

# Prognosis for Survival of Young Women with Breast Cancer by Quantitative p53 Immunohistochemistry

David E. Axelrod<sup>1</sup>, Kinsuk Shah<sup>2</sup>, Qifeng Yang<sup>3</sup> & Bruce G. Haffty<sup>3</sup>

<sup>1</sup> Department of Genetics and Cancer Institute of New Jersey, Rutgers University, New Jersey, USA

<sup>2</sup> Department of Genetics, Rutgers University, USA

<sup>3</sup> Department of Radiation Oncology, Robert Wood Johnson Medical School-UMDNJ, Cancer Institute of New Jersey, New Jersey, USA

Correspondence: David E. Axelrod, Department of Genetics and Cancer Institute of New Jersey, Rutgers University, 604 Allison Road, Piscataway, NJ 08854-8082, USA. Tel: 1-732-445-2011. E-mail: axelrod@biology.rutgers.edu

Received: March 27, 2012 Accepted: April 10, 2012 Online Published: May 1, 2012

doi:10.5539/cco.v1n1p52

URL: <http://dx.doi.org/10.5539/cco.v1n1p52>

*DEA and KS were supported by National Institutes of Health (NIH) grant U56 CA113004. BH was supported by the Breast Cancer Research Foundation. QY was supported by the National Natural Science Foundation of China (No. 81172529)*

## Abstract

p53 protein detected immunohistochemically has not been accepted as a biomarker for breast cancer patients because of disparate reports of the relationship between the amount of p53 protein detected and patient survival. The purpose of this study was to determine experimental conditions and methods of data analysis for which p53 stain intensity could be prognostic for survival of young breast cancer patients. A tissue microarray of specimens from 93 patients was stained with anti-p53 antibody, and stain intensity measured with a computer-aided image analysis system. A cut-point at one standard deviation below the mean of the distribution of p53 stain intensity separated patients into two groups with significantly different survival. These results were confirmed by Quantitative Nuclear Grade determined by DNA-specific Feulgen staining. P53 provided information beyond ER and PR status. Therefore, under the conditions reported here, p53 protein can be an effective prognostic factor for young breast cancer patients.

**Keywords:** breast cancer, p53, survival, tissue microarray, prognosis, biomarker, nuclear grade, image analysis

## 1. Introduction

Mutations in the p53 gene are common in many tumors, including breast tumors (Osborne et al., 1991). Many of these mutations result in an accumulation of the mutant form of the p53 protein (Hainaut & Hollstein, 2000). Immunohistochemistry, which detects mostly the mutant form of the p53 protein because of its greater stability, might be expected to provide useful prognostic information. However, there have disparate reports of the prognostic value of p53 detected by immunohistochemistry.

There have been reports of a positive prognostic value of p53 detected by immunohistochemistry. For instance, expression of mutant p53 protein was associated with early disease recurrence and early death in node negative breast cancer (Allred et al., 1993). A cocktail of two different antibodies was associated significantly with overall survival, although five individual antibodies were not (Elledge et al., 1994). The accumulation of p53 was significantly associated with increased risk of progression to breast cancer in women with benign breast disease (Rohan et al., 2006). In patients with ductal carcinoma *in situ* who had not received radiation therapy, the recurrence-free interval was significantly shorter for patients accumulating immunohistochemically detected p53 (Ringberg et al., 2001). Accumulation of p53 was significantly associated with increased local relapse of breast cancer following mastectomy with, or without, radiation (Zellers et al., 2000). In breast carcinoma patients having lumpectomy and radiation therapy, overexpression of mutant p53 protein was reported to be significantly prognostic for ipsilateral breast tumor recurrence (Turner et al., 2000). The accumulation of p53 of specimens on tissue microarrays was significantly associated with tumor specific survival of patients treated with hormone

and/or cytotoxic adjuvant therapy (Torhorst et al., 2001). A positive correlation of stain intensity with survival was reported when low concentrations of antibody were used (McCabe et al., 2005; Henson, 2005). A non-linear relationship between stain intensity and patient survival has been reported; medium stain intensity was associated with better survival than high or low stain intensity (Camp et al., 2004). Only the latter two reports used image analysis to quantify p53 stain intensity, the others used visual judgment to score p53 staining.

In contrast to reports of a positive prognostic value to immunohistologically detected p53, there have been reports of a lack of positive prognostic value of immunologically detected p53. P53 stain intensity, judged visually, was not correlated with overall survival, nor was there a concordance between p53 protein detected by immunochemistry and p53 gene mutations detected by DNA sequencing (Sjögren et al., 1996). P53 was found not to be a significant risk factor for local recurrence after breast-conserving therapy and radiation therapy (Elkuizen et al., 1999). The percent of tumor cells stained with antibody to p53 was not significantly associated with recurrence free or overall survival for patients who received mastectomy and chemotherapy but not radiotherapy (Reed et al., 2000). The proportion of tissues from invasive breast cancer that stained positively for p53 was only 18.7%, although the intensity of stain, judged visually, was associated with tumor histological grade (Bartley and Ross, 2002). The time to chest wall recurrence after mastectomy and radiotherapy was not significantly associated with p53 expression, although, overall and cause-specific survival after chest wall recurrence was significant (Haffty et al., 2004). In patients with early-onset breast cancer treated with surgery and radiotherapy, no significant correlation was found, between the expression of six immunochemical biomarkers, including p53, and local relapse (Choi et al, 2005).

The difference conclusions about the efficacy of using p53 immunohistochemistry for breast cancer prognosis have resulted in p53 not being recommended for use as a biomarker in breast cancer (Harris et al., 2007). These different conclusions may be related to differences between the studies in one or more factors. These factors include differences in clinical features of the patient cohorts (young or old age at diagnosis), treatment modalities (with or without adjuvant therapy), immunohistochemical methodology (different antibodies, high or low antibody concentration), methods of evaluating stain intensity (visual judgment or computer-aided measurement), and statistical analysis of measurements (one or two cut points).

In this study we have asked if p53 immunohistochemical stain intensity could be prognostic for a cohort of young breast cancer patients who have been treated with radiation and adjuvant therapy. Stain intensity was determined quantitatively by image analysis of specimens on a tissue microarray (TMA). We found that p53 protein stain intensity measured by quantitative image analysis can distinguish patients with good and poor survival, and provide information beyond that provided by the patient's estrogen receptor status and progesterone receptor status. The stratification of patients by p53 was confirmed by determining the Quantitative Nuclear Grade of their specimens, an independent technique that characterizes nuclear DNA rather than protein.

