Quantitative assessment shows loss of antigenic epitopes as a function of pre-analytic variables

Yalai Bai[1], Juliana Tolles[2], Huan Cheng[1], Summar Siddiqui[1], Arun Gopinath[1], Eirini Pectasides[1], Robert L Camp[1], David L Rimm[1] and Annette M Molinaro[2]

Pre-analytic variables, specifically cold ischemic time, have been implicated as key variables in the measurement of proteins by immunohistochemistry. To determine the significance and magnitude of antigenic loss due to pre-analytic variables, we have compared protein antigenicity in core needle biopsies, with essentially no cold ischemic time, with that in routinely processed tumor resection specimens. Two cohorts of matched core needle biopsies and tumor resections were collected with 20 matched pairs and 14 matched pairs, respectively. Both series were analyzed by quantitative immunofluorescence using the AQUA® method. Epitopes phospho-ERK, total ERK, phospho-AKT, total AKT, phospho-S6K1, total S6K1, estrogen receptor (ER), Ki67, cytokeratin and GAPDH were assessed. Detection levels for all phospho-epitopes were significantly decreased in tumor resections compared with biopsies while no significant change was seen in the corresponding total proteins. Of the other four proteins examined, ER and cytokeratin showed significant loss of antigenicity. This data suggest that measurement of phospho-protein antigenicity in formalin-fixed tissue by immunological methods is dramatically affected by pre-analytic variables. This study suggests that core needle biopsies are more accurate for assessment of tissue biomarkers.

Standard specimen preparation in anatomic pathology often involves long cold ischemic time or other pre-analytic variables. Pre-analytic variables are defined as conditions that can alter the diagnostic assessment of the tissue, particularly by immunological methods. The list of pre-analytic variables is long and can be studied on an individual basis or collectively considered as a 'black box'.[1] Individual studies have examined: (1) time interval starting from blood vessel ligation until fixation;[2] (2) the temperature of the specimen before formalin fixation;[3] and (3) the rate of formalin penetration and biological inactivation.[4] However, there are many other variables that are much harder to address, including things such as the brand and lot of formalin (with potential for minor but significant variations in concentration), the humidity in the air during the cold ischemic time, and even individual variations in surgical technique that could produce biological changes before the initiation of cold ischemia. In human tissue, the most rapid inactivation of biological processes occurs after the immediate fixation of a core needle biopsy. Core needle biopsies are a fairly standardized procedure. They facilitate quick formalin fixation with rapid biological inactivation, due to small diameter of specimens.[4] Thus, in this study, we chose core needle biopsies as the standard, representing the minimal pre-analytic variation achievable in a human tissue system. We compared core needle biopsies with tumor resection specimens in the same patients. There are hundreds or thousands of uncontrolled variables in tumor resection specimens, compared with core needle biopsies. Because it is impractical—maybe impossible—to dissect out each variable, we treated the pre-analytic variables as a 'black box' and assessed changes seen in tumor resections as compared with the defined standard, core needle biopsies.

Although the biological processes affected by pre-analytic variation are numerous, one of the most commonly assessed is phosphorylation.[5] Phospho-protein markers of

---

[1]Department of Pathology, Yale University School of Medicine, New Haven, CT, USA and [2]Division of Biostatistics, Yale University School of Public Health, New Haven, CT, USA
Correspondence: Dr AM Molinaro, PhD, Division of Biostatistics, LEPH 206, Yale University School of Public Health, 60 College Street, New Haven, CT 06520, USA.
E-mail: annette.molinaro@yale.edu

cell signaling have been broadly implicated in both translational research and clinical settings to predict the response to chemotherapy or prognosis. Response to EGFR inhibitors in lung cancer or glioblastoma and trastuzumab in breast cancer has been reported to be predicted by measurement of a series of phospho-protein markers including phospho-AKT (pAKT), phospho-ERK (pERK) and phospho-HER2 (pHER2).[6–14] However, at least in some cases, reproduction of these findings has been difficult. Furthermore, a negative result (or even a series of negative results) may never be published to refute an observation that may be highly dependent on pre-analytic conditions. As a result, phospho-protein assessment has not become a standard companion diagnostic for any therapy, even though pre-clinical data suggest it would be both accurate and specific.[5,15,16]

Recently, a group of oncologists and pathologists convened to set standards for immunohistochemical testing. When they examined the literature, they found relatively little information addressing the pre-analytic variation. Despite this, they produced a set of standards aimed at minimizing pre-analytical variation, first for Her2[17] and subsequently for estrogen receptor (ER).[18] Their effort increased awareness of the pre-analytic variation issue and prompted us to take a quantitative look at the problem, with special attention to phospho-modified epitopes. First, we performed a pilot study of a series of samples on a tissue microarray (TMA). The results of that pilot prompted the more thorough study that follows, which assesses pairs of core needle biopsies and tumor resections for antigenicity levels of several proteins and phospho-proteins.

