current clinical management of breast cancer relies on availability of robust clinicopathological and individual molecular prognostic and predictive factors to support decision making. However, the varied behaviour and response to therapy within the clinically and morphological similar classes indicate that the traditional prognostic factors currently available are insufficient to reflect the

genetic heterogeneity of breast tumours. Recent advances in high-throughput molecular technologies have further demonstrated this biological heterogeneity of breast cancer.

A seminal study by Perou *et al* (2000) identified four distinct molecular breast cancer groups based on gene expression profiles: luminal epithelial/oestrogen receptor (ER) positive, c-erb-B2 (HER2) positive, basal-like and normal breast-like. A subsequent study extended this by dividing the luminal/ER-positive group into three subtypes: luminal A, B, and C (Sorlie *et al*, 2001), but existence of the

*Correspondence: Dr AR Green; E-mail: andrew.green@nottingham.ac.uk

Received 11 June 2012; revised 2 July 2013; accepted 13 August 2013; published online 5 September 2013

© 2013 Cancer Research UK. All rights reserved 0007–0920/13

luminal C group remains uncertain (Sorlie *et al*, 2003). Sotiropoulos *et al* (2003) demonstrated six similar groups, with two basal-like subgroups and no normal breast-like group. While numerous studies have reported these and other novel molecular subtypes, and assigned a prognostic significance to these identified classes (West *et al*, 2001; van't Veer *et al*, 2002; Calza *et al*, 2006), they remain varied in their detailed classification (Hu *et al*, 2006). In addition, issues regarding the potential clinical utility of gene expression profiling include sample processing, data interpretation and analysis, reproducibility, validation, feasibility, and cost (Ein-Dor *et al*, 2006; Pusztai *et al*, 2006; Simon, 2006). Existing studies have also not addressed the stability of the proposed classifications across different case sets, the biological value of the different genes involved in the cluster designation and the proportion of cases that cannot be classified into any of the core molecular classes. Such an issue appears of critical relevance considering the need to identify the molecular features of individual tumours in routine practice.

An alternative approach is to use established robust laboratory technology, such as immunohistochemistry on formalin-fixed paraffin embedded (FFPE) patient tumour samples utilising a set of proteins with a well-defined biological and clinical relevance in breast cancer. Previously, we have applied a 25-protein biomarker panel with known relevance to breast cancer, to large numbers of cases using FFPE tissue microarrays and unsupervised clustering analysis, and have confirmed the existence and clinical significance of distinct breast cancer classes (Abd El-Rehim *et al*, 2005). In further analysis, we used a consensus methodology between three alternative clustering techniques, Hierarchical (Anderberg, 1973), K-means (Al-Daoud and Roberts, 1996), and Adaptive Resonance Theory (ART) (Carpenter and Grossberg, 1987), followed by the computation of several validity indices to independently verify the number of output clusters and address the issue of clustering stability and classification uncertainty. By examining the concordance between the clustering techniques, we determined a set of core six breast cancer classes (Soria *et al*, 2010). Concordance between clusters, assessed by conventional statistical techniques (principal component projections and boxplots) and two automated methods (an artificial neural network and a rule-extraction approach) were used to characterise these classes. This served to confirm that key biological classes of breast can be identified using an immunohistochemical panel of biomarkers and demonstrated that luminal and basal classes of breast cancer are heterogeneous and contain distinct subclasses. Of importance was the observation that only 60% of breast cancer cases clearly exhibited core class membership criteria, while the remaining 40% of cases were not assigned to a class. However, clinical adoption of such a classification system using 25 proteins determined by immunohistochemistry would be impractical due to cost and time constraints.

Further analysis was therefore required to minimise the number of markers needed to classify patients into these distinct breast cancer classes but retain the sophistication of the classification, maintain the clinical heterogeneity of classes, and reduce the unclassified tumours to a lower level while retaining usefulness for clinical decision making.

In this study, we therefore have sought to reduce the number of biomarkers necessary to classify breast cancer classes using immunohistochemistry and to examine the association between these classes with various clinical and pathological factors and patient outcome.

PATIENTS AND METHODS

Patients and laboratory methods. A series of 1073 patients from the Nottingham Tenovus Primary Breast Carcinoma Series presenting with primary operable (stages I, II and III) invasive

breast cancer between 1986 and 1998 were used. Immunohistochemical reactivity for 25 proteins, with known relevance in breast cancer including those used in routine clinical practice, were determined using standard immunohistochemical techniques on tumour samples prepared as FFPE tissue microarrays as previously described (Abd El-Rehim *et al*, 2005). Levels of immunohistochemical reactivity were determined by microscopical analysis using the modified H-score (values between 0 and 300), giving a semi-quantitative assessment of both the intensity of staining and the percentage of positive cells as previously described (Abd El-Rehim *et al*, 2005). For c-erb-B2, the American Society of Clinical Oncology/College of American Pathologists Guideline Recommendations for Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer were used for assessment (Wolff *et al*, 2007). Equivocal (2+) cases were confirmed by chromogenic *in situ* hybridisation as previously described (Garcia-Caballero *et al*, 2010).

The Nottingham Series is a well-characterised consecutive assembly of patients who were treated according to standard clinical protocols. Of the available cases, 708 (66%) cases were aged 50 years or more. At the time of diagnosis, 160 (14.9%) tumours were histological grade 1, 343 (31.9%) were grade 2, and 572 (53.2%) grade 3 (Table 1). A total of 736 (68.4%) had tumour size more than 1.5 cm. A total of 654 (60.8) patients had lymph node-negative disease and 419 (38.9%) had positive lymph nodes (332 cases with between one and three positive nodes, 87 cases with four or more positive nodes). Frequencies for histological tumour types were: 649 invasive ductal carcinomas of no special type, 171 tubular and tubular mixed carcinomas, 30 medullary carcinomas, 112 lobular carcinomas, 11 mucinous carcinomas, 37 mixed histological type, 3 papillary type carcinomas and four miscellaneous tumours. Patient management was based on tumour characteristics using Nottingham Prognostic Index (NPI) (Galea *et al*, 1992) and hormone receptor status. Patients with an NPI score of <3.4 received no adjuvant therapy, those with a NPI score >3.4 received hormone therapy if ER positive or classical cyclophosphamide, methotrexate and 5-fluorouracil if ER negative and fit enough to tolerate chemotherapy. Hormonal therapy was given to 420 patients (39%) and chemotherapy to 264 (24.5%). This study was approved by the Nottingham Research Ethics Committee 2 under the title 'Development of a molecular genetic classification of breast cancer'. The Reporting Recommendations for Tumour Marker Prognostic Studies (REMARK) criteria, recommended by McShane *et al* (2005), were followed.

