

Influence of slide aging on results of translational research studies using immunohistochemistry

Martina Mirlacher, Marlis Kasper, Martina Storz, Yvonne Knecht, Ursula Dürmüller, Ronald Simon, Michael J Mihatsch and Guido Sauter

Institute for Pathology, University of Basel, Basel, Switzerland

Several reports have shown that a long delay between cutting sections and immunohistochemical (IHC) staining can decrease the IHC reaction intensity. However, systematic large-scale studies to investigate to what extent this problem may influence the outcome of translational research studies are lacking. In this study, we used a tissue microarray (TMA) approach to investigate the influence of slide age on comparisons between the results of IHC analyses for estrogen receptor (ER), progesterone receptor (PR), cyclin D1, HER2 (HercepTest), and E-cadherin and clinical outcome in a series of 522 breast cancer patients. Old TMA sections stored for 6 months at 4°C and freshly cut sections were analyzed under exactly identical experimental conditions. As compared to results obtained on freshly cut sections, the frequency of positivity on old sections decreased from 65 to 46% for ER ($P < 0.0001$), from 33 to 18.5% for PR ($P < 0.0001$), from 16.3 to 9.6% for HER2 ($P = 0.0047$), from 45.1 to 37.7% for cyclin D1 ($P = 0.10$), and from 58.9 to 32.9% for E-cadherin ($P < 0.0001$). Despite the lower fraction of positive cases, most associations between IHC data and tumor phenotype that were observed in fresh section analysis were also found when old section data were analyzed. The results confirm that slide aging has a great influence on the intensity of IHC staining in individual cases, but they also suggest that many clinicopathological associations can be detected if suboptimally processed sections are used for IHC.

Modern Pathology (2004) 17, 1414–1420, advance online publication, 18 June 2004; doi:10.1038/modpathol.3800208

Keywords: tissue microarray; immunohistochemistry; slide aging; quality control

The collection of tissue material is an important component of many clinical trials studying the efficiency of different cancer therapies. It is hoped that molecular analyses of tissues of patients that were enrolled in studies will eventually allow a distinction of these patients that responded favorably to therapy. Tissue materials retrieved from clinical trials are often limited to a small number of unstained sections that are submitted by local pathologists and subsequently stored at a central place. A collection and central storage of tissue blocks has proved to be impractical because many institutions dislike to lose control over the tissue, which might be needed for additional diagnostic procedures to be performed years after the initial diagnosis and enrolling the patient in a study.

The current praxis of collecting unstained sections for later use is overshadowed by strong evidence suggesting that long-time storage of unstained sections will significantly deteriorate the results of molecular analyses.^{1–4} The problem is relevant for many molecular techniques including fluorescence *in situ* hybridization (FISH),⁵ RNA *in situ* hybridization,⁶ and immunohistochemistry (IHC).⁷ IHC is the most frequently used technology in translational research studies. Studies have shown that the staining intensity of immunohistochemical reactions deteriorates for most antibodies with time,^{1,2,7–12} that a significant decrease of the staining occurs as early as 2 weeks after sectioning,¹ and that sophisticated methods of slide storage only incompletely avoid deterioration of IHC stainings.^{1,13,14} These data question the utility of current slide collection schemes in clinical trials as a deterioration of IHC stainings can result in false-negative results.

No study has so far addressed the question of the influence of slide aging on the outcome of translational research studies using IHC. To investigate the impact of age-related IHC deterioration on statistical

Correspondence: Dr G Sauter, MD, Institute of Pathology, University of Basel, Schoenbeinstrasse 40, CH-4031 Basel, Switzerland.

E-mail: Guido.Sauter@unibas.ch

Received 22 January 2004; revised and accepted 6 April 2004; published online 18 June 2004

analyses, it is necessary to analyze a sufficiently large number of cases on tissue sections of different age. Obviously, performing a large IHC study twice with slides of different age would be impractical if traditional methods were used. Our recently developed tissue microarray (TMA) technology is highly suited for such methodological comparisons.^{15,16} In this method, minute samples (diameter 0.6 mm) from up to 1000 different tissues are placed on one microscope glass slide. Many minor parameters with a potential influence on the outcome of immunohistochemical analyses such as slide thickness, slide age, incubation time, reagent concentrations are fully standardized if consecutive sections from a TMA block are used in one experiment. In this study, we utilized fresh and old sections from a breast cancer TMA containing samples from 522 patients with clinical follow-up data for a systematic analysis of the influence of slide aging on the outcome of translational research projects using IHC.