MATERIALS AND METHODS
Cohorts
The formalin-fixed paraffin-embedded primary invasive breast cancer tumors were obtained from patients who underwent core needle biopsies and subsequent surgery at Yale University/New Haven Hospital from 1998 to 2005. All patients were diagnosed with infiltrating ductal carcinoma of the breast. The first cohort consisted of 20 paired core needle biopsies and tumor resections obtained from 1998 to 2000 from the archives of the Pathology Department of Yale University. This cohort was analyzed in a TMA format. The second cohort consisted of 14 paired core needle biopsies and tumor resections obtained from 2001 to 2005 from the archives of the Pathology Department of Yale University. This second cohort was analyzed in the conventional whole tissue section format. During this time period, the estimated range of time to fixation was as short as 1 h to as long as 4–5 h. Data on age; type of procedure; size of specimen grossly; size of the tumor microscopically; the status of specimen upon receipt; ER, PR and HER2 status; node status; nuclear grade and histology grade are shown in Table 1. No patients received chemotherapy or radiation before resection. The median time from core needle biopsy to tumor resection was 27.5 days. The study was approved by the institutional review board for Yale University.

Antibodies and Immunohistochemistry
The arrays and whole tissue sections were incubated at 60°C for 20 min before being deparaffinized with xylene, rehydrated, and antigen retrieved by pressure cooking for 15 min

Table 1 Clinical characteristics of resection tissues

Code	Age	Procedure	Tumor size (cm)	Specimen size (cm)	Upon receiving	ER	PR	HER2	Node	Nuclear grade	Histology grade
1R	80	M	1.1 × 0.9 × 0.5	6.2 × 4 × 1	Fresh	P	P	N	N	2	6
2R	45	M	1.8 × 1.2 × 1.1	26 × 24.5 × 6	Fresh	P	P	N	N	2	5
3R	63	L	1.5 × 1.5 × 1.4	11.6 × 5.5 × 2	Fresh	P	P	N	N	3	6
4R	53	L	0.8 × 0.6	4.7 × 3 × 2.5	Fixed	P	P	N	(N1)	2	4
5R	52	M (p)	1.3 × 0.5 × 0.4	7.5 × 7 × 1	Fresh	P	P	(w)P	N	2	5
6R	51	M (p)	0.2	7 × 5 × 2	Fresh	P	P	N	N	N/D	N/D
7R	51	L	1.1	9 × 7 × 1.8	Fresh	P	P	N	N	3	7
8R	45	L	1.8	3.5 × 3 × 1	Fresh	N	N	(w)P	N	3	3
9R	75	M	5	17 × 13 × 4	Fresh	P	P	P	N	3	3
10R	40	M (r)	3.5	24 × 22 × 5	Fresh	P	P	(w)P	(N1)	3	3
11R	44	M	0.99	17 × 16 × 2.5	Fresh	P	P	P	N	3	6
12R	88	L	4	11 × 6 × 4.5	Fresh	N	N	(w)P	N	2	7
13R	39	M (r)	8.5	18 × 14 × 7	Fresh	N	N	N	(N2)	3	2
14R	78	L	1.5 × 1.2	5 × 4.5 × 2	Fresh	P	P	N	(N1)	2	6

Abbreviations: R: resection; M: mastectomy; L: lumpectomy; (p): partial; (r): radical; P: positive; (w)P: weakly positive; N: negative; N/D: not determined.

in citrate buffer (pH = 6). Slides were pre-incubated with 0.3% bovine serum albumin in 0.1 mol/l TBS (pH = 8) for 30 min at room temperature. The procedure for pAKT staining was as follows: slides were incubated with a cocktail of pSer473-AKT antibody diluted at 1:1000 (Mouse monoclonal, clone 587F11; Cell Signaling Technology, Danvers, MA) and a wide-spectrum rabbit anti-cow cytokeratin antibody (Z0622; Dako Corp, Carpinteria, CA) diluted 1:100 in bovine serum albumin/TBS overnight at 4°C. This was followed by a 1-h incubation at room temperature with Alexa 546-conjugated goat anti-rabbit secondary antibody (A11010; Molecular Probes, Eugene, OR) diluted 1:100 in mouse EnVision reagent (K4001, Dako Corp). Cyanine 5 (Cy5) directly conjugated to tyramide (FP1117; Perkin-Elmer, Boston, MA) at a 1:50 dilution was used as the fluorescent chromogen for pAKT detection. Prolong mounting medium (Prolong Gold, P36931; Molecular Probes) containing 4′,6-diamidino-2-phenylindole was used to identify tissue nuclei. Other antibodies performed in a similar manner are listed in Table 2.

Automated Quantitative Analysis

Automated Quantitative Analysis (AQUA) allows exact measurement of protein concentration within subcellular compartments, as described in detail elsewhere.[19,20] In brief, a series of high-resolution monochromatic images were captured by the PM-2000 microscope (HistoRx). For whole tissue sections, multiple regions of interest (ROIs) containing only invasive tumor were circled by investigators trained in histopathology on the AQUA system screen based on the cytokeratin (cytoplasm) image of the immunohistochemically stained slide taken with the AQUA system. The selected ROIs were automatically overlaid with a grid by the image capturing program and each field of view (FOV) was defined automatically. For each histospot of the TMA and each FOV of the whole tissue section, in- and out-of-focus images were obtained using the signal from the 4′,6-diamidino-2-phenylindole, cytokeratin-Alexa 546 and target protein-Cy5 channel. Target protein antigenicity was measured using a channel with emission maxima above 620 nm, in order to minimize tissue autofluorescence. Tumor was distinguished from stromal and non-stromal elements by creating an epithelial tumor 'mask' from the cytokeratin signal. This created a binary mask (each pixel being either 'on' or 'off') on the basis of an intensity threshold set by visual inspection of histospots. The AQUA score of the target protein in each subcellular compartment was calculated by dividing the target protein compartment pixel intensities by the area of the compartment within which they were measured. AQUA scores were normalized to the exposure time and bit depth at which the images were captured, allowing scores collected at different exposure times to be directly comparable. Standardization and reproducibility associated with AQUA are described elsewhere.[21]