Data relating to survival were collated in a prospective manner for those patients presenting after 1989 only; including survival time, defined as the interval (in months) from the date of the primary treatment to the time of death. Both short (5 years) and long (up to 20 years)-term patient outcome was investigated with respect to the biological classes in all patients where outcome data was available. Follow-up data was available for 974 patients, with overall survival ranging from 4 to 224 months (median 123 months, mean 118 months). During this period, a total of 317 (29.5%) patients developed distant metastases while 346 patients died, 263 of them from breast cancer. Patient age ranged from 18 to 72 years (median 54 years).

Classifying methods. Using the expression of 25 proteins, determined by immunohistochemistry, breast tumours were classified as previously described (Soria *et al*, 2010). Using the combination of boxplots for the whole data and for the singular classes and following clinical judgment, the 25 markers on which the breast cancer classes were derived were reduced to 14 by dismissing those biomarkers that had 'no role' in the class definition (i.e., had an identical overall distribution even when considered in any specific class). Then, using supervised classification approaches based on the naïve Bayes classification performance (Soria *et al*, 2008), the number of markers was further

Table 1. Breast cancer biological class and clinicopathological parameters (Note that *P*-values derived using Cramer's V (Friendly, 2000).)

| Parameter | Biological class | | | | | | | Cramer's V (P-value) | |
|------------------------------------|------------------|------------|-----------|-------------------|------------------|---------------|---------------|----------------------|-----------------|
| | Luminal A | Luminal N | Luminal B | Basal-p53 altered | Basal-p53 normal | HER2 + / ER + | HER2 + / ER - | | |
| | n (%) | n (%) | n (%) | n (%) | n (%) | n (%) | n (%) | | |
| Size | | | | | | | | | |
| < 15 mm | 106 (28.6) | 45 (30.8) | 27 (22.0) | 15 (11.9) | 10 (11.5) | 6 (10.0) | 10 (11.8) | 0.199 (5.782e-07) | |
| ≥ 15 mm | 264 (71.4) | 101 (69.2) | 95 (77.2) | 111 (88.1) | 77 (88.5) | 54 (90.0) | 75 (88.2) | | |
| Total | 370 | 146 | 122 | 126 | 87 | 60 | 85 | | |
| Grade | | | | | | | | | |
| 1 | 79 (21.4) | 41 (28.1) | 20 (16.3) | 1 (0.8) | 0 (0) | 1 (1.7) | 1 (1.2) | 0.405 (0.00) | |
| 2 | 160 (43.2) | 80 (54.8) | 48 (39.0) | 4 (3.2) | 8 (9.2) | 13 (21.7) | 10 (11.8) | | |
| 3 | 131 (35.4) | 25 (17.1) | 54 (43.9) | 121 (96.0) | 79 (90.8) | 46 (76.6) | 74 (87.0) | | |
| Total | 370 | 146 | 122 | 126 | 87 | 60 | 85 | | |
| Stage | | | | | | | | | |
| 1 | 227 (61.4) | 100 (68.5) | 69 (56.1) | 82 (65.1) | 56 (64.4) | 29 (48.3) | 38 (44.7) | 0.116 (0.008) | |
| 2 | 120 (32.4) | 37 (25.3) | 37 (30.1) | 33 (26.2) | 24 (27.6) | 25 (41.7) | 33 (38.8) | | |
| 3 | 22 (5.9) | 9 (6.2) | 15 (12.2) | 11 (8.7) | 7 (8.0) | 6 (10.0) | 14 (16.5) | | |
| Total | 369 | 146 | 121 | 126 | 87 | 60 | 85 | | |
| Tumour type | | | | | | | | | |
| Invasive ductal | 202 (54.6) | 37 (25.3) | 67 (54.5) | 103 (81.7) | 70 (80.5) | 53 (88.3) | 73 (85.9) | 0.23 (0.00) | |
| Invasive lobular | 28 (7.6) | 25 (17.1) | 12 (9.8) | 0 (0) | 1 (1.1) | 1 (1.7) | 0 (0) | | |
| Medullary | 1 (0.3) | 0 (0) | 0 (0) | 14 (11.1) | 9 (10.3) | 0 (0) | 5 (5.9) | | |
| Mixed type | 112 (30.3) | 74 (50.7) | 37 (30.1) | 6 (4.8) | 6 (6.9) | 6 (10.0) | 6 (7.1) | | |
| Mucinous | 8 (2.2) | 0 (0) | 1 (0.8) | 0 (0) | 0 (0) | 0 (0) | 0 (0) | | |
| Other | 0 (0) | 0 (0) | 0 (0) | 2 (1.6) | 0 (0) | 0 (0) | 0 (0) | | |
| Papillary | 1 (0.3) | 0 (0) | 1 (0.8) | 1 (0.8) | 0 (0) | 0 (0) | 0 (0) | | |
| Tubular | 16 (4.3) | 6 (4.1) | 4 (3.3) | 0 (0) | 0 (0) | 0 (0) | 0 (0) | | |
| Total | 364 | 142 | 122 | 126 | 86 | 60 | 84 | | |
| Nottingham Prognostic Index | | | | | | | | | |
| Excellent | 54 (14.6) | 29 (19.9) | 14 (11.4) | 0 (0) | 0 (0) | 1 (1.7) | 1 (1.2) | | 0.203 (0.00) |
| Good | 93 (25.1) | 49 (33.6) | 21 (17.1) | 2 (1.6) | 5 (5.7) | 5 (8.3) | 4 (4.7) | | |
| Moderate 1 | 103 (27.8) | 34 (23.3) | 30 (24.4) | 50 (39.7) | 22 (25.3) | 18 (30.0) | 24 (28.2) | | |
| Moderate 2 | 70 (18.9) | 24 (16.4) | 35 (28.5) | 45 (35.7) | 37 (42.5) | 18 (30.0) | 26 (30.6) | | |
| Poor | 39 (10.5) | 8 (5.5) | 15 (12.2) | 20 (15.9) | 21 (24.1) | 13 (21.7) | 19 (22.4) | | |
| Very poor | 8 (2.2) | 2 (1.4) | 5 (4.1) | 7 (5.6) | 1 (1.1) | 4 (6.7) | 9 (10.6) | | |
| Total | 367 | 146 | 120 | 124 | 86 | 59 | 83 | | |

reduced by an exhaustive search down to the minimum number of biomarkers required to retain the previous classification. This formed the basis of the development of a fuzzy rule induction algorithm using the methodology previously described in Rasmani *et al* (2009). The ultimate goal of this process was to create a single algorithmic process to classify a breast tumour into one of the six clinical classes, while reducing the number of unclassified patients to a minimum. Technical class characterisation in terms of marker distribution was performed as previously described (Soria *et al*, 2010). The rule induction algorithm was subsequently validated on the expression of the 10 biomarkers determined by IHC in an additional 238 unselected cases of primary breast cancer from the Nottingham Tenovus primary breast cancer series.

