X-Tile: A New Bio-Informatics Tool for Biomarker Assessment and Outcome-Based Cut-Point Optimization

Robert L. Camp,[1] Marisa Dolled-Filhart,[2] and David L. Rimm[1]

Departments of [1]Pathology and [2]Genetics, Yale University, School of Medicine, New Haven, Connecticut

ABSTRACT

The ability to parse tumors into subsets based on biomarker expression has many clinical applications; however, there is no global way to visualize the best cut-points for creating such divisions. We have developed a graphical method, the X-tile plot that illustrates the presence of substantial tumor subpopulations and shows the robustness of the relationship between a biomarker and outcome by construction of a two dimensional projection of every possible subpopulation. We validate X-tile plots by examining the expression of several established prognostic markers (human epidermal growth factor receptor-2, estrogen receptor, p53 expression, patient age, tumor size, and node number) in cohorts of breast cancer patients and show how X-tile plots of each marker predict population subsets rooted in the known biology of their expression.

INTRODUCTION

In theory, associations between tumor biomarker expression and patient outcome should reveal the existence of biologically meaningful tumor classifications; however in practice, there is no universal method for discovering, assessing or displaying such associations. Consequently, studies generally group tumors into set divisions (*e.g.,* quartiles, deciles), which fails to reflect the underlying biology of most markers. Microarray technology has magnified this problem by accelerating the discovery of tumor markers for which there is no known biological basis. In such cases, there is no foreknowledge about how (or whether) a marker parses a population into subsets.

Although there is a range of statistical methods for cut-point selection (1, 2), we sought to produce a method that is comprehensive, based on traditional statistical tests, and yet intuitive for the oncologist. Here we describe our solution of cut-point selection using a new method that we have named "X-tile." X-tile plots provide a single, global assessment of every possible way of dividing a population into low-, medium-, and high-level marker expression. X-tile data are presented in a right triangular grid where each point represents a different cut-point. The intensity of the color of each cutoff point represents the strength of the association. The X-tile software allows the user to move a cursor across the grid and provides an "on-the-fly" histogram of the resulting population subsets along with an associated Kaplan-Meier curve.[3] This type of graphical representation can provide insight into the biological nature of a marker (*e.g.,* does it show a linear distribution relative to survival, does it reveal distinct sub-populations, or does it manifest a U-shaped relationship to outcome). Because it is statistically invalid to test multiple divisions and accept the best *P* value (3, 4), rigorous statistical evaluation is achieved by defining divisions in a "training set" and then validating them in a separate patient cohort ("validation set"). The X-tile software provides a method of dividing a single cohort into training and validation subsets for *P* value estimation when separate training and validation cohorts are not available. In addition, the software can perform standard Monte Carlo simulations (*e.g.,* cross-validation) to produce corrected *P* values to assess statistical significance of data assessed by multiple cut-points.

MATERIALS AND METHODS

Cohort Design and Tissue Microarray Construction. Tissue microarrays were constructed as described previously (5, 6). Paraffin-embedded formalin-fixed specimens of breast carcinoma were identified from the archives of the Yale University, Department of Pathology as available from 1962–1980. Two principal cohorts were used. The first consisted of 350 cases node-positive breast carcinomas, which was randomly divided into equally-sized training and validation sets. The second cohort consisted of 236 tumors (50% node-negative, 50% node-positive) similarly divided into training and validation halves. Complete treatment information was unavailable for either cohort; however, most of the node-positive patients were treated with local radiation, and approximately 15% were given chemotherapy consisting primarily of Adriamycin, Cytoxan, and 5-fluorouricil. The

Received 4/13/04; revised 7/13/04; accepted 7/28/04.

Grant support: R. Camp is supported by NIH Grant K0-8 ES11571 and by a grant from the Breast Cancer Alliance. M. Dolled-Filhart is supported by the United States Army Breast Cancer Research Grant DAMD17-03-1-0349. D. Rimm is supported by a grant from the Patrick and Catherine Weldon Donaghue Foundation for Medical Research, NIH Grant NCI R21 CA100825, and United States Army Grant DAMD-17-02-0463.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Requests for reprints: Robert L. Camp, Department of Pathology, Yale University School of Medicine. 310 Cedar St., New Haven, CT 06520. Phone: 203-785-6340; Fax: 203-785-7303; E-mail: robert.camp@yale.edu.

[3] A sample program is available for download at http://www.tissuearray.org/rimmlab/.

node-negative patients were routinely treated with surgery and/or local radiation alone. Approximately 27% of the patients subsequently received tamoxifen (post-1978). Representative regions of invasive carcinoma were selected for coring by a pathologist (R. L. C). These regions were taken from tumor areas that were not contaminated with normal stroma. Because prior studies have shown that a single core adequately represents the staining pattern of an entire slide, all studies were done with a single sample of each tumor (7, 8). All patients were followed until death or for a minimum of 30 years. Patients were deemed "uncensored" if they died of breast cancer within 30 years of their initial date of diagnosis. Patient information was collected under an approved protocol from the Yale Human Investigation Committee.

