

Quantitation of fixative-induced morphologic and antigenic variation in mouse and human breast cancers

Robert D Cardiff^{1,2}, Neil E Hubbard¹, Jesse A Engelberg¹, Robert J Munn^{1,2}, Claramae H Miller^{1,2}, Judith E Walls¹, Jane Q Chen¹, Héctor A Velásquez-García¹, Jose J Galvez³, Katie J Bell¹, Laurel A Beckett⁴, Yue-Ju Li⁴ and Alexander D Borowsky^{1,2}

Quantitative Image Analysis (QIA) of digitized whole slide images for morphometric parameters and immunohistochemistry of breast cancer antigens was used to evaluate the technical reproducibility, biological variability, and intratumoral heterogeneity in three transplantable mouse mammary tumor models of human breast cancer. The relative preservation of structure and immunogenicity of the three mouse models and three human breast cancers was also compared when fixed with representatives of four distinct classes of fixatives. The three mouse mammary tumor cell models were an ER + /PR + model (SSM2), a Her2 + model (NDL), and a triple negative model (MET1). The four breast cancer antigens were ER, PR, Her2, and Ki67. The fixatives included examples of (1) strong cross-linkers, (2) weak cross-linkers, (3) coagulants, and (4) combination fixatives. Each parameter was quantitatively analyzed using modified Aperio Technologies ImageScope algorithms. Careful pre-analytical adjustments to the algorithms were required to provide accurate results. The QIA permitted rigorous statistical analysis of results and grading by rank order. The analyses suggested excellent technical reproducibility and confirmed biological heterogeneity within each tumor. The strong cross-linker fixatives, such as formalin, consistently ranked higher than weak cross-linker, coagulant and combination fixatives in both the morphometric and immunohistochemical parameters.

Laboratory Investigation (2013) 93, 480–497; doi:10.1038/labinvest.2013.10; published online 11 February 2013

KEYWORDS: estrogen receptor; Her2; IHC4 biomarker panel; immunohistochemistry; MIB1 (Ki67); progesterone receptor; quantitative image analysis

Histopathology is in an era of cooperative studies including immunohistochemical and molecular analyses that require increased use of standard operating procedures (SOPs) and quantitative scrutiny.¹ Before applying this new technology to cancer research, practicing pathologists need to have suitable measurements of technical and biological variation that affect the analysis. In an empirically heterogeneous tissue such as neoplasia, one must establish a basis for comparing one set of observations with another. Using relatively uniform tumor models can provide the baseline data.

This project was also stimulated by the increased interest in various ‘universal fixatives’ that are designed to optimize the preservation of the molecular data yet preserve the pathology, and by the desire to reduce the use of formalin.² Practicing pathologists are now confronted with the bewildering literature containing interesting observations

on new fixatives but lacking well-controlled comparisons to validate the various claims.³ This era has also seen an abundance of new fixatives coming to market.^{2–29} The new and old fixatives have been reviewed in a 2009 report from the National Health Services Purchasing and Supply Agency.³ With the multiple variables involved, it is difficult to compare one study with another. Since any significant changes in SOP could interfere with workflow, instituting change must be justified by evidence that supports the new procedures.

The documentation for each fixative has been largely based on appropriate molecular data coupled with assessment by one or more ‘experienced’ or ‘expert’ pathologists.^{17,26} Most studies involve the use of human tissues with unknown and poorly documented characteristics. Moreover, relatively few studies involve rigorous comparisons between formalin and other fixatives.³ This rather interesting approach, in

¹Center for Comparative Medicine, University of California, Davis, CA, USA; ²Department of Pathology and Laboratory Medicine, School of Medicine, University of California, Davis, CA, USA; ³Center for Biomedical Informatics and Information Technology, National Cancer Institute, National Institute of Health, Washington, DC, USA and ⁴Division of Biostatistics, Department of Public Health Sciences, School of Medicine, University of California, Davis, CA, USA
Correspondence: Dr AD Borowsky, MD, Center for Comparative Medicine, University of California Davis, County Road 98, Hutchison Drive, Davis, CA 95616, USA.
E-mail: adborowsky@ucdavis.edu

Received 10 September 2012; revised 14 January 2013; accepted 15 January 2013

our estimation, ignores many technical and biological issues confronting the pathologist. One rarely sees measures of technical reproducibility and biological variation or acknowledges tumor heterogeneity in the context of immunohistochemistry (IHC) or fixation. Thus, the reader has little or no measure of reproducibility or variability in most studies. Therefore, an experimental design that includes recognition of reproducibility and variation was developed.

Representatives of the four major fixative classes were used: (1) strong cross-linkers (formaldehydes); (2) weak cross-linkers (glyoxals); (3) coagulants (alcohols); and (4) combinations of cross-linkers and coagulants (Table 1). At least one reagent from each of the four major fixative classes was analyzed in every experiment. In most cases, the fixative reagents were purchased commercially, and the choice of fixatives was primarily driven by availability on the retail market so that readers could obtain identical reagents, if desired. An IHC panel was performed for identification of four principal breast cancer-related antigens: Her2/ErbB2, ER, PR, and Ki67 (IHC4 biomarker panel).³⁰

All slides were scanned to generate digital whole slide images (WSI).

Quantitative image analysis (QIA) was used to provide more precise data, which should be reproducible in independent laboratories. The initial steps involved evaluation of the technology with validation of the technical and biological reproducibility and variance. The generation of quantitative data permitted statistical analysis of the morphometric parameters and IHC staining. Rank orders for each parameter were developed based on the numerical data. The observations using the mouse mammary tumors were compared, confirmed, and validated using three human breast cancers that were available for comparable fixation and processing. It was the aim of the study to determine the differences in morphology and immunophenotyping assays resulting from different fixatives. To provide this comparison, additional parameters (eg, antigen retrieval) were held constant in all experiments even though it might be possible that some fixatives would benefit from tailored optimization. We found statistically significant differences in an array of

Table 1 Fixatives used in the study

Fixative name	Designation	Mode of action	Other known constituents	Source	Comments
Neutral-buffered formalin	NBF	Strong cross-linking (formaldehyde)	Ethylene glycol, sodium acetate, methanol (1–1.5%)	American Master Tech, Lodi, CA	Buffered formalin concentrate solution diluted 1:4 in water
Z-Fix	Z-Fix	Strong cross-linking (formaldehyde)	Buffer, ionized zinc	Anatech Ltd, Battlecreek, MI	'Aqueous buffered formalin'
HistoChoice	HChoice	Weak cross-linking (glyoxal)	Glyoxal, sodium chloride, butanedial and zinc sulfate (Titford and Horenstein ²⁶)	Electron Microscopy Sciences, Hatfield, PA	
Mirsky's fixative	Mirsky	Weak cross-linking (glyoxylate)	Dipotassium phosphate	National Diagnostics, Atlanta, GA	Ingredients proprietary
Boonfix b.v.	Boon's	Coagulation (EtOH)	Acetic acid, PEG300	Denteck, Zoetermeer, The Netherlands	Ingredients proprietary
FineFIX	Fine	Coagulation (EtOH)	Distilled water, glycerol, polyvinyl alcohol, and monomeric carbohydrates (Stanta <i>et al</i> ²⁵)	Milestone Medical, Shelton, CT	Ingredients proprietary
Methacarn	MCarn	Coagulation (MeOH)	Acetic acid, chloroform	Advanced Biomedical Reagents and Technologies, Stanwood, WA	
PAXgene	PAX	Coagulation (MeOH)	Acetic acid	PreAnalytix, a Qiagen/BD Company	Ingredients proprietary
Z7 fixative	Zn7	Coagulation (zinc salts)			Used according to protocol described (Lykidis <i>et al</i> ¹⁵)
Tellysniczky/Fekete fixative	Telly's	Coagulation, cross-linking (EtOH, formaldehyde)	Acetic acid		Prepared in-house monthly using 100 ml 70% ethanol, 5 ml glacial acetic acid and 10 ml 37% formaldehyde
Prefer fixative	Prefer	Coagulation, cross-linking (EtOH, glyoxal)	Buffer	Anatech Ltd, Battlecreek, MI	Ingredients proprietary

measurements, and present the detailed and summarized findings.

MATERIALS AND METHODS

Tissue Sources

Mice

Mice were housed in a vivarium under NIH guidelines and all animal experiments followed protocols approved by the UC Davis Institutional Animal Care and Use Committee (IACUC). Tissues were sampled from two mouse strains, FVB/NJ (Charles River, Wilmington, MA) and 129S6/SvEv (Taconic, Hudson, NY). The mice were transplanted with three types of mammary tumors as described. The mouse mammary tumor cell lines were previously grown in culture and prepared for transplantation into the #2, #3, and/or #4 uncleared mouse mammary fat pads. A total of 22 mouse experiments were performed over a period of 2 years.

Cell lines

Three transplantable murine mammary tumor cell lines were used (Table 2). The SSM2 cell line was developed and provided courtesy of Drs. Robert Schreiber and Szeman Chan (Washington University School of Medicine, St Louis, MO).³¹ This cell line was developed from a mouse lacking STAT1 activity (129S6/SvEv-*Stat1^{tm1Rds}*) and had spontaneously developed an adenocarcinoma in the mammary gland. SSM2 cells were injected into 129S6/SvEv female mice. The Met-1 tumor cell line was developed in our laboratory.³² The NDL-1 cell line was also developed in our laboratory from a tumor that arose in a transgenic mouse with an active *neu* gene expressed in the mammary gland.³³ All cell lines were stored in liquid nitrogen until use and were of relatively low passage numbers anywhere from 12 to 27. Cells were grown in the appropriate culture media supplemented with 10% fetal calf serum.³² Cultures that reached at least 70% confluence were trypsinized, washed three times with PBS and counted.

Transplantation

A bolus of cells (1×10^5 to 2×10^6) was injected bilaterally into the uncleared #2, #3, and #4 fat pads of 6- to 8-week-

old mice. For the SSM2 cell line, the recipient mouse strain was 129SvEv and for the NDL and Met-1 cell lines, the recipient mouse strain was FVB/NJ. Following transplantation, mice were palpated twice weekly to monitor tumor take, which equaled or exceeded 90%. The tumors were usually palpable within 10–14 days after transplantation. The tumors were allowed to grow to 5–10 mm in greatest dimension. In our experience, tumors larger than 10 mm are frequently necrotic, resulting in fewer viable tumor cells.

Tissue harvest

To remove the tissues, mice were first anesthetized using Nembutal (60 mg/kg) and selected tissues removed from live, anesthetized animals in a specific order. Following organ removal, the mice were euthanized using an overdose of Nembutal (120 mg/kg). In experiments where multiple tissues were harvested, tumors were removed first to significantly decrease trauma and blood loss. Subsequently, the uterus was removed and then the liver. All tissues were immediately cut into 1–2 mm slices using a razor blade or scalpel and placed immediately into the experimental fixative or frozen in liquid nitrogen (see Figure 1 for experimental flow).

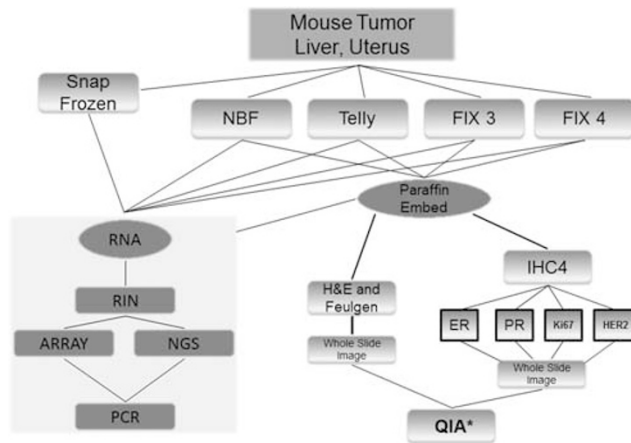


Figure 1 Work flow schematic: schematic of work flow from tissue to image analysis.