## **2. Materials and Methods**

### *2.1 Tissue Microarray*

The tissue microarray (YTMA71) was produced at the Yale Microarray Facility from formalin fixed biopsy specimens. Each disk was 0.6 mm in diameter. Included on the array were duplicate disks from 171 patients, 18 disks from normal breast ducts, and 1 disk from liver to determine orientation. Two similar microarray slides were produced from serial slices of same biopsy cores, one slide was stained with an antibody to p53 protein at the Cancer Institute of New Jersey, and the other slide was stained with the DNA specific Feulgen reaction for nuclear grade at Rutgers University. Not all disks could be used. Disks that contained no tissue or insufficient tissue were not included in the analysis; disks whose Feulgen stain color was inappropriate (light green not blue) were not included in the analysis; and disks with an inadequate number of segmented nuclear images were not included. For p53 protein, 186 disks from 93 treated patients were informative, and of these, 153 disks from 82 patients (some with only single useable disks) were also informative for Quantitative Nuclear Grade.

### *2.2 Immunohistochemistry*

One tissue microarray slide was stained with primary mouse anti-human monoclonal antibody clone DO-7, DAKO code M 7001 (Carpinteria, CA) at 1:50. This antibody labels wild-type and mutant-type p53 protein, but because of the relatively shorter half-life of the wild-type p53 protein most of the stain intensity would be expected to represent mutant-type p53. The primary antibody incubation was at room temperature for 30 minutes. Antigen retrieval was in citrate buffer for 20 minutes. The DAKO LSAB<sup>TM</sup>+ HRP kit was used for visualization.

### 2.3 Feulgen Staining

A second tissue microarray slide made from a serial cut of the same biopsy cores as for immunohistochemistry was stained with the Feulgen DNA Staining Reagent kit, Bacus Laboratories (Lombard, IL) Cat. No. SR200. The DNA specific procedure included acid hydrolysis in 5N HCl for one hour at room temperature, and reaction with Schiff reagent to develop color.

### 2.4 Image Analysis

The BLISS Imaging System (Bacus Laboratories, Lombard, IL) was used to acquire images and extract data. For immunohistochemistry the following features were measured for each disk using the TMAScore v1.0 Program: disk area, stained area, and average optical density of the stained area. The resulting average immunohistochemistry optical density (Ave IHC OD) for each disk provides a quantitative measure which takes into account both the proportion of stained nuclei and the intensity of staining of each nucleus, as previously judged visually (Allred et al., 1993). Thresholds for protein stain and background hematoxylin counterstain were selected for maximum distinction using disks with normal breast duct tissue as negative controls. Relative staining intensity determined by image analysis was confirmed as reasonable by visual inspection. For Quantitative Nuclear Grade the Cell Finder Program v0.50 was used to automatically segment images of Feulgen-stained nuclei using selected criteria of size, shape, sum optical density, and nuclear density. A gallery of automatically segmented nuclear images was visually inspected and inappropriately segmented nuclei (mostly doubles) were manually deleted. The average number of satisfactory nuclei per disk was 159. The Cell View Program v0.90 was used to extract values of 39 nuclear features in three classes: nuclear size and shape, optical density (amount of DNA, ploidy), and Markovian chromatin texture (arrangement of DNA in the nucleus) (Haralick et al., 1976; Bacus et al., 1999). The Nuclear Grade Program v0.40 was used to produce a grade for each nucleus and a mean grade for all nuclei in a disk, the latter reported as the Quantitative Nuclear Grade (Bacus et al., 1999).

### 2.5 Statistical Analysis

Data extracted by the BLISS Imaging System was saved as a Microsoft Excel v11.2.5 spreadsheet and combined with clinical data provided as a separate Excel file. Excel spreadsheets were used to reorganize data and generate subsets of data for statistical analysis. JMP v7 (SAS Institute, Cary, NC) was used to generate descriptive statistics, and to perform Kaplan-Meier survival analysis, multivariate Cox proportional hazards analysis, and statistical tests. Matlab v14.3 (Math Works, Lowell, MA) was used generate distributions of categories of patients.

## 3. Results

### 3.1 Selection of Patients

One hundred and seventy-one patients were selected who had been diagnosed with breast cancer at a young age, median of 40 years. The mean time to events was 7.7 years, standard deviation of 5.9 years. BRCA1, 2 mutation status was not informative since among the 71 patients whose status was determined only five had events. All patients were treated with radiation therapy. Some patients received adjuvant therapy and some did not. Among those patients who received adjuvant therapy we wished to determine if p53 immunohistochemistry (IHC) could distinguish those that had good survival and those that had poor survival. Of the total of 171 patients with tissue on the microarray, 93 patients who had been treated with adjuvant hormone therapy, chemotherapy, or both, and whose pairs of disks had sufficient tissue were analyzed for p53 stain intensity. Their characteristics are given in Table 1. Approval was obtained from the Institutional Review Boards.

Table 1. Patient and tumor characteristics

Age	Mean	39.0
	Median	40
	Range	25-49
T stage	T1	54
	T2	34
	T3	3
	T4	1
	Unknown	1
	Total	93
N stage	N0	56
	N1	29
	N2	2
	Unknown	6
	Total	93
Tumor size (cm)	Mean	2.1
	Median	2.0
	Range	0.5-5.0
Adjuvant therapy	No chemotherapy	9
	Chemotherapy	84
	No hormone	59
	Hormone therapy	34
	Chemotherapy and/or Hormone therapy	93

### 3.2 Evaluation of Tissue Microarray

The tissue microarray included two disk positions for each of 171 patients. However, not all disk positions contained sufficient patient tissue to be informative. Informative core positions were determined by inspecting a log-log plot of the stained area of the cores vs. the total tissue area of the cores. Disks were considered informative if they had total tissue area greater than 11,000 square microns, and a stained area greater than 800 sq. microns. Disks containing patient tissue were considered to stain positive if their average IHC optical density was greater than the average IHC optical density of disks containing normal duct tissue, Figure 1. Of the 171 patients, 93 patients had pairs of informative disks. There were a total of 186 informative disks. Measurements of p53 stain intensity of pairs of disks were similar, Spearman rank correlation coefficient = 0.427,  $p < 0.0001$ . Previous reports have concluded that two cores were sufficient to characterize each patient (Camp et al., 2000; Ganchberg et al., 2002; Gillett et al., 2000; Torhorst et al., 2001).