Statistical Analysis

The average values for target AQUA scores from duplicate histospots or from multiple ROFs were calculated and treated as independent continuous variables. The Wilcoxon's Signed Rank test was employed to assess the paired differences, considered significant at a P-value < 0.05. Error bars in the accompanying graphs represent 95% confidence interval (CIs). Power calculations based on a paired t-test were performed for each epitope (Table 3).

For each epitope of interest, a linear mixed-effects model was fit to the AQUA scores of ROFs from full histological sections. Individual model parameterizations were selected on the basis of the Akaike Information Criterion and Bayesian Information Criterion.[22] These criteria guide selection of a parsimonious model by rewarding models for close prediction of observed data while penalizing them for use of additional parameters.

Table 2 Antibodies, epitiopes, sources and dilutions

Protein	Site	Species	Clone	Dilutions	Supplier
pAKT	Ser473	Mouse mAb	587F11	1:1000	CST
AKT		Rabbit MAb	11E7	1:1000	CST
pERK1/2	Thr202/Tyr204	Rabbit mAb	20G11	1:1000	CST
ERK1/2		Mouse mAb	L34F12	1:1000	CST
pS6K1	Thr389	Mouse mAb	1A5	1:350	CST
S6K1		Rabbit mAb	49D7	1:450	CST
GAPDH		Rabbit mAb	14C10	1:500	CST
Ki67		Mouse mAb	B56	1:500	BD Pharmingen
ER		Mouse mAb	1D5	1:50	Dako
pTyrosine	Tyr	Mouse mAb	4G10	1:20000	Upstate
p53		Mouse mAb	DO-7	1:50	Dako

Table 3 Statistical comparison between core needle biopsies and tumor resections with experimental power

Epitope	*P*-value	Experiment sample size (pair)	Experiment power	Sample size needed to give power = 0.80 (pair)	Coefficient of variation for loss of epitope (%)
pAKT	0.0039	10	0.93	7	102
AKT	0.0645	10	0.74	12	138
pERK	0.0098	10	0.98	5	86
ERK	0.3750	10	0.21	90	380
pS6K1	0.0020	10	0.99	2	52
S6K1	0.2324	10	0.46	26	205
ER	0.0273	10	0.96	6	93
Ki67	0.7910	10	0.10	516	914
GAPDH	0.4973	13	0.26	79	357
Pan-cytokeratin	0.0245	14	0.88	11	133

The form of the model is $y_{ikj} = \beta_0 + b_{0i} + r_{0ij} + (\beta_1 + b_{1i})s_i + \varepsilon_{ijk}$, where $i = 1, \ldots,$ number of subjects; $j = 1, \ldots,$ number of ROI subject i; $k = 1, \ldots,$ number of FOV in ROI j; and the factor $s_i = 0$ for core needle biopsy specimens and $s_i = 1$ for tumor resection specimens. The term ε_{ijk} is the residual error. In other words, y_{ijk} represents the AQUA score of the kth FOV from the jth ROI of the ith patient. The parameters b_0, b_1, r_0 and ε_{ijk} are assumed to be normally distributed with variances σ_1^2, σ_2^2, σ_3^2 and σ_4^2, respectively. This assumption was verified for all models. All analyses were performed with the R Program for Statistical Computing,[23] nlme package.[24]

RESULTS
### Differences in Biomarker Expression in Core Needle Biopsies *vs* Surgical Tumor Resections

To determine if there is a difference in antigenicity between core needle biopsies and subsequent tumor resections, we performed a pilot study on 20 core needle biopsies and matched tumor resections on a TMA. Antibodies to ER, Ki67, p53 and phospho-proteins including pERK, pAKT and phospho-tyrosine (pTyr) were immunofluorescently stained and the results were quantified by AQUA. Scores for each core needle biopsy and tumor resection were determined by the average of two TMA spots, then plotted in pairs as shown in Figure 1. Antigenicity of pERK, pAKT, pTyr and Ki67 was decreased in tumor resections compared with that in core needle biopsies. ER and p53 exhibited no trend for reduction in tumor resections. In these TMA-based experiments, no statistically significant differences were observed possibly due to heterogeneity between fields and/or the small amount of tissue assessed in a TMA. However, the trends seen in this pilot ultimately motivated a more complete assessment.
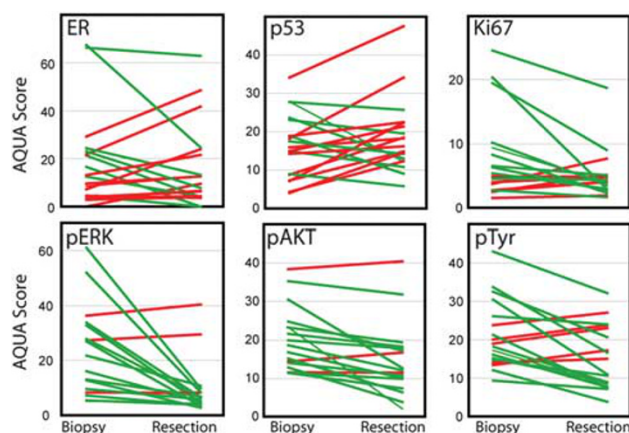
Figure 1 Differences in biomarker expression in core needle biopsies *vs* tumor resections. Twenty core needle biopsies and matched tumor resections were arrayed in TMA with two-fold redundancy. In all, 1.5 mm core from each tumor block was arrayed in a recipient block. The TMA was immunohistochemically stained with ER, p53, Ki67, pERK, pAKT and pTyr and the results were quantified using AQUA. Scores represent the average of two cores. Specimens that showed decreased staining in the resection relative to biopsy are shown in green; those with higher resection levels are shown in red.