Table 2 Description of cell lines and tumors

Tumor name	IHC staining (+ / -)				Host mouse strain	Source/Characteristics
	ER	PR	Her2	Ki67		
Met-1	-	-	-	+	FVB/N	Derived from tumor of FVB/N-Tg(MMTV-PyVmT)634Mul/J female mouse
SSM2	+	+	-	+	129S6/SvEv	Derived from spontaneous mammary tumor of 129S6/SvEv- <i>Stat1^{tm1Rds}</i> female mouse
NDL-1	-	-	+	+	FVB/N	Derived from mammary tumor of Tg(<i>Neu^{nd1}</i>) female mouse
Human 1	-	-	-	+	N/A	Invasive ductal carcinoma
Human 2	+	+	+	+	N/A	Invasive micropapillary carcinoma
Human 3	-	-	-	+	N/A	Invasive ductal carcinoma

Specimen triage

In total, 11 Met-1 tumors, 13 SSM2 tumors, and 9 NDL tumors were processed for these experiments (Figure 1). Each tumor was sliced into five to ten 1 mm slices depending on its size. The 1–2 mm slices from each tumor were divided between flash frozen (3–5 slices) and fixed in formalin (1 slice), Telly's (1 slice), and at least two additional fixatives of other fixative types (1 slice each). In some cases, the tumors were large enough to provide sample slices for up to four additional fixatives. The number and type of additional fixatives used for any given experiment were dependent on reagent availability and size of the tumor. The sample distribution is presented in a workflow diagram (Figure 1). The time for completion of these procedures did not exceed 5 min per mouse.

Human tissue sampling

Evaluation of various fixatives was also performed on three human breast cancer samples. Surgically removed samples were placed in PBS, chilled on ice, and transported from the UC Davis Medical Center (Sacramento, CA) surgical suite to the UC Davis Medical Center Pathology Department (Sacramento, CA). Samples were then cut into small, 1–2 mm pieces by a breast pathologist (ADB) and placed in the appropriate fixatives. Time lapse from surgery to initiation of fixation varied but was never longer than 45 min. All samples were collected under a protocol approved by the UCDMC Institutional Review Board.

Tissue Fixation and Processing

The 1–2 mm tissue slices were quickly placed in one of the fixatives described in Table 1. The resulting ratio was ~1 part tissue to 50 parts fixative. In addition, 3–5 tissue pieces 2 mm thick were placed in separate plastic freezer vials and immediately flash-frozen in liquid nitrogen, followed by storage at -80°C . According to manufacturer's specification, tissues fixed in PAXGENE (Qiagen, Valencia, CA) were placed in 'fixative solution 1' for 1 h followed by placement in 'fixative solution 2'. Tissues in all other fixatives remained at room temperature for 18 h. Table 1 describes the different fixatives used in the study.

Tissue embedding, H&E and Feulgen staining

The Tissue-Tek VIP autoproccessor (Sakura, Torrance, CA) was used to process tissues which were then embedded in Paraplast paraffin (melting temperature $56\text{--}60^{\circ}\text{C}$), sectioned to $5\ \mu\text{m}$ and mounted on glass slides. In all experiments, sections were stained using Mayer's hematoxylin and eosin to facilitate the histology and morphology evaluation. Other sections from the same blocks were stained with a Feulgen stain kit (American MasterTech, Lodi, CA), using the manufacturer's suggestions to enhance the nucleus, thus increasing accuracy for quantitating nuclear size and nuclear/cytoplasmic (n/c) ratio (see Figure 2a).

Immunohistochemistry

The antibodies used for this project are listed in Table 3. Because this project was primarily based on analysis of mouse mammary tumors, the standard breast cancer prognostic panel of ER, PR, Ki67, and Her2³⁰ for breast cancer was emphasized. The antibody utility in IHC was validated using both human and mouse tissues. Most of the standard antibodies used in the University of California, Davis Medical Center's clinical diagnostic laboratory could be used for the murine studies. SP1 was used for the IHC biomarker panel on the human breast cancers. However, the SP1 (anti-ER) antibodies used for standard clinical diagnostic studies did not provide a satisfactory stain for mouse ER. Therefore, SC-542 (Santa Cruz Biotechnology, Santa Cruz, CA) was substituted for SP1 for use on mouse tissues. However, SC-542 stained the cytoplasm of a small subset of neoplastic mouse cells. The regions with cytoplasmic staining were eliminated from the QIA. In addition, for anti-ErbB2 (Her2), we tested and preferred a rabbit monoclonal antibody (Lab Vision RM-2112-S) over the clinical lab's rabbit polyclonal antibody since the monoclonal reduced the non-specific background staining in mouse tissues.

All IHC was performed manually without the use of automated immunostainers. The same antigen retrieval method was used for all IHC and was performed using a Decloaking Chamber (Biocare Medical, Concord, CA) with 10 mM citrate buffer at pH 6.0, 125°C and pressure to 15 p.s.i. The total time slides were in the chamber was 45 min. Two other antigen retrieval methods were tested with a small group of samples to compare with the citrate decloaker method. For comparison for ER antigen retrieval, we tested an additional waterbath method with citrate buffer and a decloaker method using a Tris-EDTA buffer. The results of that comparison and the additional antigen retrieval methods are provided (Supplementary Table S1). An antigen retrieval comparison was also done for Her2 between the citrate decloaker method and water bath method. The results of that comparison are provided (Supplementary Table S2). Incubations with primary antibodies were performed at room temperature overnight in a humidified chamber. Normal goat serum was used for blocking. Biotinylated goat anti-rabbit (1:1000) was the secondary antibody used with a Vectastain ABC Kit Elite and a Peroxidase Substrate Kit DAB (Vector Labs, Burlingame, CA) used for amplification and visualization of signal, respectively. Tissues known to contain each assessed antigen were used as positive controls. In some instances, the tissue used in the study (eg, NDL tumor for Her2 staining) was the best positive control. Antibody deletion controls were used for every assessed antigen to confirm specific staining.

QIA and Data Acquisition

The quantitative analysis is based on slides from 51 tumors in 18 mice from 12 core experiments (see Supplementary Table S3 for details of tumor and animal distribution). All slides

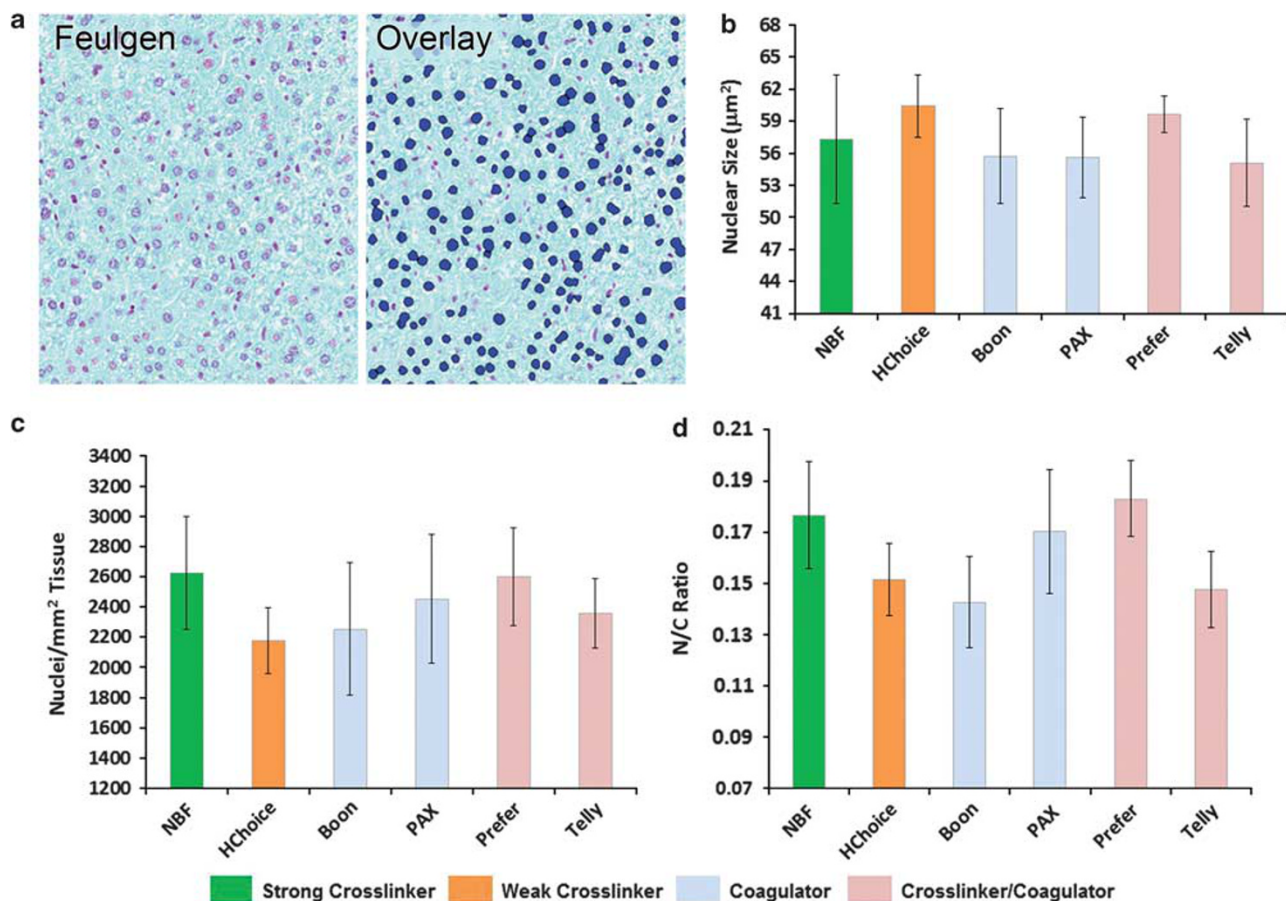


Figure 2 QIA of Feulgen-stained liver. (a) Images (Feulgen) were analyzed as described in Materials and methods. The customized algorithm identified each hepatocyte nucleus in blue (Overlay) and calculated its size. Scale bar length is 50 μm . At least one fixative from each of the fixative classes was assessed for nuclear size (b), nuclei per unit area (c), and the ratio of nucleus to cytoplasm (d). Fixative classes are identified by color.

were scanned and digitized using the Aperio ScanScope XT to capture digital WSI using the $\times 20$ objective lens at 0.5 μm /pixel and stored in the Aperio Spectrum version 10 customized for laboratory workflow. QIA was performed using Spectrum version 10 and version 11, based on the FDA-approved algorithms supplied by the manufacturer with modifications as described.

The digital WSI were analyzed using Aperio ImageScope software (<http://www.aperio.com/pathology-services/image-scope-slide-viewing-software.asp>). The data for each slide were automatically stored in the Aperio Spectrum database. The data were analyzed using Excel and the R environment for statistical computing.

Pre-analytical Annotation

Tumor morphology

The mouse tumor types used in the study are described in Table 2. These tumors were selected based on their resemblance to human breast cancer phenotypes and their stability when serially transplanted in host mice. Although the morphology and antigenicity of each transplant line mostly

remained consistent, variant areas and variant transplants were identified morphologically and eliminated from the study to reduce heterogeneity as much as possible. For example, the SSM2 from 129S6/SvEv-*Stat1*^{tm1Rds31} and Met-1 cells from Tg(PyVmT)³² had a tendency to undergo epithelial-to-mesenchymal transition (EMT) in some transplant generations and these samples were eliminated from the study.