Statistical analysis. The association between breast cancer classes and both histopathological and clinical characteristics, which were not involved in the development of the classes, was assessed using Cramer's V (Friendly, 2000) to produce *P*-values. Breast cancer-specific survival and disease-free survival (DFS) between classes

was determined using Kaplan–Meier curves. Non-parametric Kruskal–Wallis was used to test the difference of NPI across classes.

RESULTS

Breast cancer classification. Using an exhaustive search of the best combination of a reduced set of 14 biomarkers, the minimum number required for classification of breast tumours into the six classes was reduced from 25 biomarkers to 10. These biomarkers were ER, progesterone receptor, cytokeratin (CK) 5/6, cytokeratin 7/8, EGFR, HER2, c-erbB3 (HER3), c-erbB4 (HER4), p53, and Mucin 1. Using a fuzzy rule induction-derived algorithm, a total of 997 out of 1073 (93%) breast tumours were subsequently assigned to one of the six classes and contained 370, 146, 123, 126, 87, and 145 patients, respectively. The remaining patients ($n = 76$, 7%) were not assigned to any class.

Table 2. Comparison of the distribution of breast tumours by classification using 25 and 10 biomarkers

| Classification (25 biomarkers) | Classification (10 biomarkers) | | | | | |
|--------------------------------|--------------------------------|---------------|---------------|-----------------------|----------------------|-----------|
| | Luminal A (n) | Luminal N (n) | Luminal B (n) | Basal-p53 altered (n) | Basal-p53 normal (n) | HER2+ (n) |
| Luminal A | 173 | 3 | 14 | 0 | 1 | 4 |
| Luminal N | 29 | 99 | 5 | 0 | 0 | 2 |
| Luminal B | 31 | 3 | 28 | 1 | 3 | 9 |
| Basal-p53 altered | 0 | 0 | 0 | 79 | 0 | 1 |
| Basal-p53 normal | 1 | 0 | 3 | 17 | 35 | 3 |
| HER2+ | 0 | 0 | 2 | 6 | 4 | 63 |

Kappa index = 0.79; Kendall's tau correlation coefficient = 0.92.

Table 3. Performance of principal component analysis

| | Component | | | | | | | | | |
|------------------------|-----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| s.d. | 1.649 | 1.298 | 1.077 | 0.966 | 0.881 | 0.841 | 0.776 | 0.736 | 0.678 | 0.644 |
| Proportion of variance | 0.272 | 0.168 | 0.116 | 0.093 | 0.078 | 0.071 | 0.060 | 0.054 | 0.046 | 0.042 |
| Cumulative proportion | 0.272 | 0.440 | 0.557 | 0.650 | 0.727 | 0.798 | 0.858 | 0.913 | 0.96 | 1.000 |

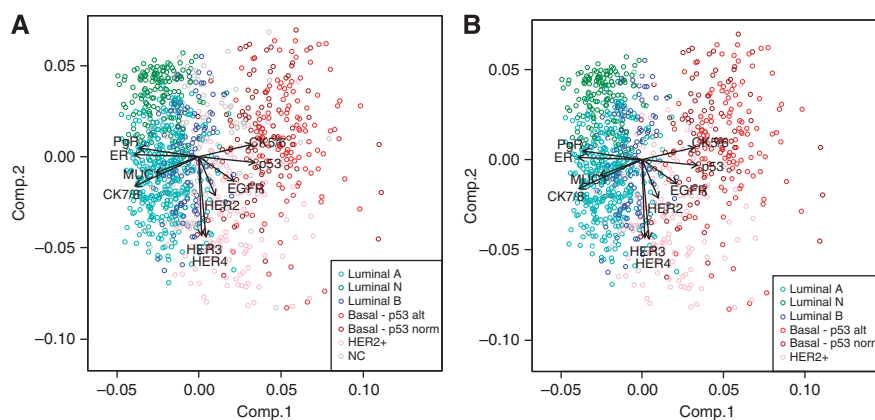


Figure 1. Biplots of classes projected on the first and second principal component axes: (A) for all patients and (B) for only patients assigned to a class.

Having removed unclassified tumours, we compared the classification of the 619 remaining tumours between the classification using 25 and 10 biomarkers. There was a good agreement between classes (Table 2, kappa index = 0.79, Kendall's tau correlation coefficient = 0.92). Any shift in classification between tumours primarily occurred within the main (luminal and basal) groups.

The classification approach was validated using an additional set of 238 patients for which IHC data for the ten biomarkers was available. Results showed that the breast cancer classes were populated with similar distribution to our initial findings (68, 61, 36, 23, 24, and 25 patients, respectively) and presented clear characteristics of the six classes.

Class characterisation. Figure 1 provides a visualisation of the separation of the classes and the association with the biomarkers. Table 3 summarises the performance of the principal component analysis where the first two components accounted for 44% of the

variance. Figure 1A shows the biplot obtained for all patients, in which those not assigned to any class (unassigned) have been coloured grey. It can be seen that these fall mainly into the centre region of the biplot. Figure 1B shows the biplot obtained for only patients assigned to all classes—except the unassigned cases. The first axis was mainly determined, on the left, by luminal markers including luminal cytokeratin (CK7/8), hormone receptors (ER and PgR), and MUC1 overexpression and, on the right, by basal cytokeratin (CK5/6) and partly by p53 overexpression. The second axis is determined, on the bottom, by HER2, HER3, and HER4 overexpression.