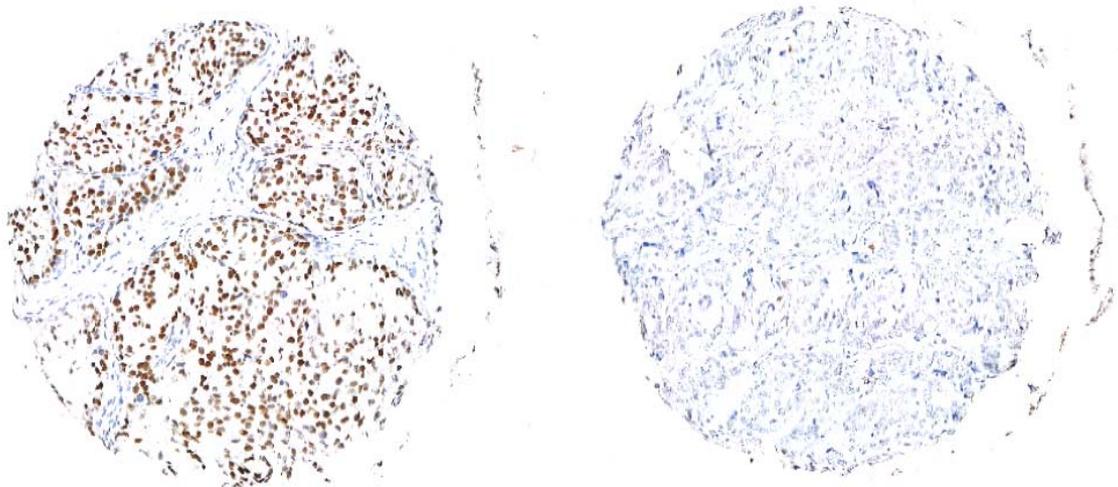


Figure 1. Adjacent disks from the tissue microarray stained with antibody to the p53 protein and counterstained with hematoxylin. Left: patient tissue with dark nuclear p53 stain, Average IHC OD = 43. Right: patient tissue showing counter stain, but p53 stain intensity similar to normal tissue

### 3.3 Distribution of p53 staining

The distribution of the p53 staining intensities (average IHC optical density for each disk) for all 186 disks is shown in Figure 2. The mean stain intensity for all disks is 42.8 and the standard deviation is 7.2. The median stain intensity is 40.0. There is a broad distribution of stain intensities, coefficient of variation = 17%. Next we wished to determine an effective cut point, or cut points, of the distribution that would distinguish patients with good survival from patients with poor survival.

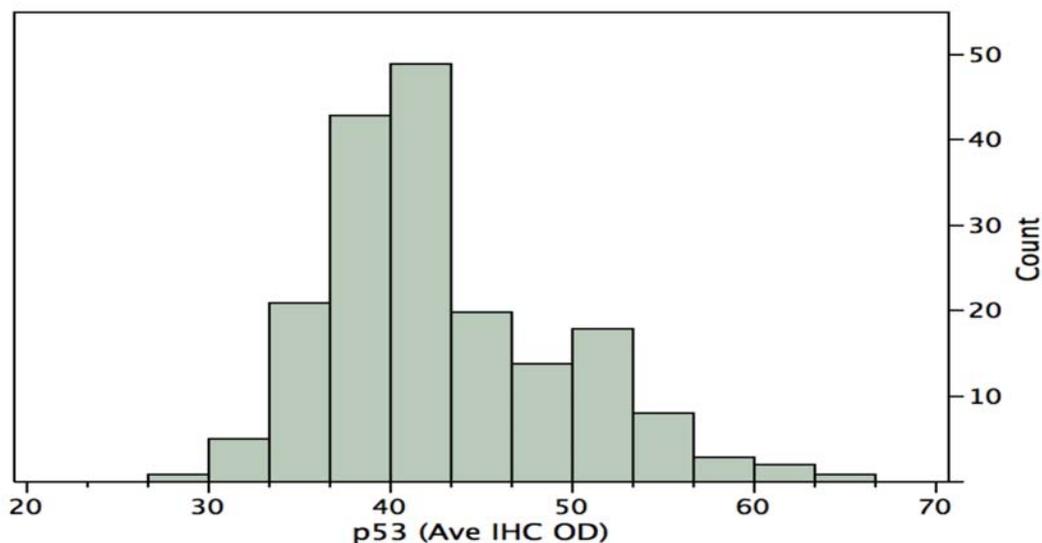


Figure 2. Distribution of p53 stain intensity. Average immunohistochemistry optical density per disk, for 186 disks of 93 patients

### 3.4 Using p53 staining to distinguish between patients with good and poor survival

Inspection of the distribution of the p53 stain intensity might suggest that the broad distribution was a composite of two subsets, for instance one subset with intensities greater than the mean, 42.8, and one subset with intensities less than the mean. Patients that expressed more or less of p53 might then be expected to differ in

survival. To test this possibility a single cut point at the mean of the p53 distribution was used to separate two subsets of patients. The difference in survival of the two subsets, Figure 3, was not significant, Log-Rank test  $p = 0.5999$ , Wilcoxon test  $p = 0.3401$ .

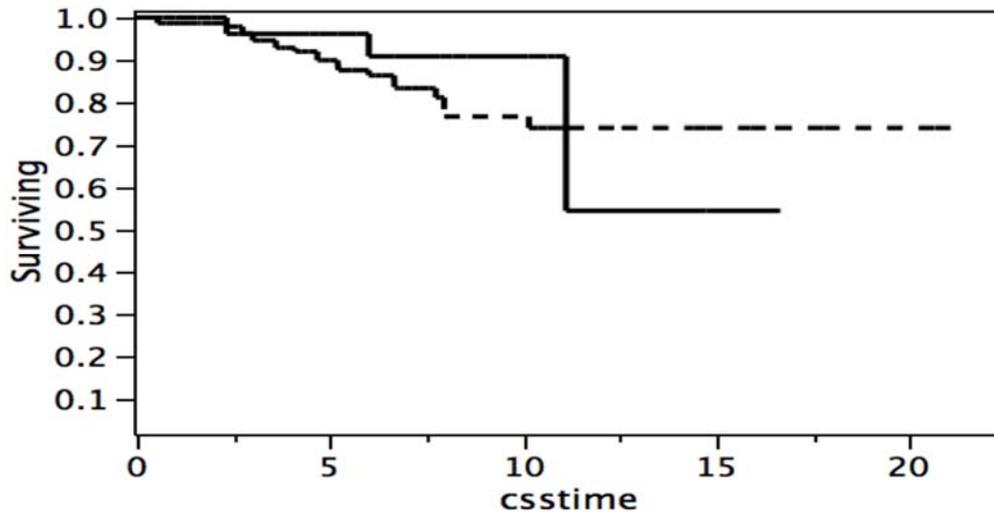


Figure 3. Kaplan-Meier survival curves, patients classified by p53 stain intensity below or above the mean of the distribution in Figure 2. Below the mean (---),  $n = 153$ ; above the mean (\_\_\_),  $n = 33$ . “csstime” = cause specific time. Log-Rank test  $p = 0.5999$ , Wilcoxon test  $p = 0.3401$

An alternative to deriving two distinct subsets by a clean separation at the mean, is to consider the possibility that the broad distribution is a composite of two overlapping normal distributions. This would be consistent with the apparent minor peak as well as the major peak of the distribution in Figure 2. The Normal Mixtures method was used to derive two normally distributed overlapping clusters, one with a mean of 38.8 and standard deviation of 3.8, and a second with a mean of 47.5 and standard deviation of 8.3. The difference in survival of the two clusters, Figure 4, was not significant, Log-Rank test  $p = 0.4488$ , Wilcoxon test  $p = 0.4641$ .