Quantitative image analysis

The manufacturer's (Aperio Technologies, Inc.) FDA-approved algorithms were initially used to quantify nuclear and membrane staining.^{34–36} Each sample was re-examined with and without the overlay containing the morphometric images to verify its accuracy. The manufacturer's product manual emphasizes that the accuracy of the algorithms is highly dependent upon the selection of the area for analysis on the WSI (annotation) and should be performed by experienced professionals. In some instances, customization of algorithm parameters was required. For example, areas of necrosis and fibrosis were excluded from tumor morphometric and IHC annotations. In another example,

the original nuclear algorithms did not identify all hepatocytic nuclei using hematoxylin staining. Feulgen stain for DNA was the most consistent nuclear stain, but the IHC nuclear algorithm required customization to exclude cells other than hepatocytes (Figure 2a). Mouse uterine myometrium (outer layer) was used as a positive control for ER and PR staining since that layer is constant through the estrous cycle.³⁷

False positives and false negatives

Pre-analytical surveys were performed to detect tissues, fixatives, and stains that produced false positives or negatives. For example, fixatives can affect the intensity of the counterstain. When samples were processed in the same batch, some fixatives varied in hematoxylin stain intensity. Light nuclear counterstaining resulted in significant 'non-detection' of visible nuclei by the algorithm, decreasing the total cell count. These regions had falsely higher 'percentages' of positive cells. Therefore, the total 'positive cell density' was calculated and used as the quantitative measure. In most cases, the rank orders were usually the same when both percentages and positive densities were used. However, a discrepancy between percent positive cells and total positive cell density stood out when the weak cross-linker fixatives were used and were particularly notable with the Ki67-stained slides. Thus, the total positive cell density was uniformly used to compare the fixatives.

Quantitation of tissue shrinkage in liver

The IHC nuclear algorithm developed by Aperio in Image-Scope was customized to limit detection to hepatocytes in Feulgen-stained liver tissue sections. Briefly, because a

Feulgen staining procedure was used, optical density (OD) parameters were adjusted. In addition, since hepatocyte cell nuclei are relatively large, the minimal accepted nuclear size was adjusted. Hepatocyte nuclei are mostly round, compact, and visually well-defined so the roundness and compactness setting were adjusted. The nuclear/cytoplasmic ratio was calculated by subtracting the total nuclear area from the total area in a selected region and dividing the nuclear area by the total area minus the nuclear area.

Quantitation of tissue shrinkage in tumor

The default settings of the Aperio Positive Pixel Count (PPC) algorithm quantify the amount of stain present in the area of analysis. Pixels that are stained, but do not fall into the positive-color range, are considered as negative stained. Since tissue shrinkage (especially in tumor) leaves empty-space artifacts, the decreased tissue area could be quantified by calculating the ratio of empty area *vs* filled area (regions where pixels are detected). Accordingly, the PPC algorithm was customized to report non-empty areas as 'strong pixels' (red on the overlay) and the empty areas as 'weak pixels' (yellow on the overlay, see Figure 3). The shrinkage ratio is then calculated by dividing the number of weak pixels by the total pixels. Several areas within each tumor section were annotated and analyzed for the degree of shrinkage. For each section, a mean shrinkage ratio was calculated and compared with other tumors among the fixatives.

QIA of ER, PR, Her2, and Ki67 staining

Stained, scanned slides were analyzed using the Aperio 'IHC Membrane' and 'IHC Nuclear' algorithms except as noted below. The areas for analysis of each were selected by

Table 3 Antibodies used for immunohistochemistry

Antibody	Source	Catalog/Lot number	Dilution used
CK5	Abcam Cambridge, MA	ab24647/522866	1:1000
CK8/18	Fitzgerald, Acton, MA	20R-CP004/P11022326	1:4000
CK14	Lab Vision Kalamazoo, MI	RB-9020-P1/9020P10068	1:1000
CK19	Epitomics Burlingame, CA	3863/YH122404C	1:1000
E-cadherin	Abcam	ab53033/827763	1:1200
N-cadherin	Epitomics	2019-1/YE011803R	1:800
ER mouse	Santa Cruz, Santa Cruz, CA	SC-542/F0110	1:800
ER (Sp1), human	Thermo Sci. Kalamazoo, MI	RM-9109-50/9101151003C	1:100
Her2, monoclonal	Lab Vision	RM-2112-S	1:1000
Her2, polyclonal	Lab Vision	RB-103-PIABX/103X504A	1:1200
Ki67 mouse, human	Lab Vision	RB-1510-PO/1510P803C	1:800
pAKT (ser473)	Cell Signaling Danvers, MA	4060S/#9	1:2000
PR mouse/human	Dako Carpinteria, CA	A0098/00057857	1:500
Vimentin	Epitomics	2707/1EPR3776	1:1000

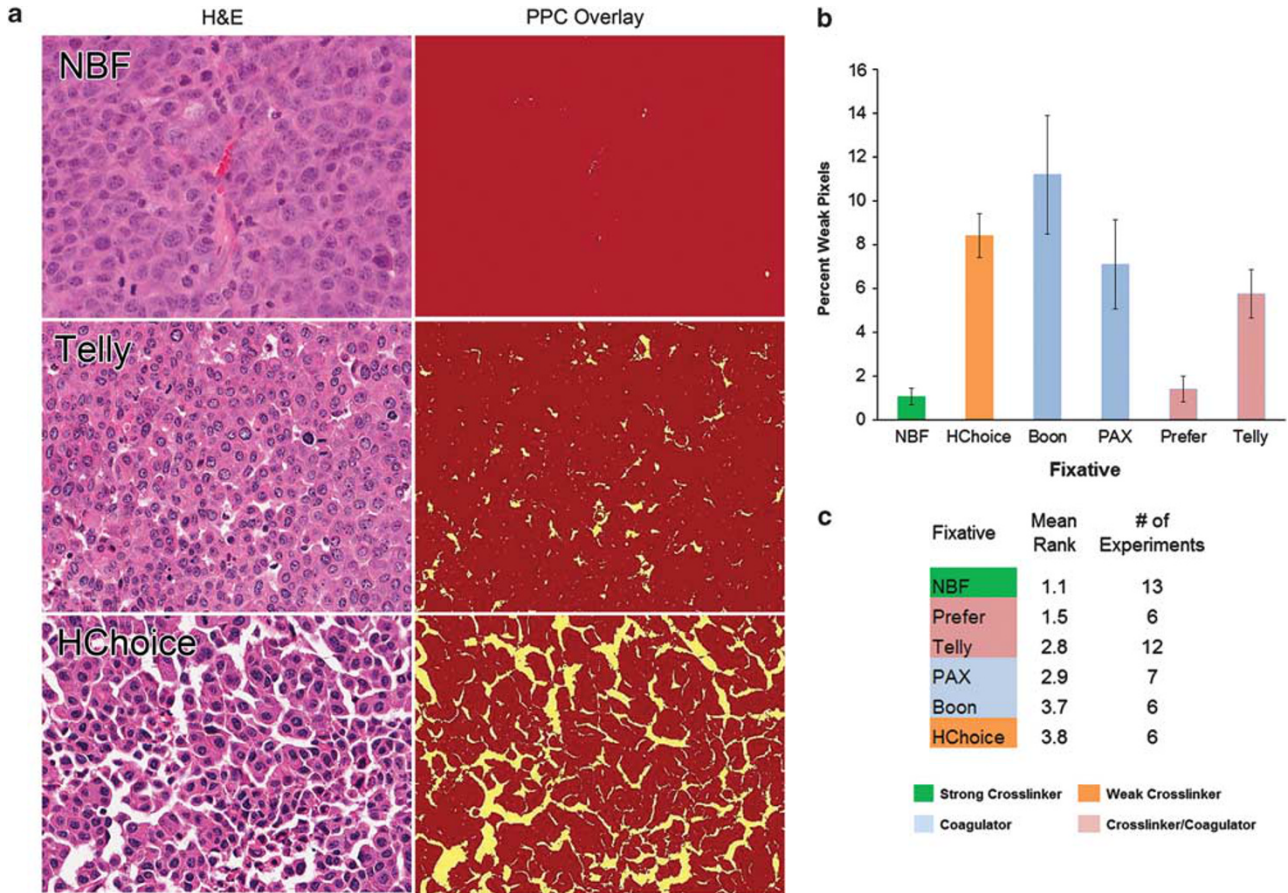


Figure 3 Fixative effect on tumor morphology. In H&E stains (a), spaces are notable but not quantifiable. Using a customized PPC algorithm, weak pixels (yellow in PPC Overlay) representing shrinkage artifact can be quantified. Scale bar length is 50 μm . Overall, significant differences among the fixatives for all experiments existed (b). The mean rank order (c) shows the distribution of fixatives with class designation. The *n* value represents the number of experiments for each of the fixatives used in the tumor shrinkage study.

experienced morphologists and validated by two pathologists (AB and RC).

Algorithm modifications

The outer layer of the mouse uterine myometrium was used for analysis of ER and PR staining since it remains constant through the various stages of the mouse estrous cycle.³⁷ At least five separate areas were selected with a modification of the Aperio ‘IHC Nuclear’ algorithm to quantitate the number and intensity of stained nuclei. To compensate for granular Ki67 staining within the nucleus of human cancer tissue, the IHC nuclear algorithm parameters were modified. The averaging radius, which controls the amount of blurring between pixels, was increased from 1 to 1.7 μm , while the ‘threshold type’ parameter was changed from ‘edge rejection’ to ‘manual’. The ‘lower Intensity threshold’ was set at 30 and the ‘upper intensity threshold’ was set at 190. This modified algorithm was used only for human tissue stained with Ki67.

Density:intensity graphs

Raw data obtained from image analysis include calculations of percentages and absolute numbers of positive and negative

cells, and the area scanned (Figures 4 and c). These data indicate the number of events recorded and can be used to assess the magnitude of differences between fixatives. The data are then normalized by conversion into density of positive cells per unit area for each fixative. The relative values were then ranked from 1 to 4 for comparison with other tumors. For more accurate representation of the number of positive stained nuclei, a novel density:intensity graph was developed (Figures 4b and d). After the algorithm was executed, the data were exported and analyzed using the R environment for statistical computing. The data represented the frequency of cells in the annotation with a given intensity in optical intensity units (OIU), binned in 240 steps from 0.0 (lightest) to 1.0 (darkest). Graphs of this data give a more detailed view of the intensity data than percent positive nuclei. To compare data from multiple annotations on the same axis, a line plot was used instead of a column-based graph. Annotations covering a larger area of tissue could not be directly compared with smaller annotations, so the data were normalized by dividing by the size of the annotation, yielding density in cells/mm². The cells to the right of the intensity cutoff of 0.125 are equivalent to the cells