Assessment of Phospho-Protein Epitope Loss in Paired Whole Tissue Sections

Toward a more comprehensive investigation, we assessed whole sections from matched pairs of core needle biopsies and tumor resection specimens. Representative pictures of pAKT, AKT, ER and GAPDH in biopsy tissues and paired tumor resection are shown in Figure 2. Both pAKT and ER show lower signal in a representative FOV from the tumor resection compared with the core needle biopsy (Figure 2a, b, e and f), while total AKT and GAPDH appear to show no change (Figure 2c, d, g and h) between the core needle biopsy and tumor resection.
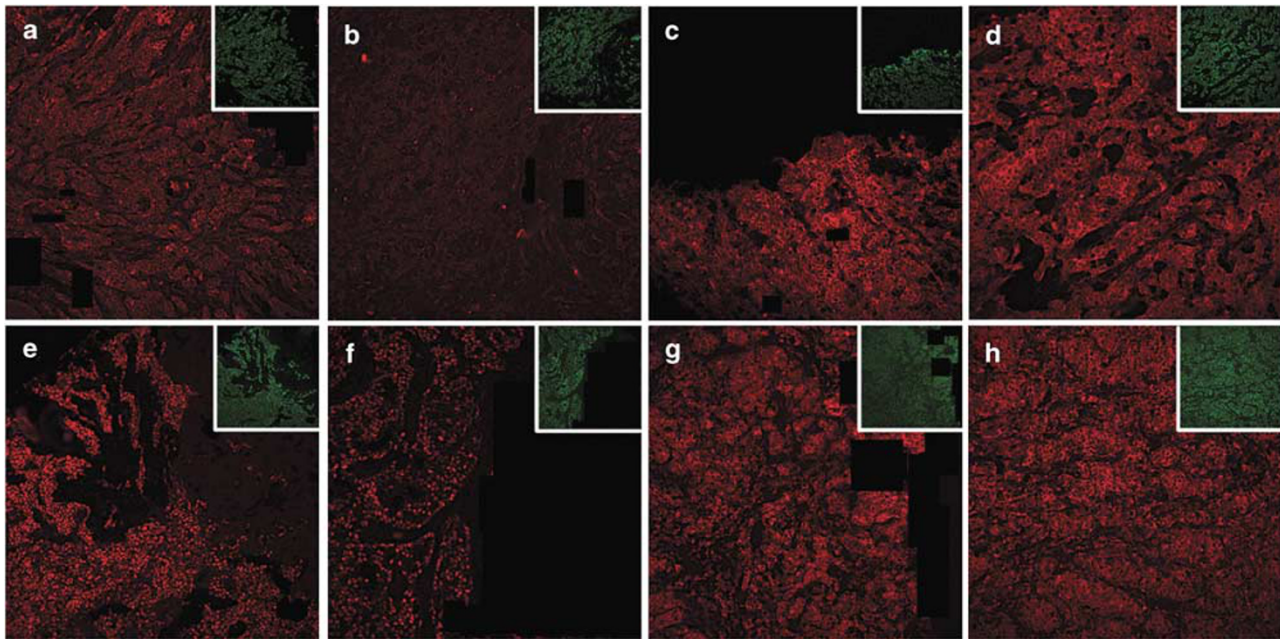
Figure 2 Comparison of biomarker staining images between biopsies *vs* tumor resections on whole tissue section slides. Representative immunofluorescence staining of pAKT (red) in CNB (**a**) and tumor resection (**b**), AKT (red) in CNB (**c**) and tumor resection (**d**), ER (red) in CNB (**e**) and tumor resection (**f**), and GAPDH (red) in biopsy (**g**) and tumor resection (**h**) was illustrated. Each corresponding cytokeratin staining is shown as inset (green). Photographs are shown at magnification of × 20.