The core molecular classes identified in this study included three luminal class tumours characterised by high luminal CK7/8 and hormone receptor (ER and PgR) expression. Luminal A and luminal B tumours showed high expression of CK7/8, ER, HER3 and HER4 but were separated by relatively lower levels of PgR expression in luminal B compared with luminal A tumours. In contrast, luminal N tumours showed differential expression of

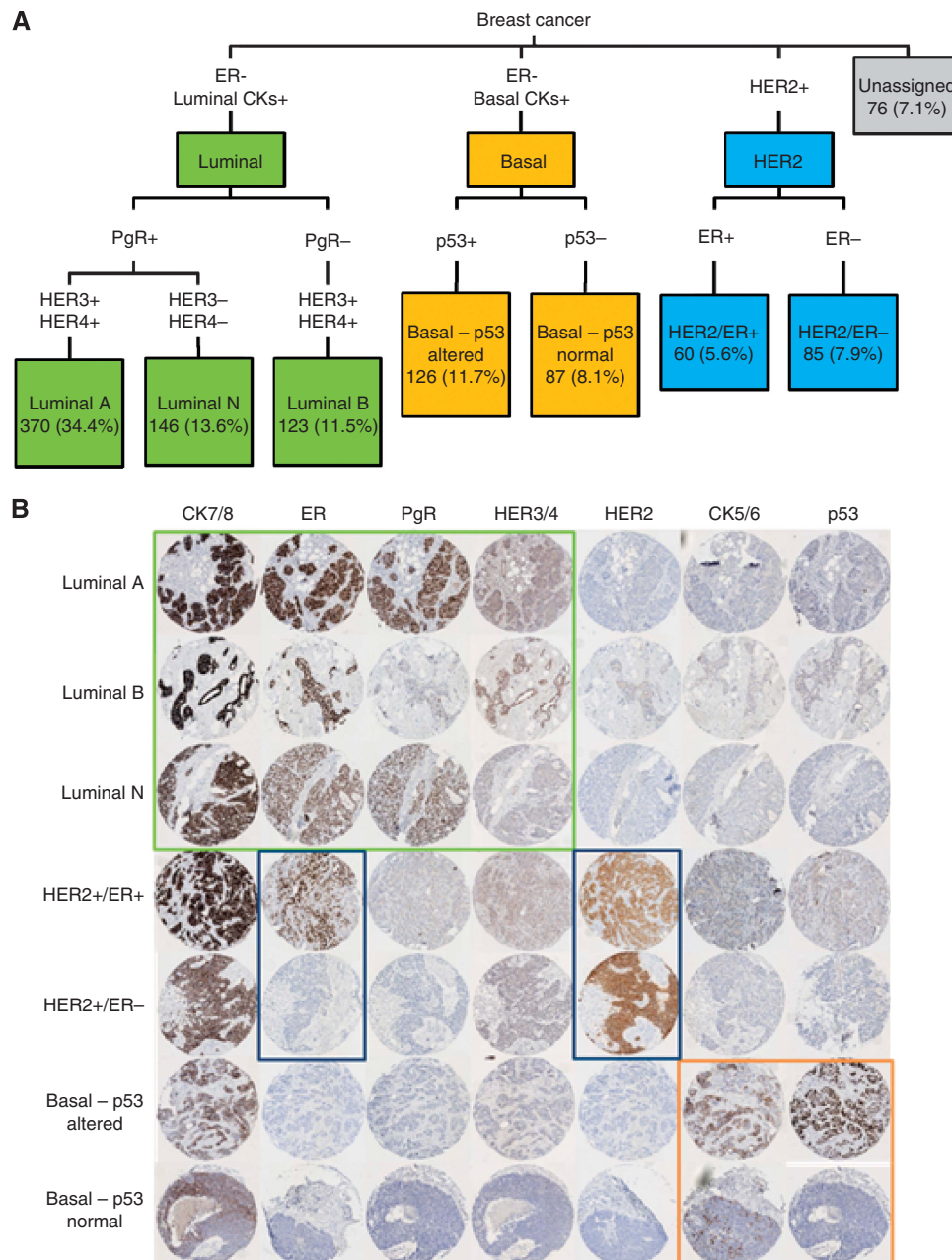


Figure 2. Breast cancer biological classes. (A) Classification and proportions of cases. (B) Representative immunohistochemical profiles of the biological classes of breast cancer.

HER3 and HER4. There were two basal classes of tumour, characterised by low luminal cytokeratin and high basal expression (CK5/6) along with showing a triple-negative phenotype (i.e., ER, PgR and HER2 negative). They were, however, separated by p53 protein expression levels where the tumours either expressed high p53 (basal-p53 altered) or low p53 (basal-p53 normal). The remaining class (HER2) was characterised by high luminal cytokeratin and HER2 overexpression but showed heterogeneity in the expression of hormone receptors. We therefore split these tumours into two subclasses those expressing ER (HER2 + /ER +) and those that showed an ER - phenotype (HER2 + /ER -). The unassigned tumours showed heterogenic expression of all ten markers. A summary of the breast cancer classes and relative biomarker expression is shown in Figure 2A and B.

Clinical characterisation of patients by class. Significant associations were found between the identified breast cancer classes with

respect to tumour grade, size, lymph node stage, and histological tumour type and were in line with expectations (Table 1). The basal and HER2 classes were significantly associated with larger tumour size, higher grade, higher stage, invasive ductal carcinoma of no special type and the poorer NPI prognostic groups. The luminal classes were significantly associated with tubular, lobular, and mixed-type tumours where luminal N tumours tended to be of lobular and mixed type compared with luminal A and B tumours, which were ductal of no special type.