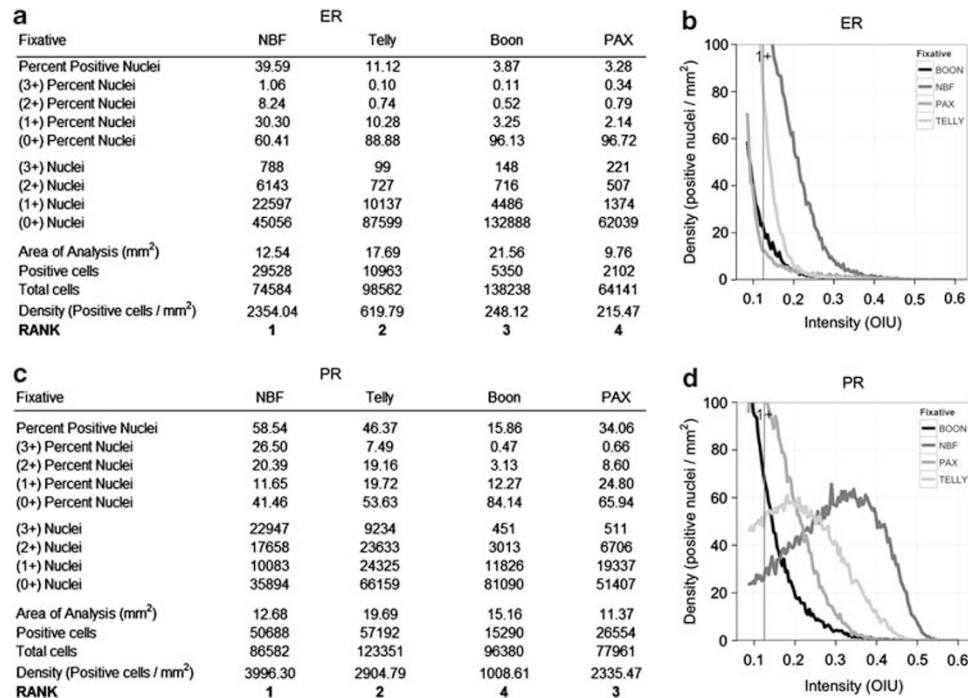


Figure 4 Tabular and graph representation of data: Conversion of tabular image analysis data of ER and PR from a single representative ER/PR-positive SSM2 tumor into density:intensity graphs. The raw data obtained from the image analysis include calculations of percentages and absolute numbers of positive and negative tumors cells, and the area scanned (a, c). The raw data indicate the number of events recorded and the magnitude of differences between fixatives. The data are then normalized by conversion into density of positive cells per unit area for each fixative. Graphs provide visual representation of the effects of the four fixatives on IHC for ER and PR (b, d).

categorized as ‘positive’ when calculating the percentage of positive cells, giving a measure of ‘positive cell density’. To compare the effects of different fixatives on staining intensity, the mean and standard deviation of the graph plots were calculated across tumors for each fixative in each experiment. In Figure 4, the density:intensity graphs provide a visual representation of the effects of four fixatives on ER and PR. Further, statistical analysis of all tumors and all stains using least square of means and Tukey–Kramer tables are also available (Supplementary Data).

Statistical Methods

Descriptive summaries of each outcome measure (mean, standard deviation, range, quartiles) were prepared to assess whether transformation would be necessary before fitting linear models. In some cases, log transformation was necessary to stabilize variances and reduce impact of long tails (outliers). For each outcome measure, fixatives were compared using analysis of variance treating experimental condition as a fixed effect that could modify the underlying mean value of the outcome measure. For each outcome measure, an overall assessment of the differences between fixatives was provided by the F test for fixative in the ANOVA table. Comparisons between individual pairs of fixatives were then carried out using Tukey’s studentized range procedure

(See Supplementary Tables S4–S15). Graphical summaries of comparisons used the least-squares.

Mean estimate rank

To compensate for well-documented biological heterogeneity of tumors, the data were also analyzed by rank.³⁸ The ordinal rank order for each value was calculated by assigning the maximum value in a given experimental group as 1. The other values for that parameter were then ranked in numerical order 1 through 4. In experiments with >4 fixatives the rank order was capped so that any fixative with a rank lower than 4 was set to 4. The use of fractional ordinates allowed direct comparison of experiments with four fixatives with experimental groups with five or more fixatives.

For any given group of experiments, the rank orders are represented as mean ranks, weighted mean rank, and/or mean class rank. For mean rank, the values for the given fixative for all tumors were summed and compared with all other sums to establish rank order. This, in essence, gives the rank-by-experiment. The ‘weighting’ was done by first establishing the rank order of the fixatives within each individual tumor and then summing these ranks for each fixative and dividing by the number of tumors to find the mean rank over all tumors in all experiments involving that fixative. This, in essence, gives rank-by-tumor. The fractional averages reflect the overall average. The mean class rank was

established by summing the rank within a class of fixatives and dividing by the number of fixatives in that class.

RESULTS

Faced with numerous variables affecting preservation and preparation of diagnostic samples,¹⁰ these studies used highly controlled SOPs and standard tissues. The evaluation of results utilized computer-based QIA rather than subjective evaluations by pathologists. Three transplantable mouse mammary tumor cell lines were used to measure biological variability with rigorous control over fixation and processing variables such as tissue characteristics, sample size, and time from excision to fixation. Next, a standard immunohistochemical panel for breast cancer was assessed (IHC4 Biomarker Panel).³⁰ The IHC4 panel uses antibodies that are among the most rigorously controlled and regulated assays in clinical practice and recommended for large-scale studies.³⁰ Finally, the QIA data involved morphometrics and IHC. These data allowed validation of the methods used with evaluation of technical and biological variability. These data were then integrated into assessment of the relative preservation characteristics or various fixatives using standard statistical measures where appropriate and non-parametric rank order comparisons.

Preliminary Experiments

Each of the initial 11 fixatives was evaluated by a pilot experiment that compared the given fixative with both neutral-buffered formalin (NBF) and Telly's. If the morphometric results were comparable, then the fixative was used in at least three additional experiments. When the morphometric results revealed major deficiencies using a given preservative such as poor antigen preservation or excessive shrinkage, the fixative was eliminated from further consideration and not used repeatedly. The full list of experimental comparisons of fixatives is primarily reported in Supplementary Data. Fixatives evaluated but eliminated

for unacceptable performance in one or more categories include FineFix, Mirsky's, Methacarn, Zn7, and Z-Fix. This approach allowed conservation of our resources, simplified analyses and concentration of efforts on a limited number of fixatives to assure more thorough statistical analysis.

Morphometrics

Liver

Since liver has been used previously to evaluate the effect of fixation,³⁹ it was used as a 'standard normal control' for comparison with the tumor tissues. The number of hepatocytic nuclei/unit area has been a common metric.³⁹ The liver provides a measure of shrinkage and a record of the natural biological variation and technical reproducibility.

With Feulgen staining, the samples within the four fixation categories were consistently similar, indicating reproducibility between experimental samples. The mean and standard deviation of nuclear size, nuclei/unit area, and n/c ratio varied in only a few instances (Figure 2). No statistically significant morphometric differences were observed between the livers of the two mouse strains (FVB/NJ and 129S6/SvEv) (Supplementary Figure S1).

Differences in mean nuclear size for liver were statistically significant ($P < 0.001$) across fixatives (Supplementary Table S4). Adjusted P -values were smaller than 0.05 only NBF compared with Telly's (Supplementary Table S4). Mean liver nuclei per mm² (Supplementary Table S5) differed significantly across fixatives ($P < 0.001$). Mean n/c ratio for liver also differed significantly across fixatives ($P < 0.001$; Supplementary Table S6). An overall mean rank based on the combined rank for all liver morphometric data is provided for each fixative (Supplementary Table S7).

Tumors

The different fixatives affected the cytology and histology of the tumor tissues when inspected visually. While the details of the cytology were notable, the major changes were in the degree of

Table 4 Final rank order of fixatives

Fixative ^a	Liver morphometrics ^b	Tumor morphometrics ^c	Myometrial ER/PR IHC ^d	Mouse tumor Ki67 IHC ^e	Mouse SSM2 tumor ER/PR IHC ^f	Mean rank of all categories
NBF	1.7	2.0	1.4	1.5	1.3	1.6
Telly's	3.0	3.0	2.2	2.1	2.5	2.5
Prefer	1.6	2.4	3.0	2.8	3	2.6
HChoice	2.7	3.0	2.0	3.2	2	2.6
PAX	2.4	2.3	3.1	3.1	3.8	2.9
Boon's	2.6	3.0	3.6	3.5	3.1	3.1

^aNNBF is a strong cross-linker, Telly's and Prefer are coagulator/cross-linkers, HChoice is a weak cross-linker, and PAX and Boon's are coagulators only.

^bMean rank of nuclear size, nuclei/mm², and n/c ratio (for breakdown, see Supplementary Tables 4–6).

^cMean rank of shrinkage, nuclear size, and nuclei/mm² (for breakdown, see Supplementary Table 7).

^dMean rank based on percent positive nuclei for ER (estrogen receptor) and PR (progesterone receptor).

^eWeighted mean rank based on density of positive nuclei of various experiments (see also Table 8).

^fWeighted mean rank based on density of positive nuclei from SSM2 mouse tumors (see also Supplementary Table 14).

shrinkage as measured by the area of space between dyscohesive tumor cells. Comparing the mean percent weak pixels (Supplementary Table S8), HistoChoice and Boon's had the most shrinkage, with moderate shrinkage for PAX, Prefer, and Telly's, and little shrinkage after NBF fixation (Figure 3).

The tumor mean nuclear size also differed significantly ($P < 0.001$) across the fixatives. After adjusting for experiment, NBF had the greatest area, while all other fixatives consistently fell below NBF and did not statistically differ ($P > 0.05$) from each other (Supplementary Table S9). In tumors, mean nuclei/mm² differed significantly ($P < 0.001$) across fixatives (Supplementary Table S10). Boon's differed significantly from NBF, HistoChoice, and Telly's, with the remaining fixatives (PAX and Prefer) not statistically distinguishable ($P > 0.05$) from either group (Supplementary Table S10). The mean rank order (Figure 3c) also indicated that strong cross-linkers had the least shrinkage.

Immunohistochemistry

ER and PR normal tissue

The uteri from each transplant recipient mouse was removed and processed and the outer layer of the myometrium used as the standard normal tissue for ER/PR IHC in each experiment. Pieces from the same uterus were placed into the various fixatives of each experiment. The image representation of the staining of ER with each fixative can be viewed (Supplementary Figure S2). LS Mean percent ER-positive nuclei differed significantly across fixatives ($P < 0.001$) with pairwise comparisons (Supplementary Table S11). The means for fixatives, adjusted for experiment, fell into two distinguishable clusters (Supplementary Table S11): HistoChoice, NBF, Telly's, and Prefer with relatively high percent positive nuclei (all above 66%), and Boon's and PAX all below 49%.

LS-Mean percent PR-positive nuclei differed significantly across fixatives ($P < 0.001$; Supplementary Table S12). Similarly to the ER, NBF, HistoChoice, Prefer, and Telly's clearly displayed higher values compared with Boon's and PAX and so the overall trends for ER and PR were similar.

Careful observation of the annotated fields during QIA of IHC-stained slides revealed that the weak cross-linking, glyoxal-based fixatives frequently had weak hematoxylin staining of the nuclei. This led to a relatively lower count of IHC-negative cells per unit area (density) and a relatively high percent of positive cells. When adjusted for unit area, the glyoxal-based fixatives had a lower density of positive cells per unit area.