Immunohistochemistry and Analysis. Quantitative analysis of expression was done on the tissue microarrays with the automated quantitative analysis (AQUA) method of analysis described previously (9) Briefly, antibodies for human epidermal growth factor receptor-2 (HER2; 1:8,000, polyclonal), estrogen receptor (ER, 1:500, monoclonal), p53 (1:50, monoclonal), and cytokeratin (1:200, AE1/AE3 monoclonal and polyclonal) were obtained from DAKO Corp. (Carpinteria, CA). Primary antibodies were applied for 1 hour followed by an horseradish peroxidase-conjugated secondary (Envision, DAKO Corp.). Tumor cells were identified with an anticytokeratin antibody, followed by an Alexa488-conjugated antimouse or antirabbit secondary (1:200, Molecular Probes, Eugene OR). 4′,6-Diamidino-2-phenylindole was added to visualize nuclei. HER2, ER, and p53 were visualized with a fluorescent Chromagen (Cy-5-tyramide, NEN Life Science Products, Boston, MA). Cy-5 (red) was used because its emission peak is well outside the green-orange spectrum of tissue autofluorescence. Monochromatic, high-resolution (1024 × 1024 pixel, 0.5 μm resolution) images were obtained of each histospot. Areas of tumor were distinguished from stromal elements by creating a mask from the cytokeratin signal. Coalescence of cytokeratin at the cell surface helped localize the cell membranes, and 4′,6-diamidino-2-phenylindole was used to identify nuclei. HER2 signal from the membrane area of tumor cells was scored on a scale of 0 to 255 and expressed as signal intensity divided by the membrane area. ER and p53 signal from nuclei were similarly scored and expressed as signal intensity divided by nuclear area; thus, all scores are expressed on a scale with three significant figures. For arrays stained for visual (manual) analysis, 3,3′-diaminobenzidine was used in place of Cy5-tyramide.

Construction of X-Tile Plots. X-tile plots are created by dividing marker data into three populations: low, middle, and high (*i.e.,* two divisions). All possible divisions of the marker data are assessed. Associations can be calculated at each division by a variety of standard statistical tests, including the log-rank test for survival and means tests for associations between other marker data. The data are represented graphically in a right-triangular grid where each point (pixel) represents the data from a given set of divisions. The vertical axis represents all possible "high" populations, with the size of the high population increasing from top to bottom. Similarly, the horizontal axis represents all possible "low" populations, with the size of the low population increasing from left to right. Data along the hypotenuse represent results from a single cut-point that divides the data into high or low subsets. Data points away from the hypotenuse up or to the left represent results from two cut-points that define an additional "middle" population in addition to the high and low subsets. The size of the middle subset increases with greater distances from the hypotenuse. Specifically, a χ^2 value is calculated for every possible division of the population shown on the grid using a color code. Coloration of the plot represents the strength of the association at each division, ranging from low (dark, black) to high (bright, green, or red). Inverse associations between marker expression and survival (*e.g.,* high expression connotes poorer survival) are colored red, whereas direct associations are colored green. The cursor can be manually moved over any cut-point to reveal survival curves. Alternatively, the program can select the optimal division of the data by selecting the highest χ^2 value. Statistical significance is assessed by using the cut-point derived from a training set to parse a separate validation set, using a standard log-rank test, with *P* values obtained from a lookup table. The calculations done by X-tile have been validated with Statview 5.0.1 (SAS Institute, Cary NC). The concept is best understood by using the interactive version of the software available for download.[4]

Generation of "Training" and "Validation" Cohorts. X-tile creates separate training and validation cohorts by first making separate lists of "censored" and "uncensored" observations, ordered by follow-up time. Patients are alternately assigned to training and validation sets by reading down the list and selecting every other patient. This technique normalizes the base survival curve for both sets. It also ensures that the same training and validation sets are constructed each time an analysis of the same marker is done, thus preventing the possibility of obtaining different *P* values after reanalysis of the same data.

RESULTS AND DISCUSSION

To illustrate the utility of X-tile, we used a large, independent cohort, abstracted from the national SEER registry dataset (10). This dataset includes information on over 160,000 breast cancer patients with a median follow-up time of 30 months, including disease-related survival, patient age, tumor size, nodal status, and node number. We analyzed three well-characterized parameters: tumor size, patient age, and node number, in X-tile plots (Fig. 1). For tumor size, this analysis revealed a robust ($P < 0.0001$), linear association with patient survival, with no evidence of distinct subpopulations (Fig. 1*A*). Thus, no matter where a cutoff is made, larger tumor size is associated with poorer survival. To assess statistical significance and avoid the problems of multiple cut-point selection, we used the X-tile program to randomly divide the total cohort into two equal training and validation sets. The optimal cut-points were determined by locating the

[4] http://www.tissuearray.org/rimmlab/.