Materials and methods

Patients

Formalin-fixed, paraffin-embedded tumor samples from 522 breast carcinomas had previously been included into a TMA. All slides from all tumors were reviewed by one pathologist (GS) to define the histologic grade according to Elston and Ellis¹⁷ (BRE) and the histologic tumor type. The pathologic stage and nodal status were obtained from the primary pathology reports. The series included 333 ductal, 82 lobular, nine medullary, 19 mucinous, 29 cribriform, 10 tubular, 14 papillary, eight apocrine, eight clear cell, two metaplastic, five atypical medullar, one histiocytic, one adenoid-cystic carcinoma, and one carcinosarcoma. There were 31.2% grade 1, 38.5% grade 2, and 30.3% grade 3 cancers. The pT stage was pT1 in 28.2%, pT2 in 53.9%, pT3 in 8.1%, and pT4 in 9.8%. Axillary lymph nodes had been examined in 458 patients (51.5% pN0, 43.9% pN1, 4.6% pN2). Raw survival information was available from the Cancer Registry of Basel. Advanced stage, high grade and positive nodal status were strongly associated with poor survival of our patients ($P < 0.0001$ each). The mean follow-up time was 90 months (range 3–152).

IHC

Two sets of slides taken from one TMA block were utilized for this study. The first set (old slides) was cut 6 months before the immunohistochemical staining and stored at 4°C. The second set was cut 1 week before IHC staining (fresh slides) and stored at room temperature. The used antibodies, their dilutions, and pretreatment conditions as well as the definitions used for definition of positivity are summarized in Table 1. For all antibodies, protocols were used that had previously been established either for routine use on 'large sections' or for scientific studies on TMAs. No further protocol optimization was attempted. All tumors were analyzed for ER, PR, and HER2. The analysis was restricted to the first two quadrants (306 tumors) of our TMA for cyclin D1 and E cadherin staining because of mechanical damage on the old sections. All TMA sections were scored by one pathologist (GS) who was unaware of the section age information in 1 day in a consecutive session to minimize intraobserver variability.

Statistics

Contingency table analysis and χ^2 tests were used to study the relationship between IHC results, grade, stage, and nodal status. Survival curves were plotted according to Kaplan–Meier. A log-rank test was applied to examine the relationship between IHC data and tumor-specific survival. Surviving patients were censored 3 months before the date of inquiry to the Cancer Registry.

Results

Frequency of Positivity

The average frequency of positive cells and the average staining intensity was lower on the old slides than in the fresh sections for all examined antibodies. Even though the staining intensity difference did not result in a different classification in all cases, significant differences were visible in almost all tumors with positive staining. Examples of tumors with a weaker staining on old slides than on fresh sections are shown in Figure 1. These cases

Table 1 Antibodies

Protein	Antibody	Source	Dilution	Pretreatment	Definition of positivity
HER-2/neu	HercepTest	DAKO	Prediluted	Waterbath 99°C	Scoring according to the manufacturer. 1+, 2+, and 3+ positive tumors were categorized as positive (only membrane staining counted)
Cyclin D1	DCS6	DAKO	1:200	Pressure cooker	Any positivity (nuclear or cytoplasmic)
E-cadherin	36B5	Novocastra	1:200	Pressure cooker	Any positivity (membraneous)
ER	6F11	DAKO	1:1000	Microwave	Nuclear positivity in $\geq 10\%$ of tumor cells
PR	PGR312	Novocastra	1:500	Microwave	Nuclear positivity in $\geq 10\%$ of tumor cells

illustrate the magnitude of the staining differences in individual cases that are beyond the possible level of analysis subjectivity. For all five antibodies, this resulted in a higher fraction of positive cases in the fresh section group as compared to the old section group (Figure 2) according to our preselected definitions. On freshly cut sections, the frequency of positivity (including weak staining) increased from 46 to 65% for ER ($P < 0.0001$), from 18.5 to 33% for PR ($P < 0.0001$), from 9.6 to 16.3% for HER2 ($P = 0.0047$), from 37.7 to 45.1% for cyclin D1 ($P = 0.10$), and from 32.9 to 58.9% for E-cadherin ($P < 0.0001$).

Association with Tumor Phenotype

We then analyzed previously established associations between gene expression and tumor phenotype. In particular, we were interested whether the known associations between PR and cyclin D1 expression with low tumor grade and positive ER status as well as the association between HER-2 positivity and high tumor grade and negative ER status could be confirmed on old and fresh TMA sections. The results of this analysis are shown in Table 2. Despite marked differences in the staining intensity of individual samples, all expected associations could be found on old and fresh TMA sections, although the level of significance was sometimes better for new section analyses than for old section results. The use of different cutoff levels for definition of positivity such as 1% positive nuclei for ER positivity or 2+/3+ for HER2 positivity did not change the overall message of our results (data not shown).