Short- and long-term patient outcome was compared by tumour class where each class was significantly associated with distinct breast cancer-specific survival (BCSS) rates (Figure 3A and B, Table 4) and DFS (Figure 3C and D, Table 5). The highest frequency of mortality due to breast cancer in the first 5 years was seen in patients whose tumours belonged to the HER2 classes, with the HER2/ER - being the worst. A lower, but still high, frequency was seen in patients with tumours from luminal B and basal

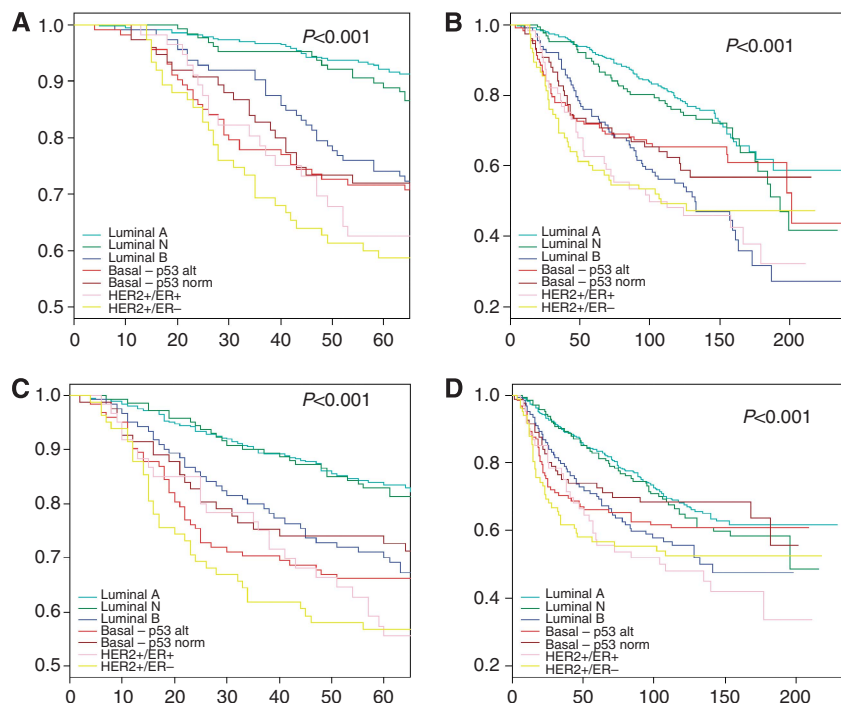


Figure 3. Breast cancer biological classes in relation to (A) 5-, (B) 20-year breast cancer-specific survival and (C) 5-, (D) 20-year disease-free survival.

Table 4. Kaplan–Meier P -values between breast cancer classes and BCSS

| | Breast cancer class | | | | | |
|------------------------|---------------------|-----------|-----------|-------------------|------------------|-----------|
| | Luminal A | Luminal N | Luminal B | Basal-p53 altered | Basal-p53 normal | HER2/ER + |
| BCSS (5 years) | | | | | | |
| Luminal N | 0.447 | | | | | |
| Luminal B | 0.508 | 0.169 | | | | |
| Basal-p53 altered | 0.00199* | 0.00147* | 0.00419* | | | |
| Basal-p53 normal | 0.0222 | 0.0152 | 0.0473 | 0.294 | | |
| HER2/ER + | 0.205 | 0.0603 | 0.553 | 0.0578 | 0.44 | |
| HER2/HER – | 0.0418 | 0.0337 | 0.0861 | 0.328 | 0.933 | 0.429 |
| BCSS (20 years) | | | | | | |
| Luminal N | 0.432 | | | | | |
| Luminal B | 7.74e-09* | 7.57e-05* | | | | |
| Basal-p53 altered | 0.00906* | 0.182 | 0.0483 | | | |
| Basal-p53 normal | 0.00594* | 0.128 | 0.144 | 0.805 | | |
| HER2/ER + | 1.57e-07* | 0.000246* | 0.616 | 0.0425 | 0.0823 | |
| HER2/HER – | 3.07e-07* | 0.00123* | 0.775 | 0.074 | 0.130 | 0.768 |

* $P < 0.01$.

classes. Luminal A and N had the lowest frequency of breast cancer-specific death. Over a period of 20 years, luminal B and HER2+/ER+ tumours showed the worst BCSS (Figure 3B). A similar pattern of outcome in the biological classes was observed for DFS as for BCSS although the basal-p53 normal class had a longer disease-free interval compared with the basal-p53 altered tumours during the first five years.

A boxplot of the NPI split by class is shown in Figure 4 illustrating significant differences between the NPI score and

biological class (overall Kruskal–Wallis $P < 0.01$). Table 6 summarises P -values for the class-by-class comparisons. It can be seen that the NPI for luminal N tumours is lower than that of the other luminal classes (luminal A and B). HER2+/ER– class of tumours have a higher NPI score than the other classes. This is an interesting observation for two reasons. First, it confirms that the NPI is providing discriminant information between the biological classes. Second, it suggests that the class divisions are providing additional information to the NPI.

Table 5. Kaplan–Meier *P*-values between breast cancer classes and DFS

| | Breast cancer class | | | | | |
|-----------------------|---------------------|-----------|-----------|-------------------|------------------|-----------|
| | Luminal A | Luminal N | Luminal B | Basal–p53 altered | Basal–p53 normal | HER2/ER + |
| DFS (5 years) | | | | | | |
| Luminal N | 0.755 | | | | | |
| Luminal B | 0.382 | 0.623 | | | | |
| Basal–p53 altered | 2.34e-06* | 5.1e-05* | 0.000635* | | | |
| Basal–p53 normal | 0.0518 | 0.184 | 0.233 | 0.108 | | |
| HER2/ER + | 0.749 | 0.668 | 0.631 | 0.000522* | 0.21 | |
| HER2/HER – | 7.93e-05* | 0.00354* | 0.00787* | 0.606 | 0.309 | 0.00408* |
| DFS (20 years) | | | | | | |
| Luminal N | 0.506 | | | | | |
| Luminal B | 0.002* | 0.0379 | | | | |
| Basal–p53 altered | 0.0826 | 0.387 | 0.467 | | | |
| Basal–p53 normal | 0.621 | 0.986 | 0.112 | 0.416 | | |
| HER2/ER + | 0.000127* | 0.00434* | 0.306 | 0.0965 | 0.0216 | |
| HER2/HER – | 0.000928* | 0.0263 | 0.488 | 0.227 | 0.0763 | 0.671 |

**P* < 0.01.

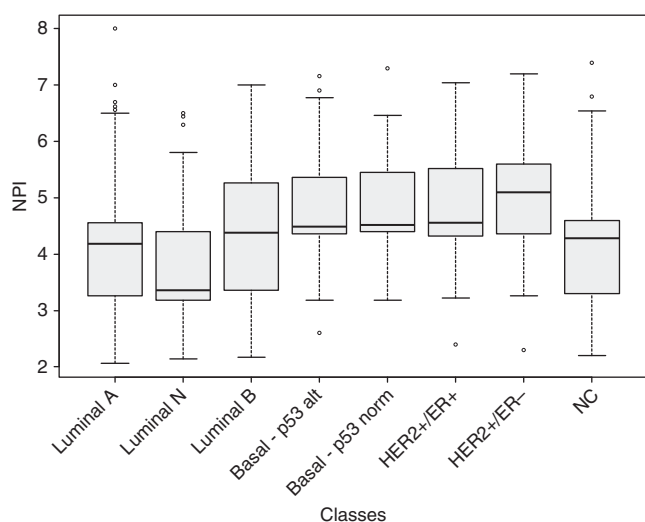


Figure 4. Boxplots of NPI by biological class.