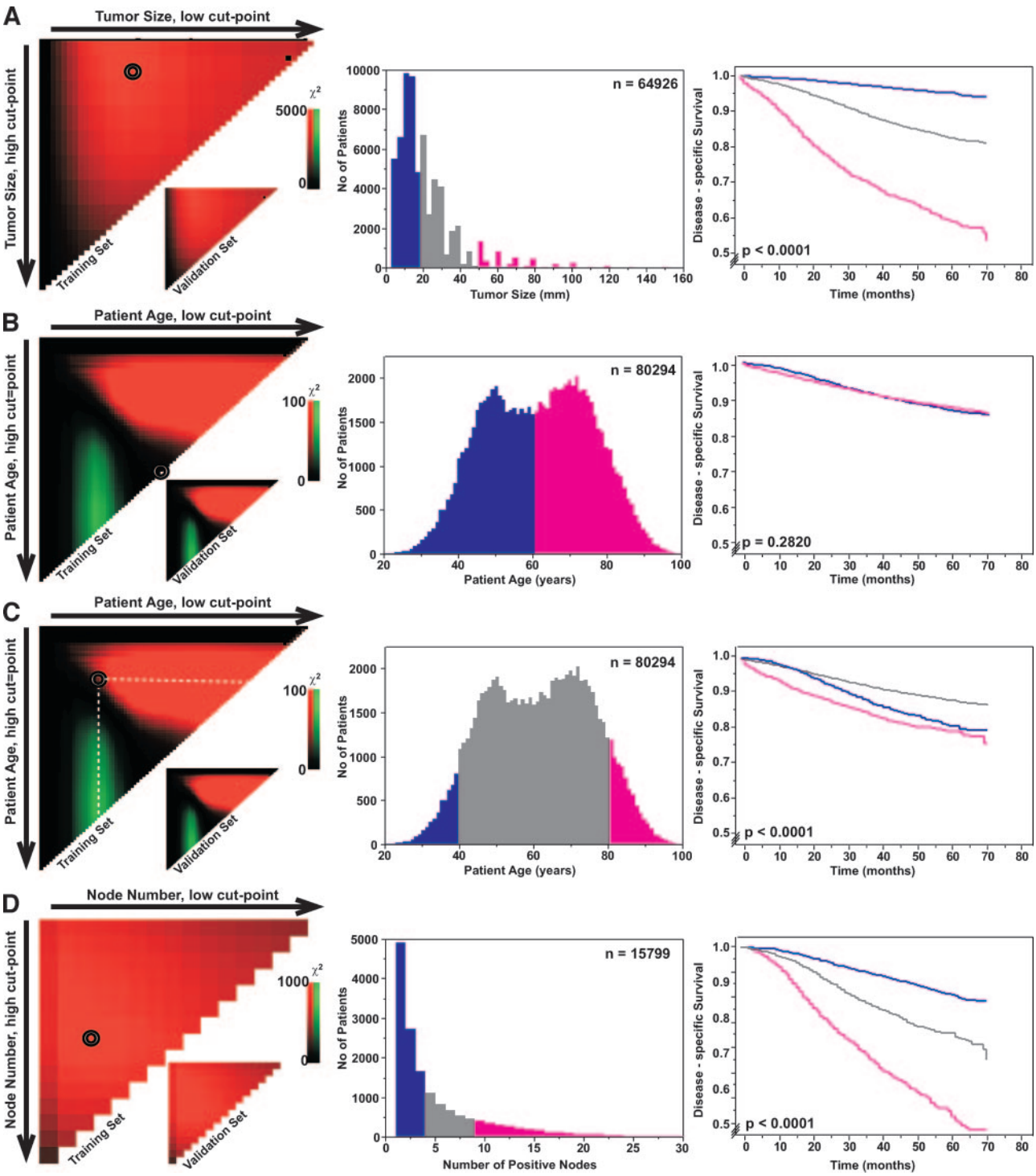


Fig. 1 X-tile analysis of survival data from the SEER registry reveals a continuous distribution based on tumor size and node number, and a U-shaped distribution based on age. X-tile analysis was done on patient data from the SEER registry, equally divided into training and validation sets. X-tile plots of training sets are shown in the *left panels*, with plots of matched validation sets shown in the *smaller inset*. The construction of X-tile plots is more fully described in Materials and Methods. The plot shows the χ^2 log-rank values produced when dividing the cohort with two cut-points, producing high, middle, and low subsets. The *X*-axis represents all potential cut-points from low to high (*left to right*) that define a low subset, whereas the *Y*-axis represents cut-points from high to low (*top to bottom*), that define a high subset. The *arrows* represent the direction in which the low subset (*X*-axis) and the high subset (*Y*-axis) increase in size. *Red coloration* of cut-points indicates an inverse correlation with survival, whereas *green coloration* represents direct associations. The optimal cut-point occurs at the *brightest pixel* (*green* or *red*). The cut-point highlighted by the *black/white circle* in the *left panels* is shown on a histogram of the entire cohort (*middle panels*), and a Kaplan-Meier plot (*right panels*; low subset = blue, middle subset = gray, high subset = magenta). *P* values were determined by using the cut-point defined in the training set and applying it to