NBF and Telly's fixatives were used in all experiments, providing an opportunity to judge the consistency of staining between experiments and to a certain extent to assess biological variation. The average percentage of positive nuclei for uteri ($n = 16$) fixed in NBF and stained for ER was 72.3% (Supplementary Table S13), and for PR, 58.6% (Supplementary Table S12). The average percentage of positive nuclei for $n = 16$ uteri fixed in Telly's and stained for ER was 67.5% (Supplementary Table S13), and for PR, 35%

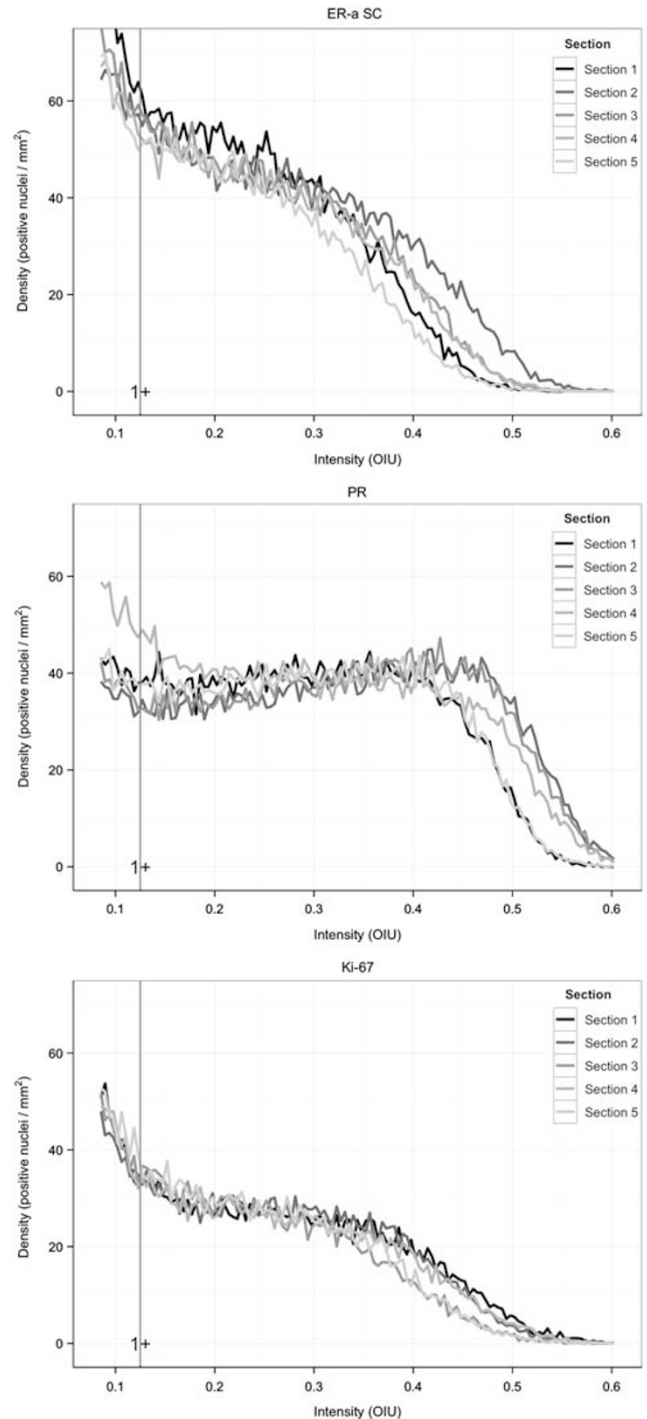


Figure 5 Technical reproducibility. Five serial sections from the same SSM2 tumor were stained for ER, PR or Ki-67 to assess technical variance. Lines represent each of the five stained slides analyzed with the nuclear algorithm and binned based on intensity (0 = lightest and 1.0 = darkest) in optical intensity units (OIU). Y axis is density in cells per mm². All cells to the right of the vertical line labeled '1+' are positive based on the default cutoff of 0.125. Note that the shape and location of the lines are very similar. Compare these graphs with those from different tumors in Figure 6.

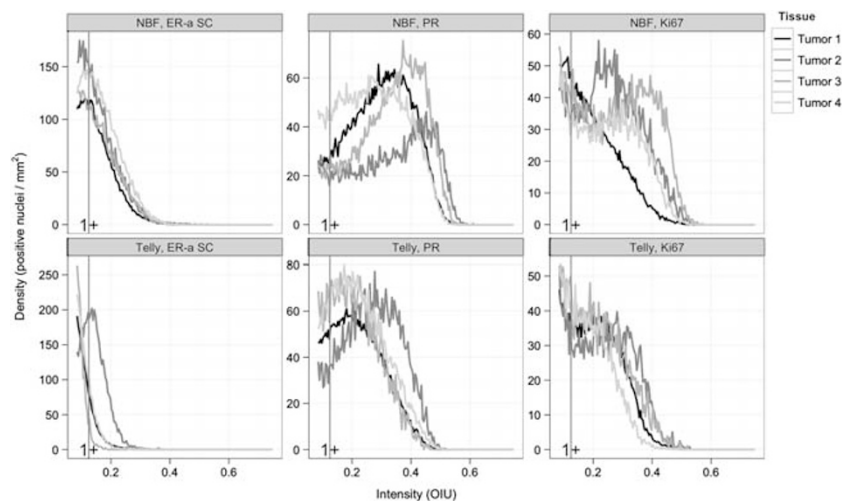


Figure 6 Biological variance within SSM2 tumors for ER, PR, and Ki67. Slides from four separate SSM2 tumors in the same mouse. Graphs represent ER, PR, and Ki67-stained slides analyzed with the nuclear algorithm and binned based on intensity (0 = lightest and 1.0 = darkest) in optical intensity units (OIU). Y axis is density in cells per mm^2 . Compare the marked difference in the four tumors as compared with the serial sections from one tumor in Figure 5. The top row represents tissue fixed in NBF and can be compared with Telly's fixed tissue (bottom row). Note that both fixatives illustrate the biological variance between tumors in the same mouse.

(Supplementary Table S12). Rank order for ER and PR was similar (Supplementary Table S13). Based on the rank order, NBF, HistoChoice, and Telly's were close and as a group led the other fixatives (average <2). These belong to the fixation categories strong cross-linker, weak cross-linker, and cross-linker with alcohol, respectively. The next cluster ranked 3 or higher. The lowest group all ranked 4 (for a summary ranking based on all measured parameters, see Table 4).

Ki67 in standard tissue

No one normal organ or tissue exhibited a fixed proportion of Ki67-positive cells. However, each run was accompanied by normal human tonsils and normal mouse lymph nodes as positive and negative (primary antibody deletion) controls. All data within a given experiment were compared in rank order.

ErbB2 standard tissues

Like Ki67, establishing a standard for ErbB2 represented a challenge without a consistently positive normal tissue. Previous tumor samples known to be ErbB2 positive were used as external positive and negative (antibody deletion) controls. The NDL mammary tumor, which overexpresses ErbB2, was chosen as the standard.

Mouse Mammary Tumors

The three mouse mammary tumor cell lines (Met-1, SSM2, and NDL-1), modeling three classes of human breast cancer (triple negative, luminal, and Her2 positive), were transplanted into syngeneic mice and allowed to grow 10 mm before harvesting for processing (Table 2).

IHC4 biomarker panel

Antigen preservation was the key element in our assessment of fixatives. Quantitative comparison of antigens was

performed using three statistical criteria and evaluated by rank order within each experiment. The usual assessment of tumor antigens is based on the opinions of trained experts who score antigenicity by percent positive tumor cells and, in some modifications, the most prevalent intensity, resulting in a composite score (Allred score⁴⁰ and H-score⁴¹ for diagnosis/prognosis). The results here are represented using a unique graph that demonstrates the density of cells in each stain-intensity bin. Although bins have been used before to show stain distribution, the unique graphs allowed more detailed analysis.

The biological variation from tumor to tumor within any given cell line was large enough to create overlaps between some fixatives in the same tumor type. Therefore, rank order was also used within the given tumor sample. The reproducibility and variability are reported below.

Reproducibility

Reproducibility and variance in biological systems is a major, but rarely documented, concern. Tumor heterogeneity can be an additional interpretive challenge.³⁶ Technical reproducibility was demonstrated by staining serial sections of the same tumor fixed with NBF. The reproducibility is recorded by the percent tumor cells staining and positive cell density. In the case illustrated (Figure 5), 15 serial sections from an SSM2 tumor were stained with anti-ER (sections 1–5), anti-PR (sections 6–10), and anti-Ki67 (sections 11–15) antibodies. The fidelity is illustrated using positive cell density:nuclear stain intensity graphs that show superimposed curves (Figure 5).

Technical reproducibility was tested by repeating the same IHC staining on the five serial sections. The graphs show the limited variance (Figure 5). Note that the overall

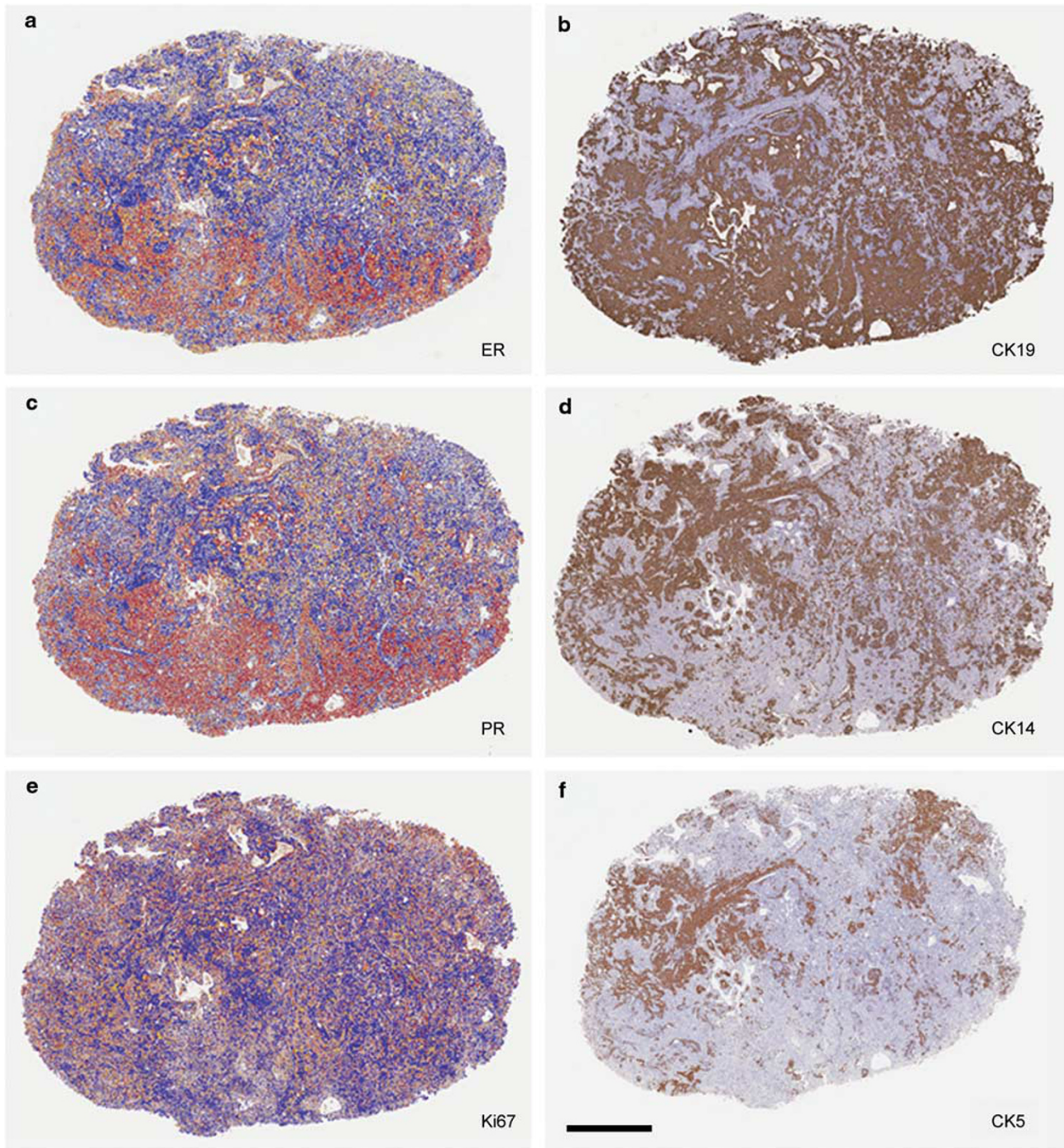


Figure 7 Illustration of tumor heterogeneity using immunohistochemistry and IHC nuclear algorithm overlay. Stained serial sections from an SSM2 ER⁺, PR⁺, Her2⁻ mammary tumor are compared. Using the algorithm overlay (**a**, **c**, **e**), nuclear antigens ER (**a**), PR (**c**), and Ki67 (**e**), stain at different intensities within the tumor. Note the more red, orange color (more intensely stained positive nuclei) in the lower part of the tumor in (**a**) and (**c**) and a more blue color (negative nuclei) in the upper part of the tumors. Ki67 staining is more evenly distributed (**e**). Supplementary Table S16 contains the QIA summary that documents the magnitude of the differences between the high and low areas for ER, PR, and Ki67. Structural antigens CK19 (**b**), CK14 (**d**), and CK5 (**f**) illustrate tumor heterogeneity with reciprocal staining. Note blue negative staining areas for CK19 (**b**) appear to be positive for CK14 (**d**) and CK5 (**f**) and some CK14-positive areas, negative for CK5. Scale bar length is 1 mm.

density:intensity results, rank order, and interpretation were similar for the serial sections.