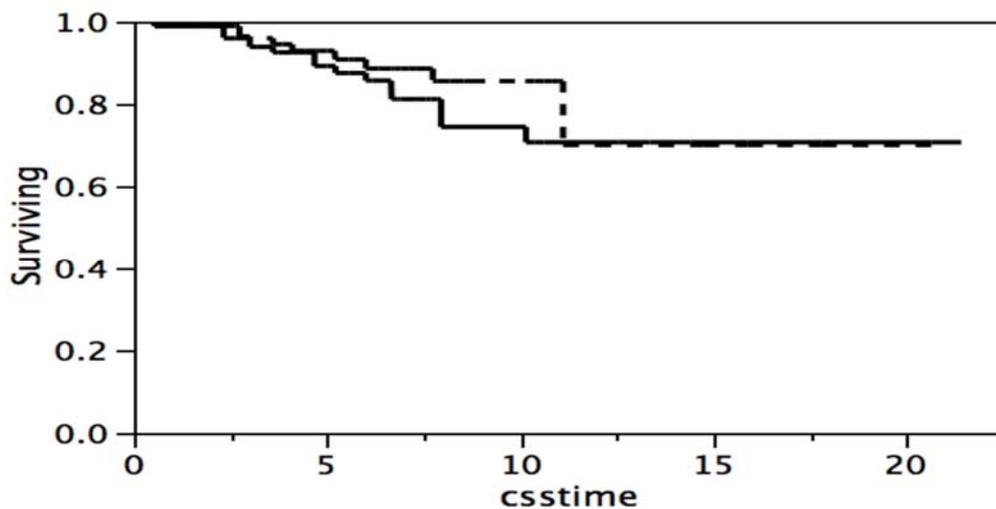


Figure 4. Kaplan-Meier survival curves, patients classified by p53 stain intensity as two overlapping normally distributed clusters. A composite of the two clusters is the distribution in Figure 2. One cluster has a mean of 38.8 (\_\_\_),  $n = 97$ ; and the other a mean of 47.5 (---)  $n = 89$ . Log-Rank test  $p = 0.4488$ , and Wilcoxon text  $p = 0.4641$ .

A broad distribution of immunohistochemical staining is frequently divided into three categories, low, medium, and high, and summarized as +1, +2, +3. To test the possibility that there are three subsets that differ in survival the broad distribution was separated using two cut points, one at one standard deviation below the mean, and a second at one standard deviation above the mean. The difference in survival of the three subsets of patients, Figure 5, was not significant, Log-Rank test  $p = 0.1138$ , Wilcoxon test  $p = 0.802$ .

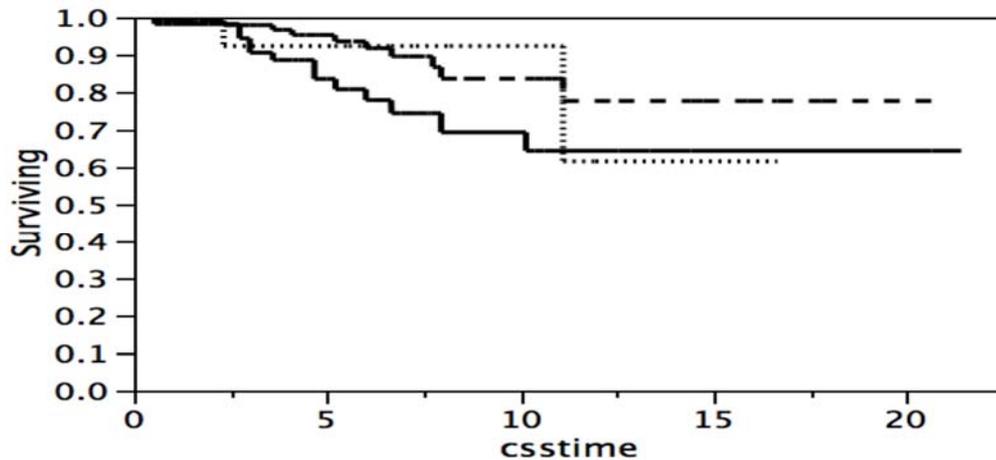


Figure 5. Kaplan-Meier survival curves, patients classified by p53 stain intensity with two cut points of the distribution in Figure 2. p53 stain intensity one standard deviation below the mean (\_\_\_\_),  $n = 57$ ; within one standard deviation of the mean (- - -),  $n = 115$ ; or one standard deviation above the mean (.....),  $n = 14$ . Log-Rank test  $p = 0.1138$ , Wilcoxon test  $p = 0.802$ .

The poor survival of the subset of patients with stain intensity one standard deviation below the mean in Figure 5 suggested the possibility that one cut point at one standard deviation below the mean might separate patients into two subsets which differed in survival. Figure 6 shows the survival of two subsets of patients, one subset with p53 stain intensities below that single cut point, and one subset above that single cut point. The difference in survival is significant, Log-Rank test  $p = 0.0372$ , Wilcoxon test  $p = 0.0247$ . It should be noted that patients expressing *less* p53 stain intensity have the *poorest* survival. This result, which might be contrary to expectations, is commented upon in the Discussion section.