Association with Survival

In a further analysis, we investigated the association between our IHC data for ER, PR, HER2 and raw patient survival. These results are shown in Figure 3. Significant association with prognosis was detected for all parameters in the fresh section analysis. The association with poor survival was always weaker if old sections were analyzed. For HER2, statistical significance was still achieved while significance was not reached for ER ($P = 0.09$) and PR ($P = 0.11$). No association with patient survival was found for E-cadherin and cyclin D1 expression, neither on fresh nor on old sections.

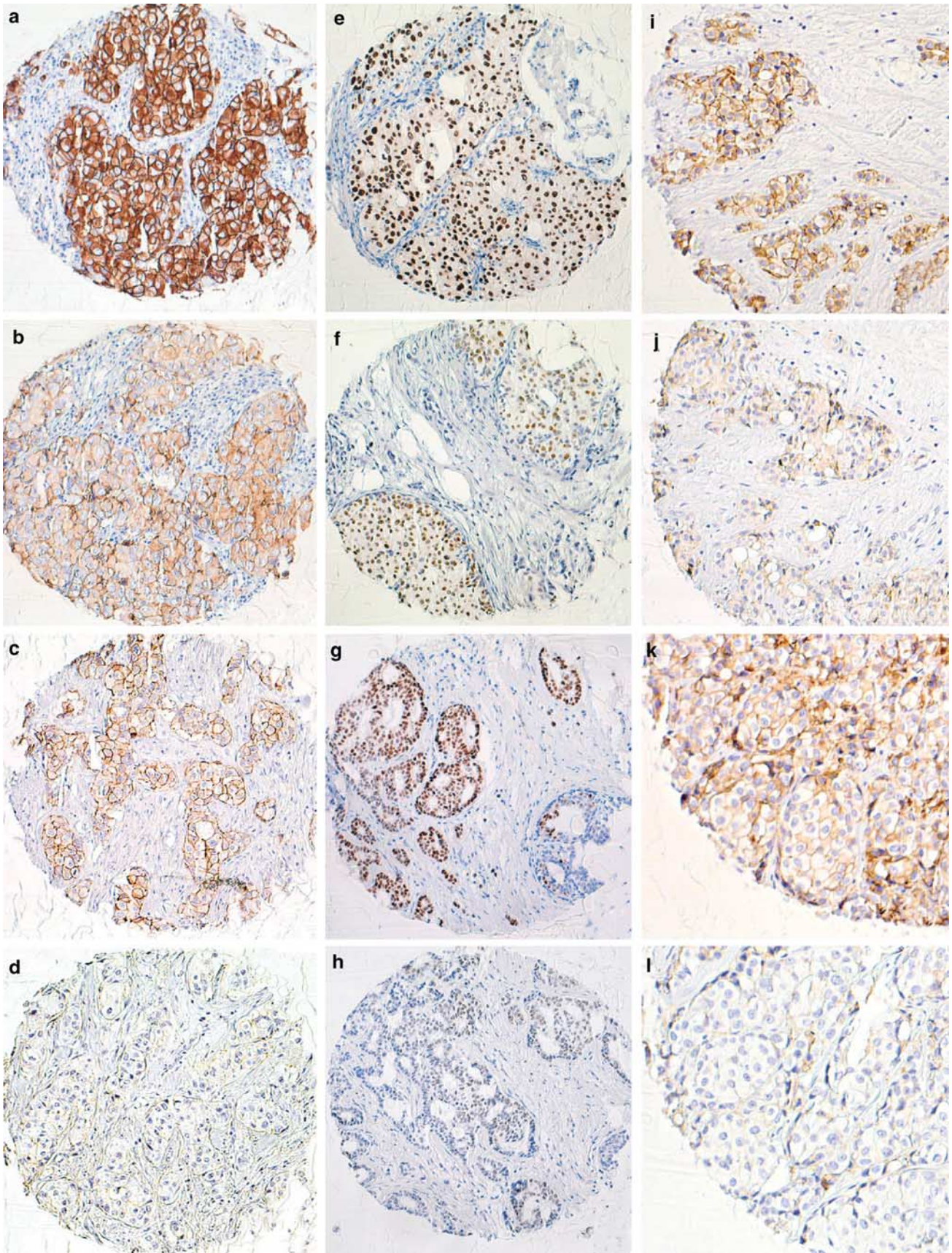
Discussion

Previous studies have demonstrated that the intensity of immunohistochemical stainings decreases for many antibodies after long-time storage of tissue sections.^{1–4,7,10–14} These studies have examined a variety of different antibodies but were always limited to a small number of tissues. Therefore, none of the earlier studies has provided data on the influence of slide aging on the results of research studies looking at associations between IHC results and other parameters. The use of TMAs enabled the rapid analysis of a sufficiently large set of ‘old’ and ‘freshly cut’ tumors and allowed us to investigate to what extent the age of tissue sections would affect associations between IHC results, tumor phenotype, and clinical outcome.