Biological variance between samples of the same tumor cell line was determined by examining multiple tumor explants

from the same passage of cultured cells transplanted into the same recipient host at the same time and harvested and processed at the same time. The example illustrated includes four SSM2 tumors from the same animal (Figure 6). In this case, as

Table 5 Adjusted P-values of pairwise comparisons (Tukey–Kramer) of least squares means for Ki67 staining in SSM2 tumors

Mean percent positive nuclei	Fixative ^a	NBF	Telly's	HChoice	Boon's	Prefer	PAX
38.67	NBF		0.1303	0.1247	0.0093^b	0.0007	<0.0001
31.48	Telly's			0.9886	0.6092	0.0984	0.0026
27.27	HChoice				0.9997	0.7022	0.3627
23.64	Boon's					0.9955	0.7017
18.43	Prefer						1.0000
15.52	PAX						

^aNBF is a strong cross-linker, Telly's and Prefer are coagulator/cross-linkers, HChoice is a weak cross-linker, and PAX and Boon's are coagulators only.

^bStatistically significant differences are in boldface.

Table 6 Adjusted P-values of pairwise comparisons (Tukey–Kramer) of least squares means for Ki67 staining in Met-1 tumors

Mean percent positive nuclei	Fixative ^a	NBF	Prefer	HChoice	Telly's	PAX	Boon's
49.42	NBF		1.0000	0.9719	0.8020	0.4087	0.0050^b
48.82	Prefer			0.9927	0.9403	0.5577	0.0189
45.62	HChoice				0.9997	0.7869	0.0498
44.19	Telly's					0.8194	0.0342
34.95	PAX						0.6766
22.68	Boon's						

^aNBF is a strong cross-linker, Telly's and Prefer are coagulator/cross-linkers, HChoice is a weak cross-linker, and PAX and Boon's are coagulators only.

^bStatistically significant differences are in boldface.

Table 7 Adjusted P-values of pairwise comparisons (Tukey–Kramer) of least squares

Mean percent positive nuclei	Fixative ^a	NBF	Telly's	HChoice	Prefer	Boon's	PAX
33.13	NBF		0.3277	0.2004	0.0055^b	0.0074	0.0095
28.93	Telly's			0.9989	0.3166	0.3211	0.1530
27.68	HChoice				0.7179	0.6789	0.3364
23.19	Prefer					1.0000	0.9634
22.88	Boon's						0.9670
18.75	PAX						

Means for Ki67 staining in ND1 tumors.

^aNBF is a strong cross-linker, Telly's and Prefer are coagulator/cross-linkers, HChoice is a weak cross-linker, and PAX and Boon's are coagulators only.

^bStatistically significant differences are in boldface.

might be expected, the variance is greater than that can be accounted for by technical variation. Biological variance among tumors was one factor that led to the use of rank order.

Intra-tumor heterogeneity, another concern, is also demonstrated in the SSM2-based experiment (Figure 5). This tumor shows intratumoral heterogeneity, which is demonstrated by comparing serial sections of one tumor using cytokeratins (CK) IHC (Figure 7). As illustrated, CK19 and CK5 stain almost completely different cell populations of cells within the SSM2 tumor. CK14 distribution seems to overlap the other two. Thus, the tumor itself has at least three demonstrable populations. Of interest, the ER and PR stain is heavily concentrated in the CK19 population but Ki67 stain is not (Figure 7). These illustrations emphasize the levels of technical reproducibility and biological variance.

Antigen Preservation

ER/PR

Fixative type affected immunohistochemical staining (Figure 4). As indicated above, the fixatives were rated by multiple criteria: standard error of the mean, rank order by percent, and rank order by density. They were further analyzed on the basis of their rank per animal and per tumor. In general, NBF ranked, or scored, as number one, but other cross-linking fixatives such as Telly's, Prefer, and HistoChoice had closely related ranks and scores (Supplementary Tables S14 and S15). The most dramatic fixative effect was observed with ER and PR. In general, the fixatives containing strong and weak cross-linkers offered the highest percentage of ER- and PR-positive cells while the coagulants often drastically reduced the staining for ER and PR (Figure 4).

Ki67

Since all of the tumors display cell division, all mouse tumor IHC results for Ki67 could be pooled for stringent parametric statistical comparisons. However, since different tumor types with different proliferative rates were used, the data are most accurately presented in relation to the tumor type (Tables 5–7). As a further refinement to account for the biological variability between tumors of different types and of the same type, the data were also analyzed by rank order for each tumor in each animal expressed as 'weighted means'. When viewed as a rank order by animal, which sorts according to the fixative and the individual animal, NBF, Prefer, Telly's, and HistoChoice are the highest ranking (Table 8).

As discussed, light counterstain might result in missed nuclei by the algorithm, thereby affecting the digital scoring. Thus, density of positive cells per mm² was used to eliminate any miscalculation of negative cells. When the Ki67 staining was ranked according to positive cell density, HistoChoice fell from second to the sixth rank using both mean rank and weighted means and elevated Telly's to second rank while NBF maintained its first rank (Table 8). Figure 8 provides visual confirmation of the results.

Table 8 Ranking table for Ki67 in mouse tumor tissue

Fixative ^a	Experiment number ^b										Mean rank ^c	Rank s.d.	Weighted mean ^d	Weighted s.d.
	6	7	8	10	11	12	13	14	15	16				
<i>Ki67 percent positive rank</i>														
NBF	2	3	1	1	1	1	1	2	1	1	1.4	0.66	1.57	0.84
Prefer						4	2	1			2.33	1.25	3	1
Telly's	1	1	2	2	2	3	4	4	3	2	2.4	1.02	2.21	0.83
HChoice						2	3	3	2		2.5	0.5	2.46	1.33
PAX	3	2	4	4							3.25	0.83	3.43	0.79
Boon's	4	4	3	3					4		3.6	0.49	3.36	0.67
<i>Ki67 positive density rank</i>														
NBF	2	2	1	1	1	1	1	2	1	1	1.3	0.46	1.54	0.79
Telly's	1	1	2	2	2	2	3	3	2	2	2	0.63	2.07	0.9
Prefer						4	2	1			2.33	1.25	2.78	1.3
PAX	3	3	3	4							3.25	0.43	3.14	0.69
HChoice						3	4	4	3		3.5	0.5	3.15	0.69
Boon's	4	4	4	3					4		3.8	0.4	3.45	0.82

^aNBF is a strong cross-linker, Telly's and Prefer are coagulator/cross-linkers, HChoice is a weak cross-linker, and PAX and Boon's are coagulators only.

^bSeparate experiments had NDL (Exp. 6, 15, and 16), SSM2 (Exp. 10, 11, and 12) or Met-1 tumors (Exp. 7, 8, 9, 13, and 14).

^cMean rank (rank = Lc) weighs experiments equally regardless of the number of tumors analyzed.

^dWeighted Mean (mean = Lc) represents the mean rank of that fixative for all tumors.

ErbB2

ErbB2 IHC was somewhat complicated and limited to one mouse tumor type (NDL), which has very high levels of antigen expression. Further, the antibodies commonly used for IHC in the clinical laboratory gave high background on mouse tissues. When using a more specific monoclonal antibody, the differences among fixatives were not statistically significant at the dilutions used (data not shown). Thus, the human sample was more informative.

Human Breast Cancers

Three human cancers were obtained under optimal conditions for processing and comparison (characteristics in Table 2). The preservation of antigens (IHC) in these three cases is illustrated in Figure 9. Since only one of the tumors expressed ER, PR, and Her2, only Ki67 could be statistically compared in all three cases (Table 9). As predicted by the parallel studies with the mouse tumors, NBF and Telly's ranked 1 and 2. The ER, PR, and Ki67 staining in human showed the same sensitivities to fixatives as in the mouse tumors with diminished staining with the coagulating fixatives (Figure 9). However, the IHC for Her2 revealed a profound fixative effect (Figure 9), the magnitude of which was documented with the density:intensity graphs (Figure 10).

DISCUSSION

The effects of fixation and processing on IHC are apparent to any experienced morphologist. A number of studies have used panels of pathologists to validate the use of one fixative over another. The morphology and IHC are frequently reported as 'comparable' to formalin. However, the reliability, reproducibility, and documentation of empirical observations between pathologists are open to question. Further, the degree of differences in the IHC preparations are difficult to document without accounting for observer bias. QIA offers a technology that promises to be more reproducible with unbiased documentation if appropriate morphometric parameters can be developed and applied.

We report here the application of QIA to morphology and IHC while dealing with a specific problem involving the effect of various preservatives on the morphology and antigenicity of breast cancers. Samples of mouse models of three different types of human breast cancer and three separate human breast cancers were evaluated by two experienced breast pathologists and by QIA. Representatives of four classes of fixatives were used and the results were analyzed using the algorithms provided by the manufacturer. The preset parameters were adjusted and false negatives and false positives were identified during the pre-analytical stage to minimize obvious technical errors. In addition, a unique Density:Intensity (D:I) graph was developed, which allowed