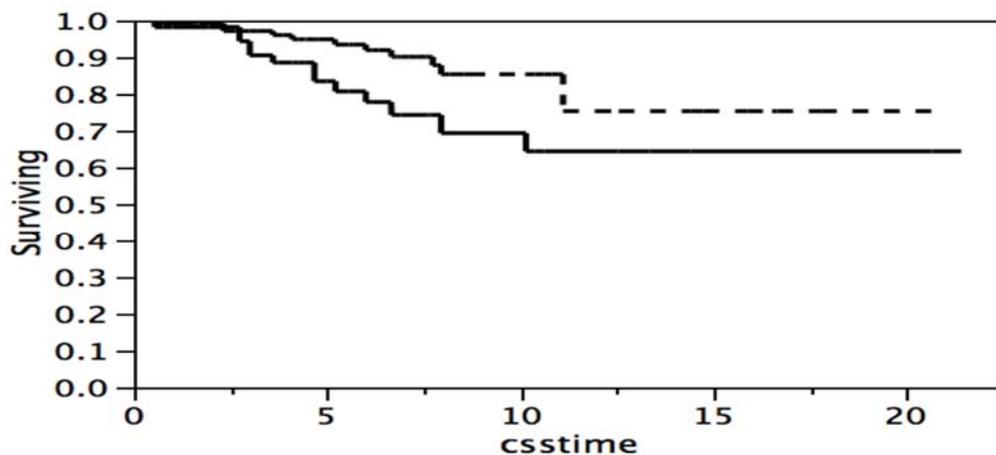


Figure 6. Kaplan-Meier survival curves, patients classified by p53 stain intensity one standard deviation below the mean, or above the mean of the distribution in Figure 2. p53 stain intensity one standard deviation below the mean (\_\_\_\_),  $n = 57$ , or above one standard deviation below the mean (- - -),  $n = 129$ . Log-Rank test  $p = 0.0372$ , Wilcoxon test  $p = 0.0247$ .

### 3.5 Comparison of p53 with Estrogen Receptor Status and Progesterone Receptor Status

The estrogen receptor (ER) status and progesterone receptor (PR) status have been recommended as tumor markers for breast cancer (Harris et al. 2007). However, p53 determined by immunohistochemistry was not previously recommended because of the disparity of reports, as described in the Introduction. Figure 6 indicates that p53 was able to stratify patients with significantly different survivals, if p53 was determined by quantitative immunohistochemistry and patients were separated with a cut point at one standard deviation below the mean of the p53 distribution. The ability of p53 determined by this procedure to contribute to ER and PR was evaluated by multivariate Cox proportional hazard analysis. Table 2 indicates that p53 can provide significant information in addition to that provided by ER status and PR status.

Table 2. Multivariate Cox proportional hazard analysis of estrogen receptor status, progesterone status, and p53

Variable	P value	Hazard Ratio (95% CI)
Estrogen receptor	0.38	0.63 (0.23-1.81)
Progesterone receptor	0.37	1.59 (0.58-4.55)
p53	0.034	2.55 (1.07-6.18)

Estrogen receptor and progesterone receptor status were each a binary variable, 0 or 1. The p53 immunohistochemistry was a binary variable, less than or greater than, one standard deviation below the mean of the p53 immunohistochemistry distribution. CI = confidence interval

### 3.6 Confirmation of p53 Classification of Patients by Quantitative Nuclear Grade

There is a danger in comparing several cut points and choosing the one for which the difference in survival between subsets of patients have the lowest p-value (Altman et al. 1994; Axelrod et al., 2009). We found that a cut point of one standard deviation below the mean of the p53 staining intensity distribution could distinguish two subsets of patients with significantly different survival. In order to have confidence in these subsets of patients we asked if the same set of patients would be identified by a different way of measuring specimens (Quantitative Nuclear Grade) and a different way of grouping the patients (k-means clustering).

Quantitative Nuclear Grade (QNG) is a composite of nuclear size, nuclear shape, nuclear DNA content, and chromatin texture. Data were acquired from a second parallel tissue microarray that had been stained with the Feulgen reaction specific for DNA, rather than stained by immunohistochemistry for p53 protein, Figure 7. Of the 93 patients that were previously analyzed on the p53 microarray, 82 patients had at least one informative disk position on each of the two parallel microarrays. Where two disks were available the one corresponding to the worst survival was used. The QNG's of the 82 patients were clustered by k-means ( $n = 3$ ), rather than by a cut point at one standard deviation below the mean as for the p53 stain intensity. The three k-means clusters were grouped into two subsets by combining two of the clusters. The two subsets of patients determined by QNG and the two subsets determined by p53 immunohistochemistry are compared in a 2 x 2 contingency table, Table 3. The chi-squared value = 7.757, and  $p = 0.0054$ . These results indicate that QNG and p53 independently classify similar groups of patients for good and poor survival. This lends confidence to the conclusion that patients with good and poor survival can be distinguished by a single cut point of the p53 stain intensity distribution at one standard deviation below the mean.

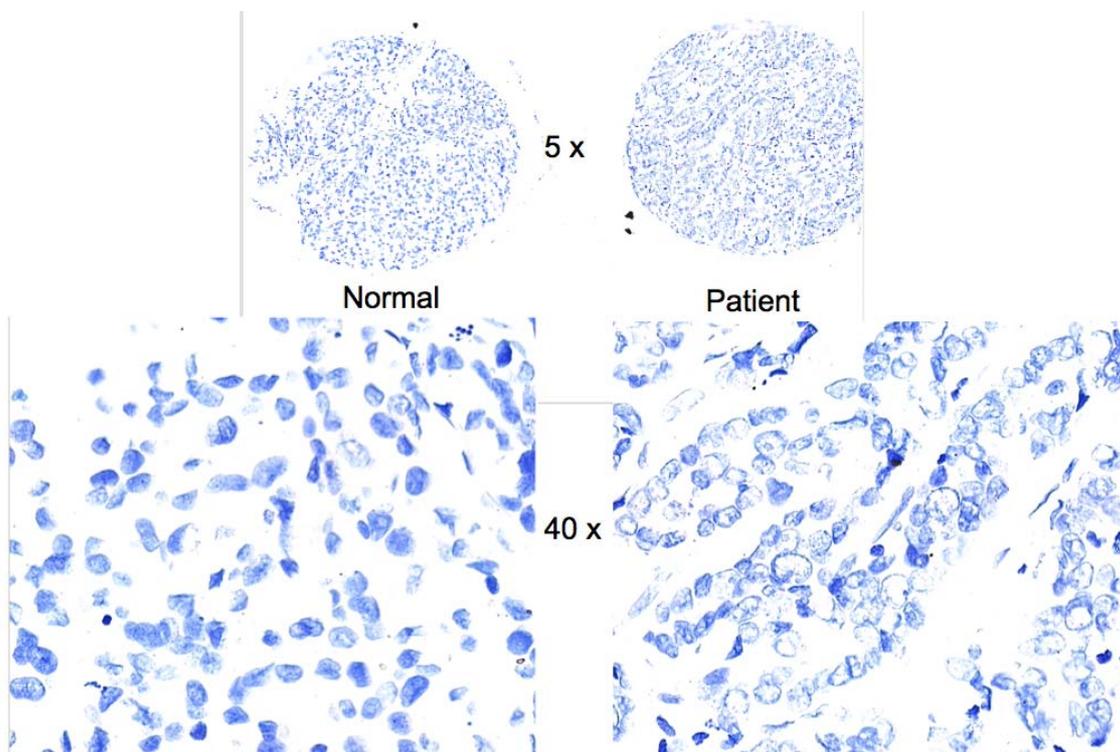


Figure 7. Disks from the tissue microarray stained with DNA-specific Feulgen reaction. Left: tissue from normal breast. Right: tissue from breast cancer patient. Top: image at low magnification used to locate each disk. Bottom: part of images at higher magnification used to extract nuclear features.