brightest pixel on the X-tile plot of the training set. Statistical significance was determined by applying this cut-point set to the validation set. Analysis of the survival data with X-tile reveals optimal cut-points of (<1.9, 1.9 to 4.9, and >4.9 cm), which are almost identical to those empirically established for the staging of breast cancer (≤2, 2 to 5, and >5 cm; ref. 11). Histogram analysis of patient age shows a bimodal distribution, with two incidence peaks above and below 59 years of age (Fig. 1*B*). Using the histogram as a guide, we divided the cohort into "old" and "young" subsets based on a cut-point of 59 years; however, this division fails to result in a statistically significant difference in survival (Fig. 1*B*). In contrast, X-tile-based optimal cut-point selection analysis of patient age reveals two distinct populations, (<40 years) and (>82 years), which exhibit poor disease-specific outcome relative to patients in the middle (40–82 years; Fig. 1*C*). This relationship has been described previously (12) and is an example of a bimodal population that is robust but is neither readily demonstrable on a histogram nor discoverable with a single cut-point analysis. It also provides an example of how X-tile can be used to discover optimal population cut-points. Finally, we use X-tile to define the optimal cut-point for the number of tumor-involved lymph nodes in patients with node-positive breast carcinoma (Fig. 1*D*). As with tumor size, X-tile shows a diffuse continuous distribution with no discernable subpopulations. X-tile identifies the optimal division of the cohort into three populations (<4, 4 to 8, and ≥9). Again these divisions are remarkably similar to those established for breast cancer staging (<4, 4 to 9, and ≥10; ref. 11).

We then did X-tile analysis on tumor biomarkers using tissue microarray cohorts available in our laboratory. HER2, a member of the epidermal growth factor family is frequently overexpressed in breast cancers and is associated with poor prognosis. X-tile analysis of the productive ability of HER2 reveals a distinct subset of HER2 high tumors with poor outcome (Fig. 2*A-C*). In contrast, a similar plot of ER shows a linear, diffuse distribution of risk across the entire plot, with no evident subpopulations (Fig. 2*D*). At virtually every division plotted, patients with ER-high tumors survive longer than patients with ER-negative tumors.

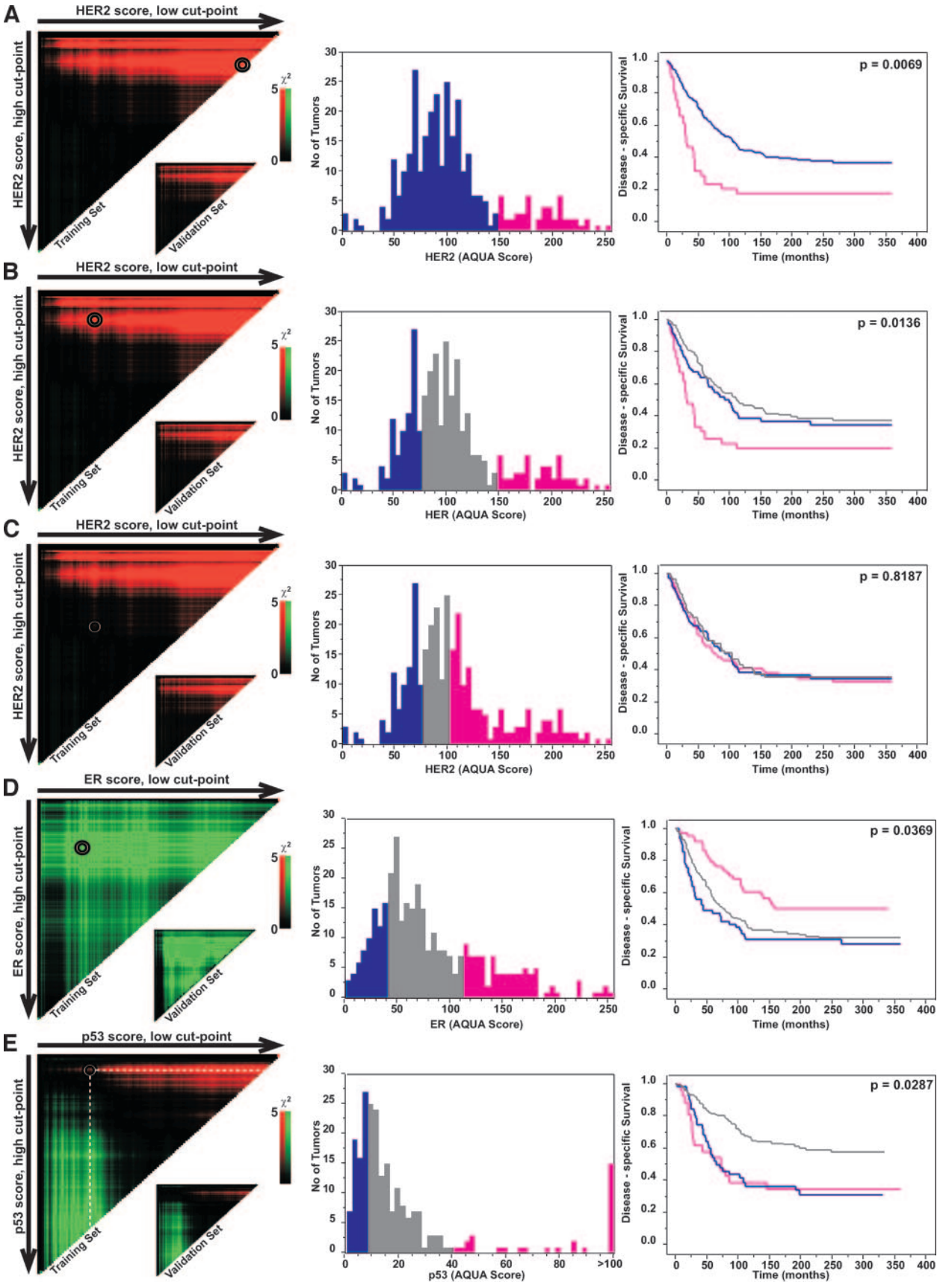
These X-tile plots represent two distinct patterns, which provide insight into the underlying biology of HER2 and ER. In the case of HER2, gene amplification is common and often results in marked overexpression of HER2 at the protein level (13). This biology is evident in the HER2 X-tile plot, which shows localized risk in a distinct subset of patients with HER2-high tumors (Fig. 2*A-C*). The discreteness of this subset suggests that a 2-way cut (high *versus* low expression) of the data is more appropriate than a 3-way split. Consequently, we determined the best single cut-point along the hypotenuse of the X-tile plot, which occurs as a HER2 AQUA score of >156.8 (test set/validation set $P = 0.0069$, Fig. 2*A*).