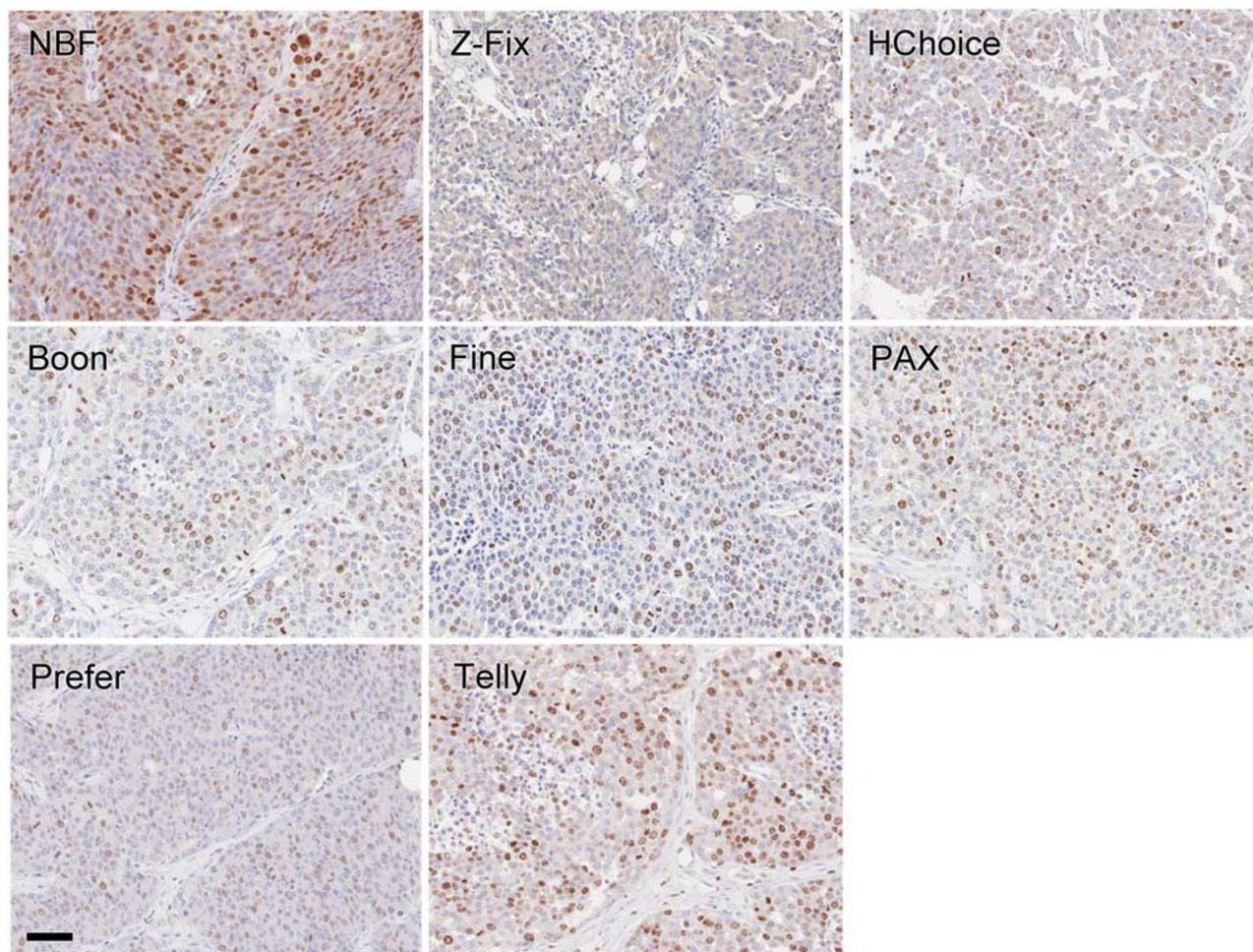


Figure 8 Effect of fixatives on Ki67 immunohistochemistry in mouse ND1 tumor. This panel provides the visual evidence of differences in both the intensity of brown positive nuclei and intensity of blue negative nuclei observed with the different fixatives. Note the difference in the counterstain that affects the number of nuclei counted as 'negative' cells. Scale bar length is 50 μ m.

detailed comparison between samples. The D:I graphs provided representation of the full range of IHC events that documented technical and biological reproducibility and variance. Intratumoral variation was also readily apparent and documented.

Bulk analysis of the raw data revealed statistically significant differences in the morphological and antigenic preservation with different fixatives. However, such analyses did not adequately represent the empirical observations of our pathologists. The bulk analyses obscured very relevant biological variation and tumor heterogeneity observed by our pathologists and documented by QIA. Therefore, an ordinal rank order was applied that accounted for variation and heterogeneity and, in our opinion, most accurately represented the empirical observations. We are confident that this approach using D:I graphs and ordinal rank order can be useful when applied to similar problems in analysis of cancer.

The major issues in this study involved the documentation of technical reproducibility and biological variance.

Pre-analytical verification of all data sets by experienced morphologists was required to detect variations in reagent reliability and false positive and false negative digital annotations. When the technical variables were carefully controlled, the density:intensity graphs documented excellent technical reproducibility. However, the element of human judgment was still required for validation of the results.

Although technical reproducibility within each experiment was excellent, greater biological variability was measurable between experiments. The magnitude of biological variance was reflected when multiple tumor transplants of the same cells were placed into the same or different mice. Although variability was demonstrable within the same transplant generation, greater variation was discernible between different transplant generations of the same tumor cell line. Intratumoral heterogeneity proved to be another source of variation.³⁶ Serial sections of a single tumor using multiple antigens illustrated intratumoral heterogeneity that would only be detected by experienced morphologists and quantitative analysis of the IHC.

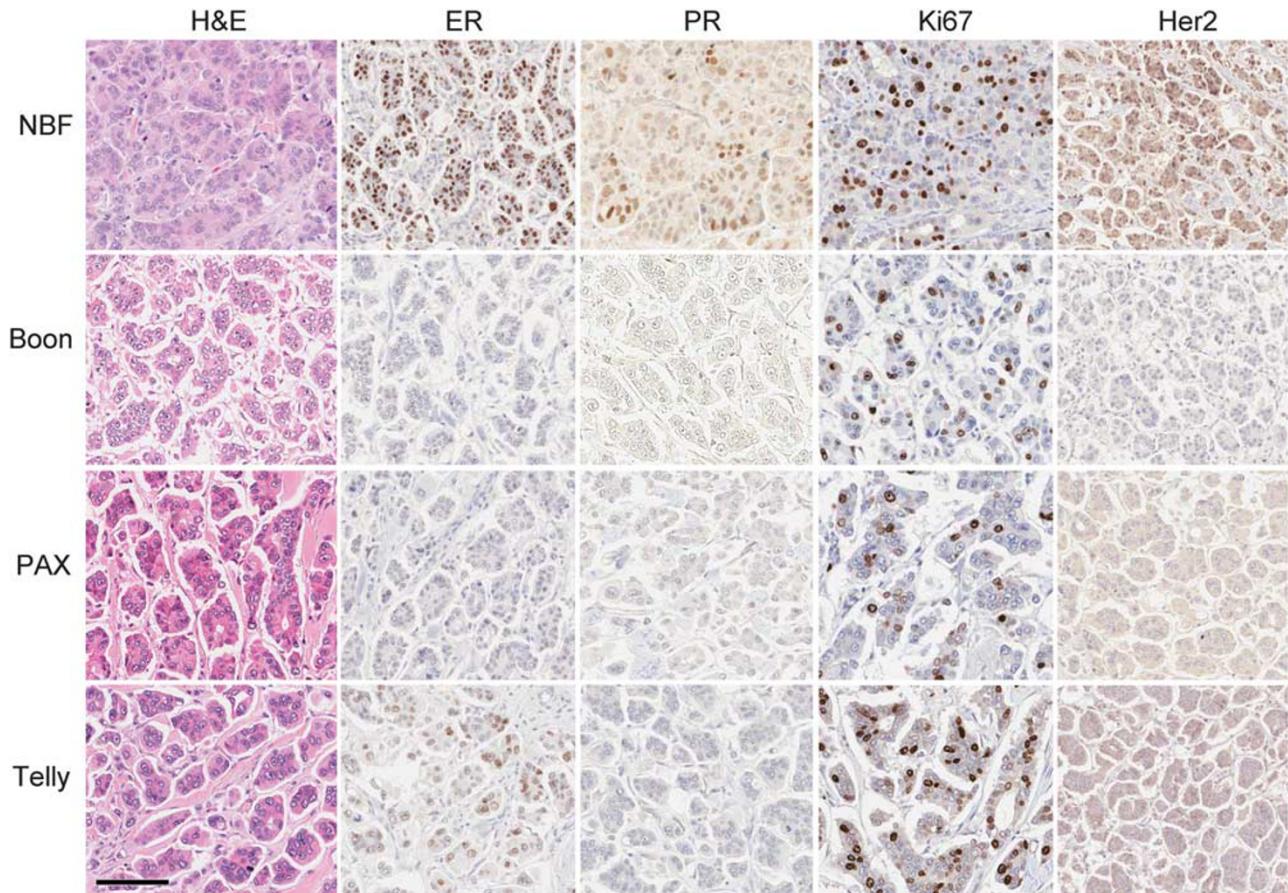


Figure 9 Effect of fixative on appearance of H&E and immunohistochemistry of IHC4-related antigens in a human breast cancer. This panel provides histological illustration of the variability in the intensity of both brown positive nuclei and background hematoxylin staining. Note that the ER, PR, and Her2 staining is lost with the coagulating fixatives Boon's and PAX. Scale bar is 100 μ m.

Table 9 Human tumor Ki67 cell density by ranking

Fixative ^a	Rank order for experiment			Rank sum	Mean rank	Rank sum by class	Mean class rank
	Human tumor 1	Human tumor 2	Human tumor 3				
NBF	1	2	1	4	1.33	4	1.33
Telly's	2	1	2	5	1.67		
Prefer	NA ^b	NA	2	2	2	7	1.75
HChoice	NA	NA	3	3	3	3	
Boon's	3	4	NA	7	3.5	7	
PAX	4	3	NA	7	3.5	7	3.5

^aNBF is a strong cross-linker, Telly's and Prefer are coagulator/cross-linkers, HChoice is a weak cross-linker, and PAX and Boon's are coagulators only.

^bNot assessed.

The QIA data were examined using a number of statistical approaches including least squares of the means (LC), and pair-wise comparisons that included bulk analysis (see Supplementary Tables S4–S15). Some of these statistical

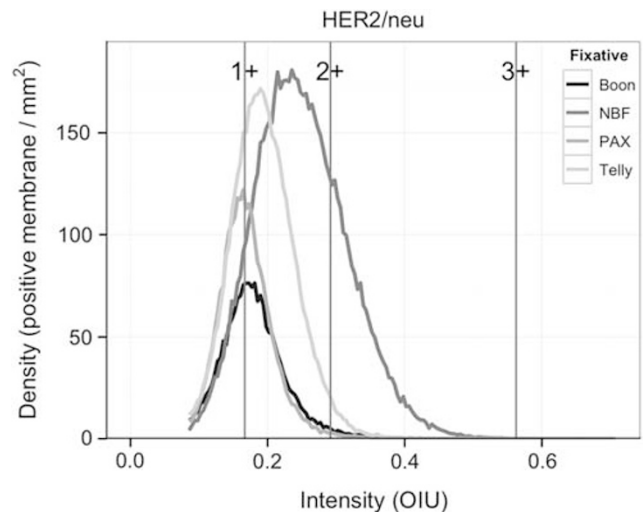


Figure 10 Fixative alteration of Her2 staining in a human breast cancer. Slides are from a single human breast cancer specimen stained for Her2. These graphs represent the Her2-stained slides analyzed with the Membrane algorithm and binned based on intensity (0 = lightest and 1.0 = darkest) in optical intensity units (OIU). Y axis is density in positive membrane cells per mm^2 . Horizontal lines represent the default cutoff values for 1+, 2+, and 3+ cells.

analyses were revealing for control tissues such as liver but they did not accurately reflect the biological variation within the three transplant tumor lines. Consequently, a non-parametric approach was chosen to compensate for internal heterogeneity as well as biological and technical variation. Thus, the quantitative data were 'normalized' by rank order within each experiment. The maximum value for each measurement in each individual experiment was compared with all other values for a given parameter within the given experiment to arrive with a weighted mean order. The ordinal rank order comparison is widely used in biology and arguably provided more simplified and coherent, interpretable data comparisons.³⁸

The individual tumors were relatively small and needed to be divided between molecular and morphological samples with multiple fixatives. Thus, only a limited number of experimental groups (fixatives) could be used in any given experiment. Therefore, NBF and Telly's solution were used in all experiments to establish reproducibility and variance. In one sense, NBF and Telly's became the 'gold standard' for other fixatives. This triage strategy left sufficient 1–2 mm slices for two additional fixatives or, in larger viable tumors, up to six additional fixatives. In most cases, each additional fixative was used at least three times in each of the tumor types. However, the initial trials with five fixatives proved them to be below acceptable standards. The observations on these five fixatives are reported in Supplementary Data but were not incorporated into the final statistical analysis.