First, differences in phospho-protein antigenicity between core needle biopsies and tumor resections were assessed paired with a second antibody that recognized the protein, independent of phosphorylation status. Tissue sections were stained with antibodies of pAKT, AKT, pERK, ERK, pS6K1 and S6K1. Then, each case pair was compared by taking the average of the AQUA scores over all FOVs in a specimen. The average number of FOV was 11 and 19 on each section of core needle biopsies and tumor resections, respectively. There were consistently and significantly lower levels of pAKT, pERK and pS6K1 ($P < 0.05$) in the tumor resections than in the core needle biopsies, as evaluated by Wilcoxon's Signed Ranks tests (Figure 3a, c and e). In contrast, the antigenicity of total AKT, ERK and S6K1 did not exhibit significant differences ($P > 0.05$) (Figure 3b, d and f). However, our power calculations indicate that we did not have a large enough sample size to detect differences for the total protein epitopes. Due to the large variation between subjects in the amount of change between core needle biopsy and tumor resection for these epitopes, a larger sample size would be required to sufficiently power the study. The ratio of the standard deviation of decrease between subjects relative to the mean decrease can be summarized numerically as the coefficient of variation. A summary of the data with *P*-values, experimental power, calculated required sample size requirements and coefficient of variation is shown in Table 3.

Assessment of Non-Phospho-Protein Epitope Loss in Paired whole Tissue Sections

To determine if antigenic liability can be observed in proteins other than phospho-epitopes, we selected some other commonly used antibodies to determine the effects of cold ischemic time on routine reagents. We tested ER, Ki67, GAPDH and cytokeratin antigenicity on paired tissue sections as above (Figure 4). To binarize ER into a positive and negative score, a cut-point was determined using a series of 37 breast cancer cases with conventional ER scores defined by the standard clinical ER test. When ranked in order, the highest AQUA score in a clinically negative test was 108 and the lowest AQUA score for a clinically positive test was 118 (unpublished data). Thus, ER was defined as positive for all cases with AQUA score > 118 on the paired whole tissue sections. Both ER and cytokeratin (AE1/AE3) showed statistically significant decreases in the tumor resection compared with the core needle biopsy. Neither Ki67 nor GAPDH showed significant reductions in the tumor resection specimens, but this may have been due to insufficient power. Ki67 especially showed a high coefficient of variation for epitope loss, likely due to inherent biologic variability, and thus over 500 subjects are required to adequately power the study (Table 3). Because cytokeratin antigenicity is measured in each AQUA protocol procedure, it was included in all the experiments. As a result, the number of FOVs assessed for this marker is much higher; however, much of the data is from serial sections, which are not truly independent FOVs.
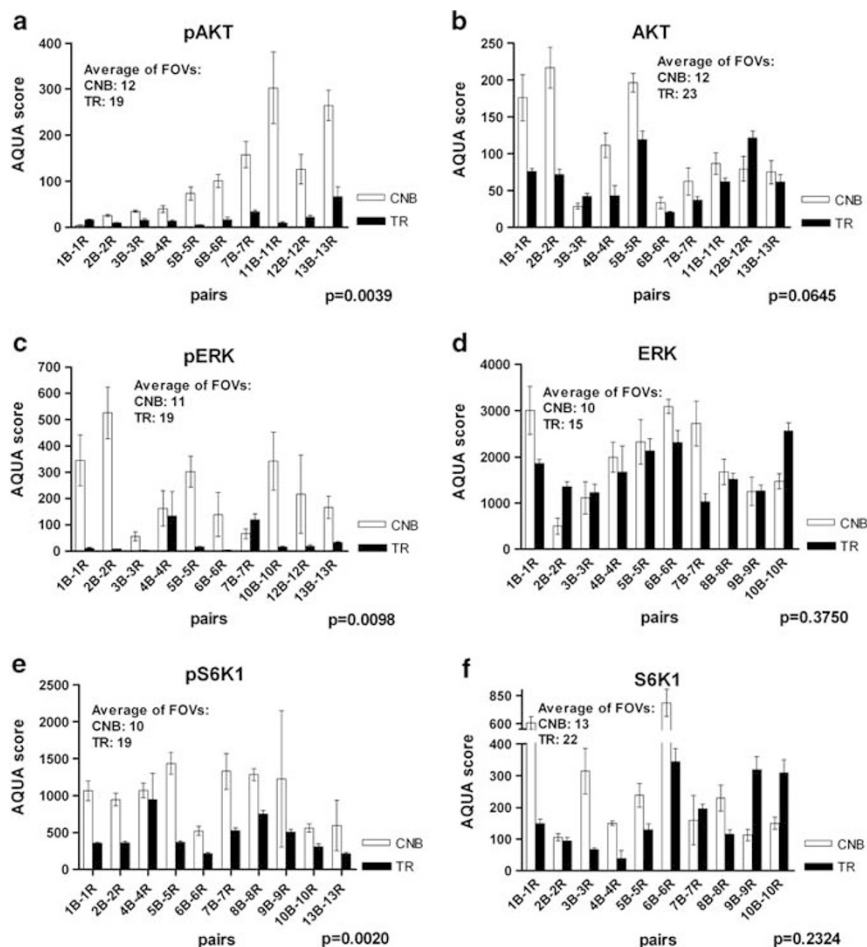
Figure 3 Expression of phospho-proteins including pAKT, pERK and pS6K1 were decreased in tumor resections compared with those in biopsies. AQUA scores of each coded biopsy (core needle biopsies, open bars) and tumor resection (tumor resection, filled bars) in staining of pAKT (a), AKT (b), pERK (c), ERK (d), pS6K1 (e), and S6K1 (f) were assessed. AQUA data represent the mean ± 95% CI. N represents the number of FOVs. Comparisons between core needle biopsies and tumor resections were performed via the Wilcoxon Signed Rank test.