DISCUSSION

We have identified core classes of breast cancer using a reduced panel of 10 protein biomarkers determined by immunohistochemistry which, we believe, are clinically meaningful and clinically well-characterised. Three of these classes have not been previously identified and, while their precise prognostic and therapeutic relevance is not yet clear, their elucidation serves as a basis for ongoing investigations to address these important factors. Of course, different clustering techniques can and will result in different clusters. It is for precisely this reason that we have previously used concordance between multiple methods to establish the core classes: the identification of stable clusters through multiple methods forms a basis of methodological validation (Soria *et al*, 2010). A further study is underway to refine these classes using a more sophisticated fuzzy rule induction algorithm (Soria *et al*, 2013).

Also of importance is the observation that 93% of breast cancer cases clearly exhibit core class membership criteria, whereas only 7.1% remain unclassified. As a matter of fact, when core classes were derived in Soria *et al* (2010) the number of unclassified patients was 413, whereas after applying Rasmani *et al*'s (2009) approach this number went down to 76. Some of these unclassified cases result from the stringent criteria used to derive class membership but it is important to recognise that other cases do exhibit characteristics of more than one class. Biologically, this observation may be explained by the complex heterogeneity of the molecular portrait of breast cancer. These mixed class cases clearly merit recognition and investigation to determine their optimal treatment strategies. However, the well-defined biological and clinical relevance of the markers used, association with clinicopathological variables, and outcome indicates the biological and clinical relevance of the current molecular classification.

The core molecular classes identified in this study are similar to those determined by gene expression profiling, but we have been able to refine the definition of the luminal and basal tumours into further distinct classes with different clinical outcome. Phenotypic classification into core luminal, basal, and HER2 classes is possible using smaller panels of three to five antibodies (Nielsen *et al*, 2004; Carey *et al*, 2006; Cheang *et al*, 2008; Blows *et al*, 2010) but such limited panels cannot further sub classify these core groups. Our study clearly demonstrates that using a larger panel of 10 biomarkers a higher level of stratification is achieved, which may have direct and important clinical relevance.

The biological characterisation of the luminal A tumours are consistent with our previously identified group 1 (Abd El-Rehim *et al*, 2005) and Ambrogi's group 1 (Ambrogi *et al*, 2006), but are more distinctly defined. These tumours show high homology to the luminal-A-type tumours, as identified in gene array studies, which are also characterised by high gene expression of luminal differentiation, ER signalling markers and those involved with EGF (epidermal growth factor) signalling (Sorlie *et al*, 2006). Luminal B tumours were also characterised by higher levels of EGFR, HER3, and HER4 but, in contrast, showed relatively lower levels of PgR (levels of ER were similar). This class was not identified by our previous study (Abd El-Rehim *et al*, 2005), but is similar to group 2 in the Ferrara series (Ambrogi *et al*, 2006). This class of tumour shows homology to the luminal B group of

Table 6. P-values of Kruskal–Wallis analysis between breast cancer classes and Nottingham Prognostic Index

| | Luminal A | Luminal N | Luminal B | Basal–p53 altered | Basal–p53 normal | HER2/ER + |
|-------------------|------------|------------|------------|-------------------|------------------|-----------|
| Luminal N | 0.002* | | | | | |
| Luminal B | 0.010 | 5.927e-06* | | | | |
| Basal–p53 altered | 1.138e-13* | <2.2e-16* | 0.0006* | | | |
| Basal–p53 normal | 1.324e-11* | 1.294e-15* | 0.0005* | 0.326 | | |
| HER2/ER + | 1.313e-07* | 3.135e-11* | 0.004* | 0.751 | 0.723 | |
| HER2/ER – | 9.309e-14* | <2.2e-16* | 4.381e-06* | 0.042 | 0.188 | 0.185 |

*P<0.01.

tumours, defined by low-to-moderate expression of luminal-specific genes, including the ER cluster (Sorlie *et al*, 2006).

Whereas gene expression profiling has determined two luminal tumour classes, we have divided those tumours with a luminal phenotype into a further class (luminal N). This novel class of tumours while having high levels of ER and PgR, have negative/low expression of the HER family, particularly HER3 and HER4. Interestingly, the luminal A and luminal N tumours were similarly associated with good prognostic factors, including smaller tumour size, grade 1 tumours, node-negative and tubular mixed carcinomas. In contrast, the third luminal class, luminal B, although phenotypically similar to the luminal A tumours (except for PgR expression), consisted of those tumours with poorer prognostic factors such as larger tumour size, higher stage, and grade. It is also apparent that HER3 and HER4 are important discriminators in our breast cancer classification, although there remains controversy as to their prognostic significance (Witton *et al*, 2003; Abd El-Rehim *et al*, 2004).

We previously identified a basal-like subtype using protein expression (group 5 (Abd El-Rehim *et al*, 2005) and Ambrogio's group 3 (Ambrogio *et al*, 2006)). However, consistent with other studies that showed that the basal-like subtype is heterogeneous (Laakso *et al*, 2006), we have now determined two basal-like classes. These classes were characterised by high expression of basal cytokeratin (CK5/6), low expression of luminal cytokeratin and a triple-negative phenotype. They were, however, separated by p53 protein expression level, which is similar to our observation in triple-negative breast cancer (Biganzoli *et al*, 2011). High frequency of tumour suppressor p53 mutations and protein expression have previously been detected in the basal-like subtype (Ellis *et al*, 1999; Sorlie *et al*, 2001; Pollack *et al*, 2002). The association between medullary carcinomas, p53 and basal tumours has been previously demonstrated (de Cremoux *et al*, 1999; Sorlie *et al*, 2001).

Those tumours with high HER2 expression were clustered into one class, which is homologous to Sorlie's HER2 group (Sorlie *et al*, 2001) and Ambrogio's group 4 (Ambrogio *et al*, 2006). This class has the worst overall survival. However, due to the heterogeneity of HER2 overexpression in both ER-positive and -negative tumours, we manually split this class into two classes those with (HER2 + / ER +) or without (HER2 + / ER –) ER expression.