In contrast, the expression of ER is highly complex and governed by a number of regulatory elements (14). There is no evidence of ER-genetic mutations that would subset tumors based on either gene amplification or deletion. This biology fits with ER's observed X-tile plot showing a diffuse distribution of risk (Fig. 2*D*). Analysis of this cohort with a continuous univariate analysis of ER expression reveals a significant correlation with survival ($P = 0.0017$), demonstrating that ER, in this node-positive cohort, is positively associated with patient outcome, when assessed as a continuous variable. Despite the continuous nature of the X-tile plot, we sought to determine the optimal division as defined by X-tile, which was found at AQUA scores of 37.7 and 117.6 (test set/validation set $P = 0.0369$). Interestingly, there is historical disagreement about how to divide tumors into ER-positive and ER-negative subsets. The continuous nature of ER would suggest that although one can determine a statistically valid cut-point for ER, the marker should really be thought of as a truly continuous (linear) marker when assessing survival.

Finally, we use an X-tile plot of p53 to elucidate tumor novel tumor subpopulations. p53 is a tumor suppressor with complex biology and a varied record of predictive and prognostic significance in breast cancer. Mutation of both p53 alleles frequently leads to overexpression of nonfunctional forms of p53 in an attempt to compensate for the loss of p53 function. Such overexpression is often attributable to missense mutations in exons 5–8 and is associated with poor prognosis (15, 16). When assessed with standard immunohistochemistry, these tumors frequently exhibit intense p53 nuclear staining. In contrast, mutations outside of these exons generally result in undetectable levels of p53 protein (p53 null mutations; ref. 17, 18). These tumors are therefore not detected by standard immunohistochemistry (19, 20). Theoretically, one should be able to distinguish among p53-negative (null mutant), p53-low (non-mutant), and p53-overexpressing (non-null mutant) tumors; however, such distinctions are beyond the sensitivity of traditional manually read "brown-stain" immunohistochemistry.

We analyzed p53 expression on a training cohort of 118 breast cancers, 50% of which were node-positive using AQUA. Correlations between p53 expression and patient survival were graphed on an X-tile plot, which clearly shows two distinct subpopulations of tumors, one of high expressers and one of low expressers, both of which do poorly relative

the validation set. *A* shows tumor size divided at the optimal cut-point, as defined by the most significant (*brightest pixel*) on the plot (19 and 49 mm, $P < 0.0001$). *Diffuse red* indicates a continuous indirect association between increasing tumor size and good prognosis. *B* shows patient age divided at the mean age (59 years, $P = 0.2820$). Note that this point appears as a *black/white circle* on the hypotenuse of the X-tile plot. The lack of a significant survival difference is reflected in the fact that this region of the X-tile plot is *black*. *C* shows patient age divided at the optimal cut-point (at the intersection of the most significant division for the high and low subsets, 48 and 70 years, $P < 0.0001$). The plot reveals a U-shaped distribution, with both a high (*red cloud*) and a low (*green cloud*) population that exhibit worse outcome. *D* shows the number of positive nodes in patients with node-positive breast carcinoma. Analysis was limited to cases in which ≥8 lymph nodes were pathologically analyzed. The optimal cut-point is shown (four and nine positive nodes, $P < 0.0001$).



to tumors expressing intermediate levels of p53 (Fig. 2*E*). The X-tile plot reveals significant, optimal cut-points of 7.5 and 46.4 ($P = 0.0287$). We speculate that the p53-negative tumors are true double-negative mutants for p53 and therefore comprise a poor-prognosis subset. In contrast, the p53-high tumors most likely represent non-null mutants, which also have a poor outcome. In this case, X-tile has elucidated an intriguing subpopulation of tumors that calls for further biochemical investigation.