Additionally, in order to standardize the scores between AQUA procedures, we scaled all CK scores to a 100-point scale. The pooled data for each pair from the nine experiments, with the average number of FOVs equal to 79 and 141 from each core needle biopsy and tumor resection, respectively, showed a statistically significant reduction in the tumor resections compared with the core needle biopsies.

Statistical Comparison Between Core Needle Biopsies and Tumor Resections with Experimental Power

After we established the statistical significance of the epitope loss between core needle biopsies and tumor resections, we used linear mixed-effects models to quantify those differences. Variation between individual subjects' magnitude of epitope loss was also significant, as demonstrated in the coefficient of variation for epitope loss (Table 3). Furthermore, loss of different epitopes was not uniform across individual subjects. For example, loss of signal from one epitope and gain of signal from another were seen in the

same individual (Supplementary Figure 1). This shows both variations as a function of tumor heterogeneity as well as variation in time to fixation.

DISCUSSION

As protein biomarkers are now being used to determine therapy, the importance of accurate measurement has increased. This has led to the observation that many events that occur before fixation of the tissue (pre-analytic variables) can be critically important. Even the best tests can only measure epitopes that are present on the slide, so the stability of these epitopes during the time between loss of circulation and fixation may be a key factor in their accurate measurement. Variables that can affect the stability of these epitopes include local pH changes, anoxic or ischemic biological reactions and loss of specificity of proteases or phosphatases. All these may result in alterations of post-translational modification and possible transcription of apoptotic factors beginning at the time the blood supply is clamped off.[5,25–28] Although this list
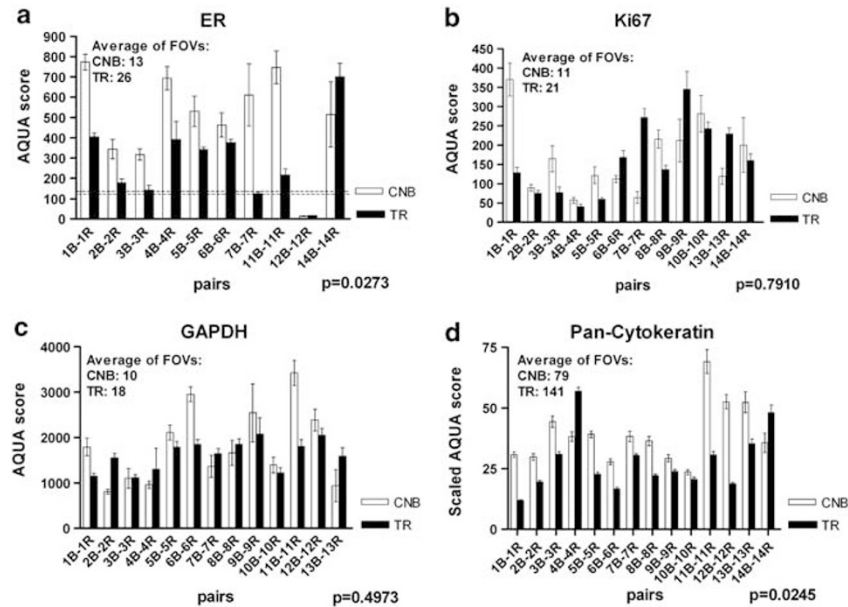
Figure 4 Epitope-specific protein degradation in tumor resections. AQUA scores of each coded biopsy (core needle biopsy, open bars) and tumor resection (tumor resection, filled bars) in staining of ER (**a**), Ki67 (**b**), GAPDH (**c**), and Pan-cytokeratin (**d**) were assessed. Raw Pan-cytokeratin scores were scaled to a 100-point scale. AQUA data represent the mean ± 95% CI. *n* represents the number of FOVs. Comparisons between core needle biopsies and tumor resections were performed via the Wilcoxon Signed Rank test.

of pre-analytic variables is incomplete, in this work we focus specifically on signaling protein phosphorylation. There is evidence that intrinsic enzymatic activity from slowly fixed endogenous phosphatase in large specimens results in the dephosphorylation of biomarkers.[29,30] The papers and findings reported here suggest that measurement of phospho-modification is essentially impossible in routinely collected surgical resection specimens.

In this work, we define core needle biopsies as a standard with the assumption that the procedure allows rapid fixation and prevents pre-analytic artifacts. While we believe this is the closest we can come to assessment of live tissue, we have no evidence that artifacts do not occur in the short time between obtaining the core and crosslinking by formalin. However, the fact that measurement of phospho-proteins spans a range of expression in our small case series suggests that at least some of the activity of signaling proteins can be assessed using this technique. This suggests that if phospho-specific markers are critical to companion diagnostic testing, surgical protocols should be altered to include a core needle biopsy of the lesion immediately before surgery for diagnostic testing purposes.

If core needle biopsies' assessment is considered to reveal the 'truth', it would be interesting to determine if some protein expression marker could be used to normalize or adjust for degradation due to pre-analytic variables. As cytokeratin staining was done on each specimen as part of the AQUA process, we tested whether normalization of the phospho-protein AQUA scores by the degradation seen in

cytokeratin in the tumor resections would more accurately reflect the levels of phospho-protein seen in the core needle biopsies. Our data indicated that AQUA scores of pAKT, pERK, pS6K1 and ER were still lower after pan-CK normalization (data not shown). While it may be possible in the future to find a method of normalization or qualification of tissue for companion diagnostic assay assessment, we believe it will not be possible to normalize for post-translational modifications. We are currently working on intrinsic controls for assessment of tissue quality.