In conclusion, we have previously applied different clustering techniques with validity indices, and used cluster consensus to derive a classification that is robust across different multivariate procedures. This has highlighted the dangers of relying on a single clustering technique, as has often been the case in other studies, particularly in the interpretation and management of high-throughput assay results. As a consequence, we have now further refined the classification of breast cancer based on a panel of 10 proteins assessed by immunohistochemistry to identify distinct biological classes of breast cancer. Using 10 biomarkers has produced a more realistic distribution of breast cancer patients into the three main classes of luminal, basal and HER2, which were established in Sorlie *et al* (2001) and confirmed by subsequent

papers (Abd El-Rehim *et al*, 2004). The fuzzy approach used to derive the classification with 10 markers has produced a breast cancer classification consistent with the proportion of cancer subtypes reported in other studies, while the previous classification with 25 markers assigned, for example, only 7% of patients to the HER2 + group. We have shown that, in addition to the luminal A, luminal B and HER2 classes identified using gene expression analysis, there are three further biologically and clinically relevant classes of breast cancer, namely the luminal N, basal–p53- altered, and basal–p53 normal classes. Furthermore, we have confirmed the complex biological heterogeneity of breast cancer through the identification of cases exhibiting characteristics of mixed class. Studies are underway to further validate these classes and to enable the creation of a clinically usable algorithm for prospective classification, taking into account current therapeutic strategies.

ACKNOWLEDGEMENTS

This study was, in part, funded by the Breast Cancer Campaign (ARG, IOE) and was supported by the BIOPATTERN FP6 Network of Excellence (FP6-IST-508803) and the BIOPTRAIN FP6 Marie-Curie EST Fellowship (FP6-007597). We thank Chris Sheehan, Caroline Sykes and Sharon Crouch.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Abd El-Rehim DM, Ball G, Pinder SE, Rakha E, Paish C, Robertson JF, Macmillan D, Blamey RW, Ellis IO (2005) High-throughput protein expression analysis using tissue microarray technology of a large well-characterised series identifies biologically distinct classes of breast cancer confirming recent cDNA expression analyses. *Int J Cancer* **116**(3): 340–3450.
- Abd El-Rehim DM, Pinder SE, Paish CE, Bell JA, Rampaul RS, Blamey RW, Robertson JF, Nicholson RI, Ellis IO (2004) Expression and co-expression of the members of the epidermal growth factor receptor (EGFR) family in invasive breast carcinoma. *Br J Cancer* **91**(8): 1532–1542.
- Al-Daoud M, Roberts S (1996) New methods for the initialisation of clusters. *Pattern Recogn Lett* **17**(5): 451–455.
- Anderberg MR (1973) Cluster Analysis for Applications. *Monographs and Textbooks on Probability and Mathematical Statistics*. Academic Press, Inc.: New York.
- Ambrogio F, Biganzoli E, Querzoli P, Ferretti S, Boracchi P, Alberti S, Marubini E, Nenci I (2006) Molecular subtyping of breast cancer from traditional tumor marker profiles using parallel clustering methods. *Clin Cancer Res* **12**(3 Pt 1): 781–790.