Table 3. Comparison of patients classified for survival by Quantitative Nuclear Grade and by p53 immunohistochemistry

Quantitative Nuclear Grade (Feulgen, DNA)	p53 (Immunohistochemistry, Protein)		
	Survival Poor	Survival Good	
Survival Poor	22	15	37
Survival Good	13	32	45
	35	47	82

Chi-squared = 7.757,  $p = 0.0054$

#### 4. Discussion

We have shown that p53 immunohistochemical stain intensity of tissue can distinguish young breast cancer patients with good and poor survival, if a number of factors are taken into account. These include, choice of a cohort of similarly treated patients, several tissue samples from each tumor to account for possible intra-tumor heterogeneity, construction of a tissue microarray which allows comparison of a broad range of patient specimens, uniform immunohistochemical treatment of the samples from various patients on the same tissue microarray, quality control of selection of informative disks with adequate amounts of tissue, quality control of stain of patient tissue compared to normal tissue, extraction of quantitative data about stain intensity using computer-aided image analysis, and appropriate choice of a cut-point to separate patients with different p53 stain intensities. Some of the benefits and possible pitfalls in using tissue microarrays have been reviewed (Camp et al., 2008).

We chose among 171 young patients with breast cancer, a cohort of 93 who were treated with radiation therapy and adjuvant hormone or chemotherapy. This was intended to reduce the confounding effect on survival of differences in therapy that might have resulted from mixing patients treated with adjuvant therapy and patients not treated with adjuvant therapy.

The tissue microarray consisted of two samples from each tumor. Some previous studies have concluded that two are sufficient to characterize each patient (Camp et al., 2000; Ganchberg et al., 2002; Gillett et al., 2000; Torhorst et al., 2001). However, other studies have suggested that for some immunohistochemical markers, including p53, that intra-tumor heterogeneity might require multiple tumor samples for tissue microarrays (Nassar et al., 2010). Heterogeneity of biomarkers stain intensity has been observed across multiple areas of whole tissue sections (Tolles et al., 2011). We have shown that there can be intra-tumor heterogeneity between different ducts of the same patient (Miller et al., 2010). In the current study, we determined that the correlation coefficient of p53 stain intensity between two of disks from two cores from the same patient was 0.427. This suggests that the two samples from any one patient would be expected to be similar but not identical. Constructing tissue microarrays with disks from more than two cores, and selecting the disk corresponding the worst survival might improve the assessment of each patient.

The intensity of p53 staining by immunohistochemistry was determined by computer-aided image analysis. There was a broad inter-tumor heterogeneity, coefficient of variation = 17%. It is unlikely that this broad distribution was due to instrument error since repeated measurements of stain intensity of the same area were indistinguishable. Rather, it is likely that this broad distribution reflects the inter-tumor heterogeneity between patients. Immunohistochemistry stain intensity is frequently interpreted as being proportional to the protein that interacts with the primary antibody. However, the relationship between stain intensity and amount of protein is not always so simple. Some mutations in p53 do not result in positive immunohistochemical staining and some positive p53 immunohistochemical staining can occur in the absence of p53 mutations (Rohan et al., 2006). The importance of immunological methodology has been demonstrated dramatically by the study of McCabe et al. (2005) which showed that at low concentrations of primary antibody patients expressing less p53 had better survival, but that at high concentrations of primary antibody patients expression less p53 had poorer survival. We also observed, that under the immunohistochemistry protocol used in our studies, that the patients expressing less p53 stain intensity had the poorer survival.

We examined several methods of grouping patients based on their quantitative stain intensity and determined if the groups differed in survival. Previously, immunohistochemistry stain intensity was evaluated by visual inspection and specimens reported as belonging in two groups (high and low, greater or less than the average), or into three groups (low, medium, and high, or +1, +2, and +3). Sometimes patient specimens were compared with standards made from cell lines. More recently, immunohistochemical stain intensities have been measured by computer-aided image analysis using commercially available instruments. Such instruments provide objective and reproducible measurements. Some of instrument manufactures provide analysis of data similar to that previously reported for visual inspection, e.g. in two groups (high or low stain intensity) or three groups (+1, +2, +3 stain intensity). Camp et al. (2004) have devised a computer program that divides the stain intensity distribution with many sets of two cut-points each producing three subsets, and selects the optimal cut-point values based on the survival of the patients. In our study we used stain intensity measured by quantitative image analysis, and compared the ability of different methods of grouping specimens. We found that neither one cut-point at the mean, nor one cut-point to separate two possible overlapping distributions was successful in distinguishing the survival of the two groups of patients. Also, two cut-points producing three groups of patients did not result in three groups of patients that differed in survival. However, a single cut-point at one standard deviation below the mean did produce two groups of patients that differed in survival. This method of determining an effective cut-point is useful since it is so simple to calculate. P53 determined in this way provides additional information beyond that provided by the patient's estrogen receptor status and progesterone receptor status.

Comparing several different cut-points and choosing one apparently best cut-point can be misleading. There may be no one uniquely optimal cut-point for a continuous distribution (Axelrod et al., 2009) and multiple testing may result in an inflation of a type I error (Altman et al., 1994). With these considerations in mind, we chose to confirm the groups of patients determined by the one cut-point at one standard deviation below the mean with an entirely different method of grouping patients. In addition to characterizing each patient by p53 protein stain intensity by immunohistochemistry we also characterized each patient by DNA Feulgen staining to determine a quantitative nuclear grade. Nuclear grade is one of the most important features proposed for histological classification of breast lesions (Jaffer & Bleiweiss, 2002). We, and many others, have shown that image analysis of nuclear features (size, shape, chromatin texture, and ploidy) can be prognostic for survival of breast cancer

patients (Chapman et al., 2007). There have been previous reports of an association between p53 immunohistochemistry and nuclear features consistent with our observations, although a relationship with patient survival was not determined (Haroske et al., 1996; Friedrich et al., 1997). In this study we have shown that the protein based method and the DNA based method identify similar groups of patients with good and poor survival.