Of the targets studied in this work, only ER is in clinical usage today. The current standard of practice for ER has recently been reviewed and an ASCO/CAP committee has issued standards that include issues related to cold ischemic time.[18] Our work supports the need for those standards as, shown here, the levels of ER in tumor resection specimens were significantly lower than in core needle biopsy specimens. It is common practice to assess ER and PR on core needle biopsy specimens; nonetheless, it is not required. Although none of the cases in this study changed from a designation of ER positive to a designation of ER negative between the core needle biopsy and the tumor resection, three cases were very close to negative in the tumor resection while clearly positive in the core needle biopsy (Figure 4a). In a broader collection of cases, it is highly likely we would see changes in patient ER status resulting in changes in endocrine therapy as a function of which specimen was assayed.

While this work shows quantitative loss of detection for a range of biomarkers, there are several limitations. Specifically,

in each setting, we sampled specimens only from 10 to 14 patients. While statistical significance is observed, it would be valuable to validate these results in other laboratories. Importantly, our power calculations demonstrate that, for every epitope without a significant difference between core needle biopsy and tumor resection, the sample size was insufficient. Thus, further investigation is needed to establish the affect of pre-analytical variables on measurement of non-phospho markers, such as Ki67. This data has led to funding that will assess larger cohorts in the future, including measured time to fixation, which is not available for these specimens. An additional weakness of this work, and essentially all work associated with assessment of core needle biopsies, is representation. In each case, we were able to quantify only a limited number of FOVs from each core needle biopsy. Furthermore, we have no way of determining whether a core needle biopsy was representative of an entire tumor. Our previous work suggests that there is substantial heterogeneity in ER expression when multiple tissue blocks are assessed in larger cases.[31,32] Heterogeneity may explain the observation that rare cases show higher average expression in the resection than in the core biopsy. For example, case 14 in Figure 4a shows higher average levels in the resection, even though all other cases trend in the opposite direction. Finally, another weakness of this work is the fact that that the core needle biopsy was not taken at the same time as the tumor resection. As this was a retrospective collection of tissue, we collected core needle biopsies that were taken as part of routine patient management and which were followed by resections. It is possible that assessment of tumors where the core needle biopsy and tumor resection are taken simultaneously would show different results. A further limitation is that while AQUA is now used in one clinical laboratory, it is not the common standard for assessment of antibody/antigen interactions. However, the immunofluorescence method is identical to DAB-based chromogenic studies in detection of antigen. That is, the primary antibody step is the same. The difference is in visualization, where the use of fluorescent visualization methods is easier to accurately quantify. Thus, it is important to emphasize that the conclusions we draw from this work are a function of the tissue assessed, not a function of the method of assessment. This is confirmed by work published, while this manuscript was under review, that showed qualitative assessment of loss of pAKT and pERK using chromogenic stains.[33] Nonetheless, work is underway toward the goal of developing quantitative methods that work for chromogenic visualization methods that can be used to assess this and other expression patterns in tissue.

In summary, this work quantitatively assesses the effect of delayed formalin fixation on surgical resections using biopsies as standardized controls. Although it is a limited set of experiments, all three phospho-antibodies tested showed significant degradation in the resection specimens. These results indicate that conventional resection tissues with uncontrolled cold ischemic time should not be used for companion diagnostic testing, at least for the five antibodies shown to be affected here. This work also suggests that any antibody planned for use on resection specimens should be validated for resistance to cold ischemic time.

Supplementary Information accompanies the paper on the Laboratory Investigation website (http://www.laboratoryinvestigation.org)

DISCLOSURE/CONFLICT OF INTEREST
Drs Rimm and Camp are founders, stockholders and consultants to HistoRx Inc.