- Biganzoli E, Coradini D, Ambrogi F, Garibaldi JM, Lisboa P, Soria D, Green AR, Pedriali M, Piantelli M, Querzoli P, Demicheli R, Boracchi P, Nenci I, Ellis IO, Alberti S (2011) p53 status identifies two subgroups of triple-negative breast cancers with distinct biological features. *Jpn J Clin Oncol* **41**(2): 172–179.
- Blows FM, Driver KE, Schmidt MK, Broeks A, van Leeuwen FE, Wesseling J, Cheang MC, Gelmon K, Nielsen TO, Blomqvist C, Heikkilä P, Heikkinen T, Nevanlinna H, Akslen LA, Begun LR, Foulkes WD, Couch FJ, Wang X, Cafourek V, Olson JE, Baglietto L, Giles GG, Severi G, McLean CA, Southey MC, Rakha E, Green AR, Ellis IO, Sherman ME, Lissowska J, Anderson WF, Cox A, Cross SS, Reed MW, Provenzano E, Dawson SJ, Dunning AM, Humphreys M, Easton DF, Garcia-Closas M, Caldas C, Pharoah PD, Huntsman D (2010) Subtyping of breast cancer by immunohistochemistry to investigate a relationship between subtype and short and long term survival: a collaborative analysis of data for 10,159 cases from 12 studies. *PLoS Med* **7**(5): e1000279.
- Calza S, Hall P, Auer G, Bjohle J, Klaar S, Kronenwett U, Liu ET, Miller L, Ploner A, Smeds J, Bergh J, Pawitan Y (2006) Intrinsic molecular signature of breast cancer in a population-based cohort of 412 patients. *Breast Cancer Res* **8**(4): R34.
- Carey LA, Perou CM, Livasy CA, Dressler LG, Cowan D, Conway K, Karaca G, Troester MA, Tse CK, Edmiston S, Deming SL, Geradts J, Cheang MC, Nielsen TO, Moorman PG, Earp HS, Millikan RC (2006) Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *JAMA* **295**(21): 2492–2502.
- Carpenter G, Grossberg S (1987) ART2: Stable selforganization of pattern recognition codes for analogue input patterns. *Appl Optics* **26**: 4919–4930.
- Cheang MC, Voduc D, Bajdik C, Leung S, McKinney S, Chia SK, Perou CM, Nielsen TO (2008) Basal-like breast cancer defined by five biomarkers has superior prognostic value than triple-negative phenotype. *Clin Cancer Res* **14**(5): 1368–1376.
- de Cremoux P, Salomon AV, Liva S, Dendale R, Bouchind'homme B, Martin E, Sastre-Garau X, Magdelenat H, Fourquet A, Soussi T (1999) p53 mutation as a genetic trait of typical medullary breast carcinoma. *J Natl Cancer Inst* **91**(7): 641–643.
- Ein-Dor L, Zuk O, Domany E (2006) Thousands of samples are needed to generate a robust gene list for predicting outcome in cancer. *Proc Natl Acad Sci USA* **103**(15): 5923–5928.
- Ellis IO, Pinder SE, Lee AH, Elston CW (1999) A critical appraisal of existing classification systems of epithelial hyperplasia and in situ neoplasia of the breast with proposals for future methods of categorization: where are we going? *Semin Diagn Pathol* **16**(3): 202–208.
- Everitt BS (2002) *The Cambridge Dictionary of Statistics*. Cambridge University Press.
- Friendly M (2000) *Visualizing Categorical Data*. SAS Institute: Cary, NC.
- Galea MH, Blamey RW, Elston CE, Ellis IO (1992) The Nottingham Prognostic Index in primary breast cancer. *Breast Cancer Res Treat* **22**(3): 207–219.
- Garcia-Caballero T, Grabau D, Green AR, Gregory J, Schad A, Kohlwe E, Ellis IO, Watts S, Mollerup J (2010) Determination of HER2 amplification in primary breast cancer using dual-colour chromogenic in situ hybridization is comparable to fluorescence in situ hybridization: a European multicentre study involving 168 specimens. *Histopathology* **56**(4): 472–480.
- Hu Z, Fan C, Oh DS, Marron JS, He X, Qaqish BF, Livasy C, Carey LA, Reynolds E, Dressler L, Nobel A, Parker J, Ewend MG, Sawyer LR, Wu J, Liu Y, Nanda R, Tretiakova M, Ruiz Orrico A, Dreher D, Palazzo JP, Perreard L, Nelson E, Mone M, Hansen H, Mullins M, Quackenbush JF, Ellis MJ, Olopade OI, Bernard PS, Perou CM (2006) The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics* **7**: 96.
- Kamangar F, Dores GM, Anderson WF (2006) Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. *J Clin Oncol* **24**(14): 2137–2150.
- Laakso M, Tanner M, Nilsson J, Wiklund T, Erikstein B, Kellokumpu-Lehtinen P, Malmstrom P, Wilking N, Bergh J, Isola J (2006) Basolateral carcinoma: a new biologically and prognostically distinct entity between basal and luminal breast cancer. *Clin Cancer Res* **12**(14 Pt 1): 4185–4191.
- McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM (2005) Reporting recommendations for tumor marker prognostic studies. *J Clin Oncol* **23**(36): 9067–9072.
- Nielsen TO, Hsu FD, Jensen K, Cheang M, Karaca G, Hu Z, Hernandez-Boussard T, Livasy C, Cowan D, Dressler L, Akslen LA, Ragaz J, Gown AM, Gilks CB, van de Rijn M, Perou CM (2004) Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. *Clin Cancer Res* **10**(16): 5367–5374.
- Parkin DM, Bray F, Ferlay J, Pisani P (2001) Estimating the world cancer burden: Globocan 2000. *Int J Cancer* **94**(2): 153–166.
- Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lonning PE, Borresen-Dale AL, Brown PO, Botstein D (2000) Molecular portraits of human breast tumours. *Nature* **406**(6797): 747–752.
- Pollack JR, Sorlie T, Perou CM, Rees CA, Jeffrey SS, Lonning PE, Tibshirani R, Botstein D, Borresen-Dale AL, Brown PO (2002) Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. *Proc Natl Acad Sci USA* **99**(20): 12963–12968.
- Pusztai L, Mazouni C, Anderson K, Wu Y, Symmans WF (2006) Molecular classification of breast cancer: limitations and potential. *Oncologist* **11**(8): 868–877.
- Rasmani K, Garibaldi J, Shen Q, Ellis I (2009) Linguistic Rulesets Extracted from a Quantifier-based Fuzzy Classification System. *Proc IEEE Int Conf Fuzzy Syst*, 11–13 December 2008. San Diego California, USA, pp 1204–1209.
- Simon R (2006) Validation of pharmacogenomic biomarker classifiers for treatment selection. *Cancer Biomark* **2**(3–4): 89–96.
- Soria D, Biganzoli E, Ellis I, Garibaldi JM (2008) A comparison of three different methods for classification of breast cancer data. *Seventh International Conference on Machine Learning and Applications* 619–624.
- Soria D, Garibaldi J, Green AR, Powe D, Nolan CC, Lemetre C, Ball G, Ellis I (2013) A quantifier-based fuzzy classification system for breast cancer patients. *Artificial Intelligence Med* **58**(3): 175–184.
- Soria D, Garibaldi JM, Ambrogi F, Green AR, Powe D, Rakha E, Macmillan RD, Blamey RW, Ball G, Lisboa PJ, Etschells TA, Boracchi P, Biganzoli E, Ellis IO (2010) A methodology to identify consensus classes from clustering algorithms applied to immunohistochemical data from breast cancer patients. *Comput Biol Med* **40**(3): 318–330.
- Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Eystein Lonning P, Borresen-Dale AL (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA* **98**(19): 10869–10874.
- Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, Deng S, Johnsen H, Pesich R, Geisler S, Demeter J, Perou CM, Lonning PE, Brown PO, Borresen-Dale AL, Botstein D (2003) Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci USA* **100**(14): 8418–8423.
- Sorlie T, Wang Y, Xiao C, Johnsen H, Naume B, Samaha RR, Borresen-Dale AL (2006) Distinct molecular mechanisms underlying clinically relevant subtypes of breast cancer: gene expression analyses across three different platforms. *BMC Genomics* **7**: 127.
- Sotiriou C, Neo SY, McShane LM, Korn EL, Long PM, Jazaeri A, Martiat P, Fox SB, Harris AL, Liu ET (2003) Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *Proc Natl Acad Sci USA* **100**(18): 10393–10398.
- van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH (2002) Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415**(6871): 530–536.
- West M, Blanchette C, Dressman H, Huang E, Ishida S, Spang R, Zuzan H, Olson Jr. JA, Marks JR, Nevins JR (2001) Predicting the clinical status of human breast cancer by using gene expression profiles. *Proc Natl Acad Sci USA* **98**(20): 11462–11467.
- Witton CJ, Reeves JR, Going JJ, Cooke TG, Bartlett JM (2003) Expression of the HER1-4 family of receptor tyrosine kinases in breast cancer. *J Pathol* **200**(3): 290–297.
- Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, Dowsett M, Fitzgibbons PL, Hanna WM, Langer A, McShane LM, Paik S, Pegram MD, Perez EA, Press MF, Rhodes A, Sturgeon C, Taube SE, Tubbs R, Vance GH, van de Vijver M, Wheeler TM, Hayes DF (2007) American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J Clin Oncol* **25**(1): 118–145.

This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution-NonCommercial-Share Alike 3.0 Unported License.