## 5. Conclusions

Immunohistochemistry procedures are convenient, and several biomarkers have been recommended for breast cancer. However, because of the disparate reports of the relationship between p53 stain intensity detected by immunohistochemistry and patient outcome, p53 is not yet among the biomarkers recommended for routine clinical practice (Harris et al., 2007). By taking into account several experimental factors and methods of data analysis we have found that p53 stain intensity can stratify young breast cancer patients for survival, and provide information beyond that provided by ER status and PR status.

## Authors' Contributions

DEA and BH conceived of the study. QY and BH provided the tissue microarray and patient data. DEA and KS extracted the data, did the statistical analyses, and prepared the manuscript. All authors approved of the final manuscript. No competing interests.

## References

- Allred, D. C., Clark, G. M., Elledge, R., Fuqua, S. A. W., Brown, R. W., Chamness, G. C., ... McGuire, W. L. (1993). Association of p53 protein expression with tumor cell proliferation rate and clinical outcome in node-negative breast cancer. *Journal of the National Cancer Institute*, *85*, 200-206. <http://dx.doi.org/10.1093/jnci/85.3.200>
- Altman, D. G., Lausen, B., Sauerbrei, W., & Schumacher, M. (1994). Dangers of using "optimal" cut-points in the evaluation of prognostic factors. *Journal of the National Cancer Institute*, *86*, 829-835. <http://dx.doi.org/10.1093/jnci/86.11.829>
- Axelrod, D. E., Miller, N., & Chapman, J.-A. (2009). Avoiding pitfalls in the statistical analysis of heterogeneous tumors. *Biomedical Informatics Insights*, *2*, 11-18. <http://www.la-press.com/avoiding-pitfalls-in-the-statistical-analysis-of-heterogeneous-tumors-article-a1374>
- Bacus, J. W., Boone, C. W., Bacus, J. V., Follen, M., Kelloff, G. J., Kagan, V., & Lippman, S. M. (1999). Image morphometric nuclear grading of intrepithelial neoplastic lesions with applications to cancer chemoprevention trials. *Cancer Epidemiology Biomarkers & Prevention*, *8*, 1087-1094. Retrieved from <http://cebp.aacrjournals.org/content/8/12/1087.full.pdf+html?sid=c6172a14-4445-4899-bc69-82bdc0433163>
- Bartley, A. N., & Ross, D. W. (2002). Validation of p53 immunohistochemistry as a prognostic factor in breast cancer in clinical practice. *Archives of Pathology and Laboratory Medicine*, *126*, 456-458. <http://www.archivesofpathology.org/doi/pdf/10.1043/0003-9985%282002%29126%3C0456%3AVOPIAA%3E2.0.CO%3B2>
- Camp, R., Dolled-Filhart, M., & Rimm, D. (2004). X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clinical Cancer Research*, *10*, 7252-7259. <http://dx.doi.org/10.1158/1078-0432.CCR-04-0713>
- Camp, R. L., Neumeister, V., & Rimm, D. L. (2008). A decade of tissue microarrays: Progress in the discovery and validation of cancer biomarkers. *Journal of Clinical Oncology*, *26*, 5630-5637. <http://dx.doi.org/10.1200/JCO.2008.17.3567>
- Camp, R. L., Charette, L. A., & Rimm, D. L. (2000). Validation of tissue microarray technology in breast carcinoma. *Laboratory Investigation*, *80*, 1943-1949. <http://dx.doi.org/10.1038/labinvest.3780204>
- Chapman, J.-A.W., Miller, N. A., Lickley, H. L. A., Qian, J., Christens-Barry, W. A., Fu, Y., ... Axelrod, D. E. (2007). Ductal carcinoma in situ of the breast (DCIS) with heterogeneity of nuclear grade: prognostic effects of quantitative nuclear assessment. *BMC Cancer*, *7*, 174-183. <http://dx.doi.org/10.1186/1471-2407-7-174>
- Elkhuizen, P. H. M., Voogd, A. C., van den Broek, L. C. J. M., Tan, I. T. C., van Houwelingen, H. C., Leer, J.-W. H., & van de Vijver, M. J. (1999). Risk factors for local recurrence after breast-conserving therapy for invasive carcinomas: A case-control study of histological factors and alterations in oncogene expression. *International Journal of Radiation Oncology and Radiation Physics*, *45*, 73-83.