Although X-tile plots were designed for use with continuous data, they can also be used to find appropriate cut-points in categorical (nominal) data such as that obtained by manual scoring of brown-stained immunohistochemical slides. To show this, we manually scored our HER2 and ER-stained tissue microarrays on a scale of 0 to 3+ and expressed this data in X-tile plots (Fig. 3*A* and *B*). Note that each possible cut-point is represented by a colored square. As with continuous data, the optimal cut-point is determined by finding the square of the highest color intensity on a training set and then using that cut-point on a separate validation set to test for statistical significance.

In addition to visualizing survival data, X-tile plots can be used display other correlative data, such as the association between two markers. For instance, X-tile can divide a cohort into subsets based on the expression of one marker and analyze the variance of the mean expression of another marker in each of these subsets. Figure 3*C* and *D* show the association between HER2 and ER expression in our cohort of breast cancer patients. In 3*C*, we used HER2 to divide the cohort and compared the mean ER expression of each subset using an ANOVA. As with the survival analysis shown in Fig. 2, the association between HER2 and ER expression shows a subset of high HER2-expressing tumors that have low mean ER expression (Fig. 3*C*). Interestingly, HER2 *versus* ER X-tile plot of training set data defines the optimal cut-point between high and low HER2 expressers at an AQUA score of 156.8, identical to that obtained when survival analysis is used (Fig. 2). This cut-point is statistically significant ($P = 0.0056$) when assessed on a validation set. In contrast, using ER to divide the population and analyzing HER2 mean expression provides a more diffuse X-tile plot (panel D) where increasingly lower ER expression correlates with increasingly higher HER2 expression. The optimal cut-point for this X-tile plot occurs at 45.7 ($P = 0.0005$). These examples illustrate that as with survival X-tile plots, correlative marker plots can also find biologically meaningful tumor subsets.

Recently, numerous algorithms have been developed in efforts to analyze results from DNA based microarrays. These include algorithms based on clustering (21), genetic permutation (22), or χ^2 Automatic Interaction Detector (23). Although these systems can analyze tumor subsets post hoc to assess differences in survival, none use survival time and censor information for the initial classification of tumor subsets. Furthermore, these algorithms generally treat marker data as continuous or nominalize data into nonbiological categories (*e.g.,* quartiles or manual scoring on a 0–3+ scale). Our study would suggest that whereas this approach might be appropriate for ER, tumor size, or node number, it clearly is not appropriate for HER2, p53, or patient age, which exhibit nonlinear expression. Indeed, the actual HER2 expression level on a linear scale is far less important than the fact that a tumor has a HER2 expression level above a certain cut-point. Although the calculations presented in an X-tile plot are not novel in and of themselves, performing a similar analysis with conventional techniques would require the performance of thousands of individual survival analyses and the exhaustive comparison of each to determine where an optimal cut-point might occur.

In summary, X-tile plots present a new tool for (*a*) the assessment of biological relationships between a biomarker and outcome; and (*b*) the discovery of population cut-points based on marker expression. Clinically, the ability to effectively define such subpopulations is important for the development of novel therapeutics that may be effective for only a limited subset of tumors. Furthermore, this method may have value in other fields (materials testing, actuarial questions, etc.) where cut-point selection is complicated by time-dependent assessment of outcome.

Fig. 2 X-tile plots of prognostic markers HER2, ER, and p53 on breast cancer tissue microarray cohorts. *A-C* show the X-tile analysis of HER2 staining on a cohort of 142 node-positive breast cancers. The X-tile plot of the 30-year disease-related survival of the training set is represented, with the X-tile plot of a separate validation cohort of 143 tumors in the *inset* (*A-C*, *left*). HER2 staining was quantitatively determined with an automated analysis system (AQUA; ref. 9). Note the cloud of high χ^2 values for a subpopulation of high expressers (~15% of the total population). As long as the size of the high subpopulation remains constant, the survival difference remains high, despite changes in the size of the lower population (*A* and *B*, *left*, *middle*, *right*). In contrast, expanding the size of the high subpopulation beyond ~18% reduces survival difference (*C*, *left*, *middle*, *right*), as more and more low expressing tumors dilute the higher subpopulation. Using this X-tile plot, the maximum χ^2 value for a single cut-point occurs at an AQUA score of 156.8 (along the hypotenuse of the plot, *A*). The *red color* of the cloud illustrates the fact that higher HER2 expressing tumors do worse. This optimal cut-point was validated by dividing a separate cohort and was found to be statistically significant ($P = 0.0136$, *A*, *right*). *D* shows the X-tile plot of ER. One hundred and eighteen node-positive tumors were assessed for ER status, and results were displayed on an X-tile plot. ER staining was quantitatively determined with AQUA. The X-tile plot of ER shows diffuse correlation between ER expression and 30-year disease-related survival (*D*, *left*). The *green color* of the plot illustrates the fact the higher ER-expressing tumors do better. X-tile discovers optimal cut-points at 37.7 and 117.6 on a training set (*outlined circle*, *D*, *left*). These cut-point were applied to a validation set ($N = 118$, inset) and found to be statistically significant ($P = 0.0369$, *D*, *right*). *E* shows the X-tile plot of p53 expression done on a cohort of 111 breast carcinomas (50% node-negative) with 30-year disease-related follow-up data. The X-tile plot reveals a U-shaped distribution, with both a high (*red cloud*) and a low (*green cloud*) population that exhibit worse outcome (*E*, *left*). The optimal population division occurs at AQUA scores of 7.5 and 46.4. When these cut-points are applied to a separate cohort of 111 breast cancer specimens (*inset*), they are found to be statistically significant ($P = 0.0287$, *E*, *right*). The p53 histogram is truncated at an AQUA score of 100 to visualize the low/moderate expressers. Tumors with p53 scores of >100 are shown in a *single bar* (*E*, *middle*).