1. Werner M, Chott A, Fabiano A, et al. Effect of formalin tissue fixation and processing on immunohistochemistry. Am J Surg Pathol 2000;24:1016–1019.
2. De Cecco L, Musella V, Veneroni S, et al. Impact of biospecimens handling on biomarker research in breast cancer. BMC Cancer 2009;9:409.
3. Micke P, Ohshima M, Tahmasebpoor S, et al. Biobanking of fresh frozen tissue: RNA is stable in nonfixed surgical specimens. Lab Invest 2006;86:202–211.
4. Start RD, Cross SS, Smith JH. Assessment of specimen fixation in a surgical pathology service. J Clin Pathol 1992;45:546–547.
5. Baker AF, Dragovich T, Ihle NT, et al. Stability of phosphoprotein as a biological marker of tumor signaling. Clin Cancer Res 2005;11:4338–4340.
6. Kirkegaard T, Witton CJ, McGlynn LM, et al. AKT activation predicts outcome in breast cancer patients treated with tamoxifen. J Pathol 2005;207:139–146.
7. Modi S, DiGiovanna MP, Lu Z, et al. Phosphorylated/activated HER2 as a marker of clinical resistance to single agent taxane chemotherapy for metastatic breast cancer. Cancer Invest 2005;23:483–487.
8. Cappuzzo F, Magrini E, Ceresoli GL, et al. Akt phosphorylation and gefitinib efficacy in patients with advanced non-small-cell lung cancer. J Natl Cancer Inst 2004;96:1133–1141.
9. Haas-Kogan DA, Prados MD, Lamborn KR, et al. Biomarkers to predict response to epidermal growth factor receptor inhibitors. Cell Cycle 2005;4:1369–1372.
10. Haas-Kogan DA, Prados MD, Tihan T, et al. Epidermal growth factor receptor, protein kinase B/Akt, and glioma response to erlotinib. J Natl Cancer Inst 2005;97:880–887.
11. Baselga J, Tripathy D, Mendelsohn J, et al. Phase II study of weekly intravenous recombinant humanized anti-p185HER2 monoclonal antibody in patients with HER2/neu-overexpressing metastatic breast cancer. J Clin Oncol 1996;14:737–744.
12. Svensson S, Jirstrom K, Ryden L, et al. ERK phosphorylation is linked to VEGFR2 expression and Ets-2 phosphorylation in breast cancer and is associated with tamoxifen treatment resistance and small tumours with good prognosis. Oncogene 2005;24:4370–4379.
13. Cappuzzo F, Finocchiaro G, Metro G, et al. Clinical experience with gefitinib: an update. Crit Rev Oncol Hematol 2006;58:31–45.
14. Yamashita H, Nishio M, Kobayashi S, et al. Phosphorylation of estrogen receptor alpha serine 167 is predictive of response to endocrine therapy and increases postrelapse survival in metastatic breast cancer. Breast Cancer Res 2005;7:R753–R764.
15. Mann GB, Fahey VD, Feleppa F, et al. Reliance on hormone receptor assays of surgical specimens may compromise outcome in patients with breast cancer. J Clin Oncol 2005;23:5148–5154.
16. Hodi Z, Chakrabarti J, Lee AH, et al. The reliability of assessment of oestrogen receptor expression on needle core biopsy specimens of invasive carcinomas of the breast. J Clin Pathol 2007;60:299–302.

17. Wolff AC, Hammond ME, Schwartz JN, *et al*. 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 2007;25:118–145.

18. Hammond ME, Hayes DF, Dowsett M, *et al*. American Society of Clinical Oncology/College of American Pathologists Guideline Recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. J Clin Oncol 2010;28: 2784–2795.

19. Camp RL, Chung GG, Rimm DL. Automated subcellular localization and quantification of protein expression in tissue microarrays. Nat Med 2002;8:1323–1327.

20. Moeder CB, Giltnane JM, Harigopal M, *et al*. Quantitative justification of the change from 10% to 30% for human epidermal growth factor receptor 2 scoring in the American Society of Clinical Oncology/College of American Pathologists guidelines: tumor heterogeneity in breast cancer and its implications for tissue microarray based assessment of outcome. J Clin Oncol 2007;25: 5418–5425.

21. Gustavson MD, Molinaro AM, Tedeschi G, *et al*. AQUA analysis of thymidylate synthase reveals localization to be a key prognostic biomarker in 2 large cohorts of colorectal carcinoma. Arch Pathol Lab Med 2008;132:1746–1752.

22. Pinheiro JCP, Bates DM. Mixed-Effects Models in S and S-PLUS. Springer: Berlin, 2000.

23. R_Core_Development_Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing: Vienna, Austria, 2009.

24. Pinheiro JC, Bates D, DebRoy S, *et al*. Linear and Nonlinear Mixed Effects Models. R package version 3.1-96 edn; 2009.

25. Kingsbury AE, Foster OJ, Nisbet AP, *et al*. Tissue pH as an indicator of mRNA preservation in human post-mortem brain. Brain Res Mol Brain Res 1995;28:311–318.

26. Cross SS, Start RD, Smith JH. Does delay in fixation affect the number of mitotic figures in processed tissue? J Clin Pathol 1990;43:597–599.

27. Srinivasan M, Sedmak D, Jewell S. Effect of fixatives and tissue processing on the content and integrity of nucleic acids. Am J Pathol 2002;161:1961–1971.

28. Allen J, Winterford C, Axelsen RA, *et al*. Effects of hypoxia on morphological and biochemical characteristics of renal epithelial cell and tubule cultures. Ren Fail 1992;14:453–460.

29. Espina V, Edmiston KH, Heiby M, *et al*. A portrait of tissue phosphoprotein stability in the clinical tissue procurement process. Mol Cell Proteomics 2008;7:1998–2018.

30. Lin TA, Lin TN, He YY, *et al*. Effects of focal cerebral ischemia on inositol 1,4,5-trisphosphate 3-kinase and 5-phosphatase activities in rat cortex. Biochem Biophys Res Commun 1992;184:871–877.

31. Chung GG, Zerkowski MP, Ghosh S, *et al*. Quantitative analysis of estrogen receptor heterogeneity in breast cancer. Lab Invest 2007;87: 662–669.

32. Chen C, Peng J, Xia H, *et al*. Quantum-dot-based immunofluorescent imaging of HER2 and ER provides new insights into breast cancer heterogeneity. Nanotechnology 2010;21:95101.

33. Pinhel IF, Macneill FA, Hills MJ, *et al*. Extreme loss of immunoreactive p-Akt and p-Erk1/2 during routine fixation of primary breast cancer. Breast Cancer Res 2010;12:R76.