- [http://dx.doi.org/10.1016/S0360-3016\(99\)00158-3](http://dx.doi.org/10.1016/S0360-3016(99)00158-3)
- Elledge, R. M., Clark, G. M., Fuqua, S. A. W., Yu, Y.-Y., & Allred, D. C. (1994). P53 protein accumulation detected by five different antibodies: Relationship to prognosis and heat shock protein 70 in breast cancer. *Cancer Research*, *54*, 3752-3757. <http://cancerres.aacrjournals.org/content/54/14/3752.full.pdf+html?sid=5feaf59f-efd2-4b57-9741-745756c7534b>
- Friedrich, K., Dimmer, V., Haroske, G., Meyer, W., Theissig, F., & Kunze, K. D. (1997). Correlation between p53 status, DNA ploidy, nuclear morphology in breast cancer. An image cytometry study. *Analytical Cell Pathology*, *15*, 85-97. <http://iospress.metapress.com/content/mel722fk37d2pb70/fulltext.pdf>
- Ganchberg, D., Di Leo, A., Rouas, G., Järvinen, T., Verhest, A., Isola, J., ... Larsimont, D. (2002). Reliability of the tissue microarray based FISH for evaluation of the HER-2 oncogene in breast carcinoma. *Journal of Clinical Pathology*, *55*, 315-317. <http://jcp.bmj.com/content/55/4/315.full.pdf+html?sid=f30b5062-4503-4805-9958-c8be795ccf6d>
- Gillett, C. E., Springall, R. J., Barnes, D. M., & Hanby, A. M. (2000). Multiple tissue core arrays in histopathology research: a validation study. *Journal of Pathology*, *192*, 549-553. [http://dx.doi.org/10.1002/1096-9896\(2000\)9999:9999<::AID-PATH721>3.0.CO;2-0](http://dx.doi.org/10.1002/1096-9896(2000)9999:9999<::AID-PATH721>3.0.CO;2-0)
- Haffty, B. G., Hauser, A., Choi, D. H., Parisot, N., Rimm, D., King, B., & Carter, D. (2004). Molecular markers for prognosis after isolated postmastectomy chest wall recurrence. *Cancer*, *100*, 252-263. <http://dx.doi.org/10.1002/cncr.11915>
- Hainaut, P., & Hollstein, M. (2000). P53 and human cancer: The first ten thousand mutations. *Advances in Cancer Research*, *77*, 81-137. [http://dx.doi.org/10.1016/S0065-230X\(08\)60785-X](http://dx.doi.org/10.1016/S0065-230X(08)60785-X)
- Haralick, R. M., Shaanmugam, K., & Dinstein, I. (1973). Texture features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, *SMC-3*, 610-621. <http://dx.doi.org/10.1109/TSMC.1973.4309314>
- Haroske, G., Dimmer, V., Friedrich, K., Meyer, W., Thieme, B., & Theissig, F. (1996). Nuclear image analysis of immunohistochemically stained cells in breast carcinomas. *Histochemistry and Cell Biology*, *105*, 479-485. <http://dx.doi.org/10.1007/BF01457662>
- Harris, L., Fritsche, H., Mennel, R., Norton, L., Ravdin, P., Taube, S., ... Bast Jr., R. C. (2007). American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. *Journal of Clinical Oncology*, *25*, 5287-5312. <http://dx.doi.org/10.1200/JCO.2007.14.2364>
- Henson, D. E. (2005). Back to the drawing board on immunohistochemistry and predictive factors. *Journal of the National Cancer Institute*, *97*, 1796-1997. <http://dx.doi.org/10.1093/jnci/dji449>
- Jaffer, S., & Bleiweiss, I. J. (2002). Histological classification of ductal carcinoma in situ. *Microscopy Research and Technique*, *59*, 92-101. <http://dx.doi.org/10.1002/jemt.10180>
- McCabe, A., Dolled-Filhart, M., Camp, R. L., & Rimm, D. L. (2005). Automated quantitative analysis (AQUA) of in situ protein expression, antibody concentration, and prognosis. *Journal of the National Cancer Institute*, *97*, 1808-1815. <http://dx.doi.org/10.1093/jnci/dji427>
- Miller, N. A., Chapman, J.-A. W., Qian, J., Christens-Barry, W. A., Fu, Y., Yuan, Y., ... Axelrod, D. E. (2010). Heterogeneity between ducts of the same nuclear grade involved by duct carcinoma in situ (DCIS) of the breast. *Cancer Informatics*, *9*, 209-216. <http://www.la-press.com/heterogeneity-between-ducts-of-the-same-nuclear-grade-involved-by-article-a2250>
- Nassar, A., Radhakrishnan, A., Cabrero, I., Cotsonis, G., & Cohen, C. (2010). Intratumoral heterogeneity of immunohistochemical marker expression in breast carcinoma: a tissue microarray-based study. *Applied Immunohistochemistry & Molecular Morphology*, *18*, 433-441. <http://graphics.tx.ovid.com.libproxy2.umdj.edu/ovftpdfs/FPDDNCOBCECKGF00/fs046/ovft/live/gv023/00129039/00129039-201010000-00006.pdf>
- Osborne, R. J., Merlo, G. R., Mitsudomi, T., Venesio, T., Liscia, D. S., Cappa, A. P. M., ..., Minna, J. D. (1991). Mutations in the p53 gene in primary human breast cancers. *Cancer Research*, *51*, 6194-6198. <http://cancerres.aacrjournals.org/content/51/22/6194.full.pdf+html?sid=4f61c7df-0b72-4bd9-adb5-8d61fa0b900c>
- Reed, W., Hannisdal, E., Boehler, P. J., Gundersen, S., Host, H., & Nesland, J. M. (2000). The prognostic value of p53 and c-erb B-2 immunostaining is overrated for patients with lymph node negative breast carcinoma. *Cancer*, *88*, 804-813. [http://dx.doi.org/10.1002/\(SICI\)1097-0142\(20000215\)88:4<804::AID-CNCR11>3.0](http://dx.doi.org/10.1002/(SICI)1097-0142(20000215)88:4<804::AID-CNCR11>3.0)

CO;2-Y

- Ringberg, A., Anagnostaki, L., Anderson, H., Idvall, I., & Fernö, M. (2001). Cell biological factors in ductal carcinoma *in situ* (DCIS) of breast-relationship to ipsilateral local recurrence and histopathological characteristics. *European Journal of Cancer*, *37*, 1514-1522. [http://dx.doi.org/10.1016/S0959-8049\(01\)00165-4](http://dx.doi.org/10.1016/S0959-8049(01)00165-4)
- Rohan, T. E., Li, S. Q., Hartwick, R., & Kandel, R. A. (2006). P53 alterations and protein accumulation in benign breast tissue and breast cancer risk: A cohort study. *Cancer Epidemiology, Biomarkers & Prevention*, *15*, 1316-1329. <http://dx.doi.org/10.1158/1055-9965.EPI-06-0195>
- Sjögren, S., Inganäs, M., Norberg, T., Lindgren, A., Nordgren, H., Holmberg, L., & Bergh, J. (1996). The p53 gene in breast cancer: Prognostic value of complementary DNA sequencing versus immunohistochemistry. *Journal of the National Cancer Institute*, *88*, 173-182. <http://dx.doi.org/10.1093/jnci/88.3.173>
- Tolles, J., Bai, Y., Baquero, M., Harris, L. N., Rimm, D. L., & Molinaro, A. M. (2011). Optimal tumor sampling for immunostaining of biomarkers in breast carcinoma. *Breast Cancer Research*, *13*, R51. <http://dx.doi.org/10.1186/bcr2882>
- Torhorst, J., Bucher, B., Kononen, J., Hass, P., Zuber, M., Köchi, O. R., ... Sauter, G. (2001). Tissue microarrays for rapid linking of molecular changes to clinical endpoints. *American Journal of Pathology*, *159*, 2249-2256. [http://dx.doi.org/10.1016/S0002-9440\(10\)63075-1](http://dx.doi.org/10.1016/S0002-9440(10)63075-1)
- Turner, B. C., Gumbs, A. A., Carbone, C. J., Carter, D., Glazer, P. M., & Haffty, B. G. (2000). Mutant p53 protein overexpression in women with ipsilateral breast tumor recurrence following lumpectomy and radiation therapy. *Cancer*, *88*, 1091-1098. [http://dx.doi.org/10.1002/\(SICI\)1097-0142\(20000301\)88:5<1091::AID-CNCR21>3.3.CO;2-P](http://dx.doi.org/10.1002/(SICI)1097-0142(20000301)88:5<1091::AID-CNCR21>3.3.CO;2-P)
- Zellars, R. C., Hilsenbeck, S. G., Clark, G. M., Allred, D. C., & Herman, T. S. (2000). Chamness, G.C., Elledge, R.M. Prognostic value of P53 for local failure in mastectomy-treated breast cancer patients. *Journal of Clinical Oncology*, *18*, 1906-1913. Retrieved from <http://jco.ascopubs.org/content/18/9/1906.full.pdf+html?sid=06f74d16-f557-4825-be85-d39342ea6250>