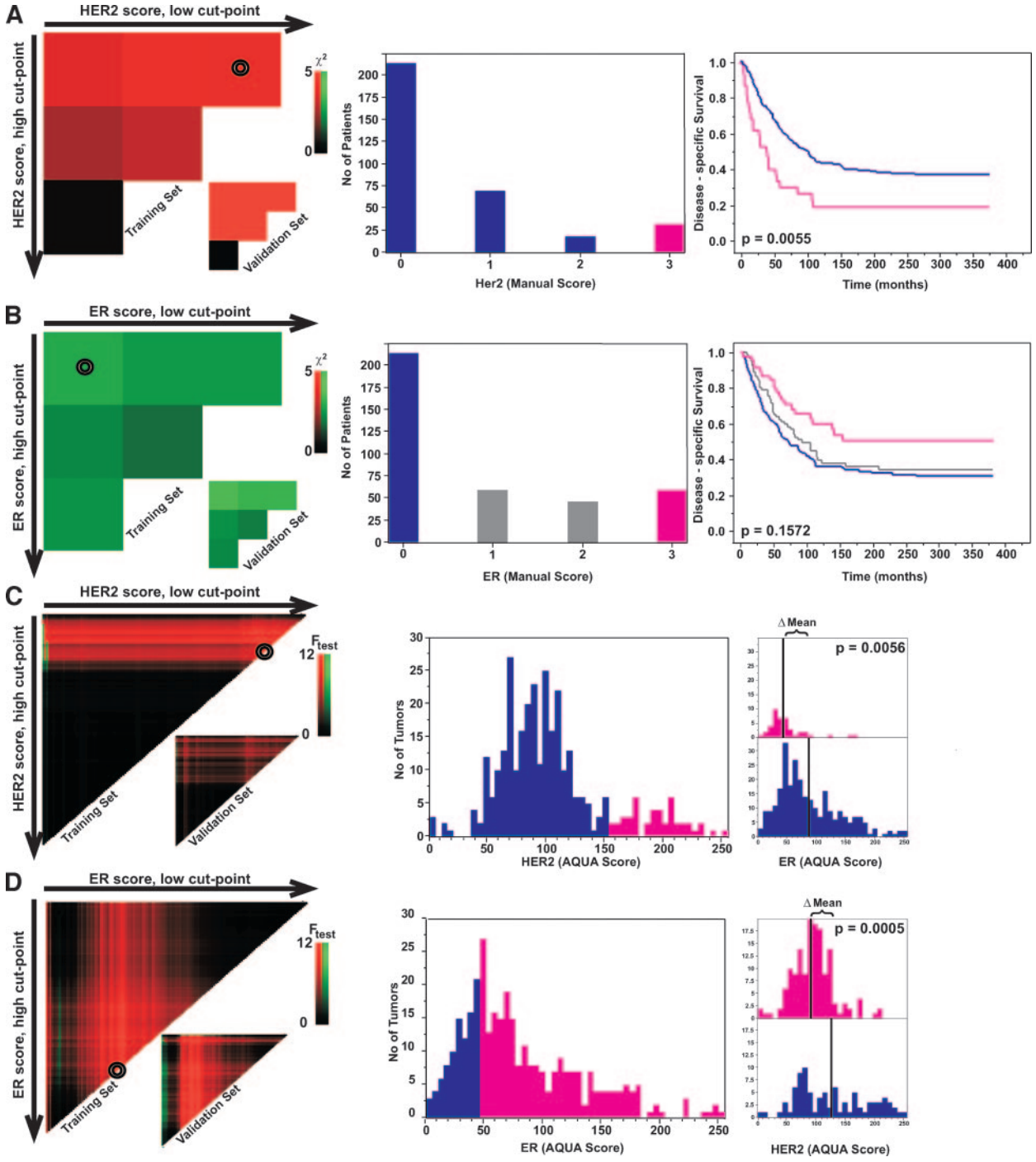


Fig. 3 X-tile can be used with categorical (nominal) data and can subset populations based on correlative associations between markers. The tissue microarrays described in Fig. 1 were stained for HER2 and ER and scored manually on a scale of 0 to 3+ (*A* and *B*, respectively). *A* shows an X-tile plot of the training set data are shown on the *left*, with the validation set shown in the *inset*. The optimal 2-population cut-point for HER2 was determined to be between a score of 2+/3+. This cut-point showed statistical significance when assessed on the validation cohort ($P = 0.0055$). *B* shows the X-tile plot of manual ER scores. The optimal 3-population cut-points were also assessed and found to be between 0/1+ and 2+/3+. These cut-points, however, were not statistically significant when assessed on a separate validation cohort ($P = 0.1572$). *C* shows an X-tile plot of ANOVA F-test values for the mean ER expression of subsets defined by their HER2 expression. The *red color* shows an inverse correlation between ER and HER2. The plot shows a subset of HER2-high tumors that are ER-low. The *middle panel* shows the optimal 2-population cut-point for this subset (at 156.8). The *right panel* shows histograms of the ER expression of the HER2-high (*top*) and HER2-low (*bottom*) subsets. The mean ER AQUA score of each subset (86.9 and 48.1) is indicated by a *vertical black line*. The *P* value (0.0056) for this mean difference was determined by subsetting a separate validation cohort. *D* shows an X-tile plot of ANOVA F-test values for the mean HER2 expression of subsets defined by ER expression. The *increasing gradient of color from right to left* indicates that lower and lower levels of ER are associated with high and higher HER2. The optimal 2-population cut-point for this plot occurs at 45.7, producing a *P* value of 0.0005 using the validation set.

ACKNOWLEDGEMENTS

The authors thank Jay Cowan, Bonnie King, and Daniel Zelterman for help with this report.

REFERENCES

1. Mazumdar M, Glassman JR. Categorizing a prognostic variable: review of methods, code for easy implementation and applications to decision-making about cancer treatments. Stat Med 2000;19:113–32.

2. Mazumdar M, Smith A, Bacik J. Methods for categorizing a prognostic variable in a multivariable setting. Stat Med 2003;22:559–71.

3. Hilsenbeck SG, Clark GM, McGuire WL. Why do so many prognostic factors fail to pan out? Breast Cancer Res Treat 1992;22: 197–206.

4. Altman DG, Lausen B, Sauerbrei W, Schumacher M. Dangers of using "optimal" cutpoints in the evaluation of prognostic factors. J Natl Cancer Inst (Bethesda) 1994;86:829–35.