Overall, the sum of the data indicates that NBF ranked highest in preservation of both morphology and antigenicity. Using a combination of morphometric and immunohistochemical data and SOPs, NBF was consistently ranked number 1 or 2. NBF was also ranked number 1 when all experimental ranks were aggregated. The combination fixatives, containing both a cross-linker and an alcohol (Telly's and Prefer), were close seconds. The glyoxal-based fixatives (HistoChoice) generally suffered with relatively poor performances in the morphometric categories. All alcohol-based fixatives performed poorly in IHC for ER and PR and resulted in altered morphology. In general, these differences were readily apparent upon empirical inspection by pathologists. However, the magnitude of variation is more accurately documented by QIA.

In a rarely performed comparative study with several optimally processed human breast cancers, the same relative patterns of staining as revealed by the density:intensity graphs were demonstrable, resulting in similar rank orders for the various classes of fixatives. Although the number of available samples limited the statistical analysis, the comparison with the more extensive mouse experiments illustrated the robustness of the approach.

We have created an approach and a database designed to provide testable standards based on SOPs. While the value of quantitative IHC for disease management remains controversial,⁴² we believe that this type of rigorous analysis

with new technologies, under SOP, is critical to continue advancement in histopathology.¹ This technology can rigorously test whether current standards of analysis have prognostic or predictive clinical or biological significance. We believe that morphometric analysis will become an important tool for comparison of image-based data. Clearly, the use of QIA will facilitate comparisons of fixatives and operational systems from different laboratories. However, these types of morphometric comparisons currently require careful supervision by qualified professionals, and have yet to be fully automated.

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

ACKNOWLEDGEMENTS

We acknowledge the genesis of these studies as arising from our participation in two cooperative organizations. First, the Athena Breast Health Network (a University of California-wide collaboration aiming to integrate clinical care and research and drive innovation in patient-centered prevention, screening, and treatment of breast cancer) and, second, the National Cancer Institute's Mouse Models of Human Cancers Consortium. Athena program needs include the preservation of structural, antigenic, and molecular analysis quality maximizing the utility of increasingly small specimens for research and standard of care diagnosis. We wish to acknowledge the support of Dr Laura Esserman, Athena Breast Health Network and University of California, San Francisco, who inspired and encouraged this project. The SSM2 cells were a gift from Drs RD Schreiber and S Chan, Washington University, Department of Pathology, St Louis, MO. We thank Dr Sarah Elson, Athena, for critical review of the manuscript and Dr Josh Miller for thoughtful commentary during the development of the project. The research described here was supported by the Athena Breast Health Network, through the support of the Safeway Foundation and the University of California Office of the President (RC, JE, AB, NH, and JW); National Cancer Institute's Mouse Models of Human Cancers Consortium (U01 CA141582, U01 CA141541, U01 CA105490-01) (RC, AB, QC, JG, HV). Statistical support was made possible by the National Center for Research Resources (NCR), a component of the National Institutes of Health (NIH), and NIH roadmap for Medical Research (UL1 RR024146) (LB, YJL); its contents are solely the responsibility of the authors and do not necessarily represent the official view of NCR or NIH.

DISCLOSURE/CONFLICT OF INTEREST

The authors declare no conflict of interest.

1. Dunstan RW, Wharton Jr KA, Quigley C, *et al*. The use of immunohistochemistry for biomarker assessment—can it compete with other technologies? *Toxicol Pathol* 2011;39:988–1002.
2. Kothmaier H, Rohrer D, Stacher E, *et al*. Comparison of formalin-free tissue fixatives: a proteomic study testing their application for routine pathology and research. *Arch Pathol Lab Med* 2011;135:744–752.
3. NHS. Evidence review: non-formalin fixatives. In: National Health Service Purchasing and Supply Agency CFe-bp editor. Reading, UK, Vol. CEP090919 (2009), pp1–36.
4. Belief V, Boissiere F, Bibeau F, *et al*. Proteomic analysis of RCL2 paraffin-embedded tissues. *J Cell Mol Med* 2008;12:2027–2036.
5. Boon ME, Kok LP. Theory and practice of combining coagulant fixation and microwave histoprocessing. *Biotech Histochem* 2008;83:261–277.
6. Chung JY, Braunschweig T, Williams R, *et al*. Factors in tissue handling and processing that impact RNA obtained from formalin-fixed, paraffin-embedded tissue. *J Histochem Cytochem* 2008;56:1033–1042.

7. Cox ML, Schray CL, Luster CN, *et al*. Assessment of fixatives, fixation, and tissue processing on morphology and RNA integrity. *Exp Mol Pathol* 2006;80:183–191.
8. Delfour C, Roger P, Bret C, *et al*. RCL2, a new fixative, preserves morphology and nucleic acid integrity in paraffin-embedded breast carcinoma and microdissected breast tumor cells. *J Mol Diagn* 2006;8:157–169.
9. Dotti I, Bonin S, Basili G, *et al*. Effects of formalin, methacarn, and fineFIX fixatives on RNA preservation. *Diagn Mol Pathol* 2010;19:112–122.
10. Espina V, Mueller C, Edmiston K, *et al*. Tissue is alive: new technologies are needed to address the problems of protein biomarker pre-analytical variability. *Proteomics Clin Appl* 2009;3:874–882.
11. Gazziero A, Guzzardo V, Aldighieri E, *et al*. Morphological quality and nucleic acid preservation in cytopathology. *J Clin Pathol* 2009;62:429–434.
12. Giannella C, Zito FA, Colonna F, *et al*. Comparison of formalin, ethanol, and Histochoice fixation on the PCR amplification from paraffin-embedded breast cancer tissue. *Eur J Clin Chem Clin Biochem* 1997;35:633–635.
13. Gomez-Fernandez C, Daneshbod Y, Nassiri M, *et al*. Immunohistochemically determined estrogen receptor phenotype remains stable in recurrent and metastatic breast cancer. *Am J Clin Pathol* 2008;130:879–882.
14. Jensen UB, Owens DM, Pedersen S, *et al*. Zinc fixation preserves flow cytometry scatter and fluorescence parameters and allows simultaneous analysis of DNA content and synthesis, and intracellular and surface epitopes. *Cytometry A* 2010;77:798–804.
15. Lykidis D, Van Noorden S, Armstrong A, *et al*. Novel zinc-based fixative for high quality DNA, RNA and protein analysis. *Nucleic Acids Res* 2007;35:e85.
16. Marchionni L, Wilson RF, Marinopoulos SS, *et al*. Impact of gene expression profiling tests on breast cancer outcomes. *Evid Rep Technol Assess* 2007;160:1–105.
17. Mikaelian I, Nanney LB, Parman KS, *et al*. Antibodies that label paraffin-embedded mouse tissues: a collaborative endeavor. *Toxicol Pathol* 2004;32:181–191.
18. Mook S, Schmidt MK, Weigelt B, *et al*. The 70-gene prognosis signature predicts early metastasis in breast cancer patients between 55 and 70 years of age. *Ann Oncol* 2009;21:717–722.
19. Nassiri M, Ramos S, Zohourian H, *et al*. Preservation of biomolecules in breast cancer tissue by a formalin-free histology system. *BMC Clin Pathol* 2008;8:1.
20. O'Leary TJ, Fowler CB, Evers DL, *et al*. Protein fixation and antigen retrieval: chemical studies. *Biotech Histochem* 2009;84:217–221.
21. Otali D, Stockard CR, Oelschlager DK, *et al*. Combined effects of formalin fixation and tissue processing on immunorecognition. *Biotech Histochem* 2009;84:223–247.
22. Paizs M, Engelhardt JI, Siklos L. Quantitative assessment of relative changes of immunohistochemical staining by light microscopy in specified anatomical regions. *J Microsc* 2009;234:103–112.
23. Prento P, Lyon H. Commercial formalin substitutes for histopathology. *Biotech Histochem* 1997;72:273–282.
24. 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.
25. Stanta G, Mucelli SP, Petrer F, *et al*. A novel fixative improves opportunities of nucleic acids and proteomic analysis in human archive's tissues. *Diagn Mol Pathol* 2006;15:115–123.
26. Titford ME, Horenstein MG. Histomorphologic assessment of formalin substitute fixatives for diagnostic surgical pathology. *Arch Pathol Lab Med* 2005;129:502–506.
27. Vincek V, Nassiri M, Nadji M, *et al*. A tissue fixative that protects macromolecules (DNA, RNA, and protein) and histomorphology in clinical samples. *Lab Invest* 2003;83:1427–1435.
28. Xie R, Chung JY, Ylaya K, *et al*. Factors influencing the degradation of archival formalin-fixed paraffin-embedded tissue sections. *J Histochem Cytochem* 2011;59:356–365.
29. Viertler C, Groelz D, Gündisch S, *et al*. A new technology for stabilization of biomolecules in tissues for combined histological and molecular analyses. *J Mol Diagn* 2012;14:458–466.
30. Cuzick J, Dowsett M, Pineda S, *et al*. Prognostic value of a combined estrogen receptor, progesterone receptor, Ki-67, and human epidermal growth factor receptor 2 immunohistochemical score and comparison with the Genomic Health recurrence score in early breast cancer. *J Clin Oncol* 2011;29:4273–4278.
31. Chan SR, Vermi W, Luo J, *et al*. STAT1-deficient mice spontaneously develop estrogen receptor alpha-positive luminal mammary carcinomas. *Breast Cancer Res* 2012;14:R16.
32. Borowsky AD, Namba R, Young LJ, *et al*. Syngeneic mouse mammary carcinoma cell lines: two closely related cell lines with divergent metastatic behavior. *Clin Exp Metastasis* 2005;22:47–59.
33. Siegel PM, Ryan ED, Cardiff RD, *et al*. Elevated expression of activated forms of Neu/ErbB-2 and ErbB-3 are involved in the induction of mammary tumors in transgenic mice: implications for human breast cancer. *EMBO J* 1999;18:2149–2164.
34. Lloyd MC, Allam-Nandyala P, Purohit CN, *et al*. Using image analysis as a tool for assessment of prognostic and predictive biomarkers for breast cancer: How reliable is it? *J Pathol Inform* 2010;1:29.
35. Mulrane L, Rexhepaj E, Penney S, *et al*. Automated image analysis in histopathology: a valuable tool in medical diagnostics. *Expert Rev Mol Diagn* 2008;8:707–725.
36. Potts SJ, Krueger JS, Landis ND, *et al*. Evaluating tumor heterogeneity in immunohistochemistry-stained breast cancer tissue. *Lab Invest* 2012;92:1342–1357.
37. Noe M, Kunz G, Herberth M, *et al*. The cyclic pattern of the immunocytochemical expression of oestrogen and progesterone receptors in human myometrial and endometrial layers: characterization of the endometrial-subendometrial unit. *Hum Reprod* 1999;14:190–197.
38. Lehmann EL, D'Abbrera HJM. *Nonparametrics: Statistical Methods Based On Ranks*. Rev. 1st edn (, New York, 2006) xvi463p.
39. Fox CH, Johnson FB, Whiting J, *et al*. Formaldehyde fixation. *J Histochem Cytochem* 1985;33:845–853.
40. Allred DC, Carlson RW, Berry DA, *et al*. NCCN Task Force Report: estrogen receptor and progesterone receptor testing in breast cancer by immunohistochemistry. *J Natl Compr Canc Netw* 2009;7(Suppl 6):S1–S21, quiz S22–23.
41. McClelland RA, Finlay P, Walker KJ, *et al*. Automated quantitation of immunocytochemically localized estrogen receptors in human breast cancer. *Cancer Res* 1990;50:3545–3550.
42. Nadji M. Quantitative immunohistochemistry of estrogen receptor in breast cancer: 'much ado about nothing!'. *Appl Immunohistochem Mol Morphol* 2008;16:105–107.