5. Kononen J, Bubendorf L, Kallioniemi A, et al. Tissue microarrays for high-throughput molecular profiling of tumor specimens [see comments]. Nat Med 1998;4:844–7.

6. Rimm DL, Camp RL, Charette LA, Olsen DA, Provost E. Amplification of tissue by construction of tissue microarrays. Exp Mol Pathol 2001;70:255–64.

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

8. Torhorst J, Bucher C, Kononen J, et al. Tissue microarrays for rapid linking of molecular changes to clinical endpoints. Am J Pathol 2001; 159:2249–56.

9. Camp RL, Chung GG, Rimm DL. Automated subcellular localization and quantification of protein expression in tissue microarrays. Nat Med 2002;8:1323–7.

10. Surveillance, Epidemiology, and End Results (SEER) Program Public-Use. In, YR1992_1997.SEER11 ed: National Cancer Institute, DCCPS, Cancer Surveillance Research Program, Cancer Statistics Branch (1973–1997); 2000.

11. Greene FL, Page DL, Fleming ID, et al., editors. AJCC cancer staging manual, 6th edition. New York: Springer; 2002.

12. Chung M, Chang HR, Bland KI, Wanebo HJ. Younger women with breast carcinoma have a poorer prognosis than older women. Cancer (Phila) 1996;77:97–103.

13. Menard S, Fortis S, Castiglioni F, Agresti R, Balsari A. HER2 as a prognostic factor in breast cancer. Oncology 2001;61(Suppl 2):67–72.

14. McDonnell DP, Norris JD. Connections and regulation of the human estrogen receptor. Science (Wash D C) 2002;296:1642–4.

15. Hartmann A, Blaszyk H, McGovern RM, et al. p53 gene mutations inside and outside of exons 5–8: the patterns differ in breast and other cancers. Oncogene 1995;10:681–8.

16. van Slooten HJ, van De Vijver MJ, Borresen AL, et al. Mutations in exons 5–8 of the p53 gene, independent of their type and location, are associated with increased apoptosis and mitosis in invasive breast carcinoma. J Pathol 1999;189:504–13.

17. Hashimoto T, Tokuchi Y, Hayashi M, et al. p53 null mutations undetected by immunohistochemical staining predict a poor outcome with early-stage non-small cell lung carcinomas. Cancer Res 1999;59: 5572–7.

18. Bodner SM, Minna JD, Jensen SM, et al. Expression of mutant p53 proteins in lung cancer correlates with the class of p53 gene mutation. Oncogene 1992;7:743–9.

19. Watanabe J, Nishiyama H, Okubo K, et al. Clinical evaluation of p53 mutations in urothelial carcinoma by IHC and FASAY. Urology 2004;63:989–93.

20. Skilling JS, Sood A, Niemann T, Lager DJ, Buller RE. An abundance of p53 null mutations in ovarian carcinoma. Oncogene 1996;13: 117–23.

21. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA 1998;95:14863–8.

22. Holland JH. Building blocks, cohort genetic algorithms, and hyperplane-defined functions. Evol Comput 2000;8:373–91.

23. Kass GV. An exploratory technique for investigating large quantities of categorical data. Appl Stats 1980;29:119–27.