Our data show the expected clear differences between freshly cut and old sections in the staining intensity for all five antibodies. However, despite a clearly decreased immunoreactivity, most of the expected associations between molecular findings and clinical or pathological parameters could still be found in the old section analysis. Importantly, this indicates that slides stored for a long period of time are still valuable for research purposes. The obvious ability to gather meaningful biologic information despite suboptimal staining quality also reflects a general phenomenon in IHC studies. If the same protein is studied with different antibodies, protocols, or scoring criteria, investigators often find highly discrepant data with respect to the absolute number of positive cases but a much better concordance of their results with respect to clinico-pathological associations. The use of slides with an old—but standardized—age is comparable to other protocol variations that may either increase or decrease the sensitivity of an IHC setting.

Major research efforts are warranted to seek solutions that allow a reliable IHC analysis after long-term tissue storage. The slide aging problem is likely to become more important in the future, because of the increasing importance of retrospective diagnostic molecular tissue analysis and also because of the increasing use of TMAs in research. Possible strategies to overcome the slide aging problem may include protocol adjustments, perhaps involving a harsher antigen-retrieval procedure. No attempts to improve existing protocols were made in this study, mainly because of a shortage of old TMA sections. It is also possible that not all antibodies for one protein will show a similar degree of age dependence. Bertheau *et al*² found no obvious

Figure 1 IHC staining results on old and fresh TMA sections. These images show examples of IHC stainings obtained on consecutive sections from individual tumors. (a) HER2 staining on freshly cut section. (b) HER2 staining on old section (same case as (a)). (c) HER2 staining on freshly cut section. (d) HER2 staining on old section (same case as (c)). (e) ER staining on freshly cut section. (f) ER staining on old section (same case as (e)). (g) ER staining on freshly cut section. (h) ER staining on old section (same case as (g)). (i) E-cadherin on freshly cut section. (j) E-cadherin on old section (same case as (i)). (k) E-cadherin on freshly cut section. (l) E-cadherin on old section (same case as (k)).



impact of the age of the section on the staining results of seven of their 11 analyzed antibodies. Improved methods for section storage would be another possible solution to the aging problem. However, previous studies had suggested that IHC deterioration cannot satisfactorily be prevented by section storage under paraffin coating or at low temperatures.¹⁸

Jacobs *et al*¹ showed a significant decrease of staining intensity already after 2 weeks of storage, suggesting that deterioration of IHC starts early after sectioning. All tissue samples on our 'old slides' were of identical age (6 months) in this study. This is in contrast to the situation within clinical trials where tumor tissue sections are continuously collected over a period of several months or years. It will be important to determine to what extent IHC quality decreases between 2 weeks and 6 months of storage and also after more than 6 months storage time. In the optimal case, a plateau will be reached for IHC deterioration after a certain amount of time.

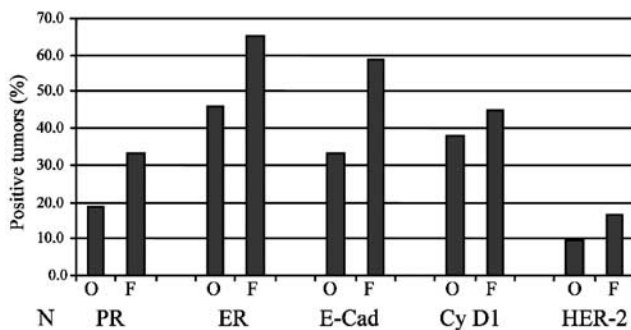


Figure 2 Influence of slide aging on the fraction of positive cases. For each antibody, the frequency of positive cases is shown as separate bars for old (O) and fresh (F) sections.

In such a scenario, the slide aging effect would be identical for all sections older than the time period in which immunoreactivity keeps decreasing. Appropriate protocol adjustments might then enable IHC analysis of old sections with sufficient quality.

TMA's constitute an optimal tool for such protocol developments. TMA's are highly suited for comparing different procedures, because of a maximal standardization of many parameters that are otherwise difficult to control. An entire TMA study does often involve the analysis of only one slide. In that case, section thickness, slide age, the type and concentration of all reagents, the temperature of incubation or during slide storage and all incubation times are absolutely identical. Even if multiple TMA sections are included in a TMA study, most of these parameters are still identical since the entire study will typically be conducted on 1 day using identical reagents. It can be speculated that other minor parameters such as section thickness or variations in incubation times that are often allowed in IHC protocols (for example: incubation 'over night') may have—at least for some antibodies—a significant influence on the outcome of IHC analyses. Overall, the influence of the protocols on IHC data is paramount. For example, the use of three different antibodies resulted in a more than fivefold difference in the positivity rate of the EGFR gene in a recent study.¹⁹ As compared to the impact of the experimental procedures on TMA results, the influence of the staining interpretation is only minimal. The small size of the tissue area to be analyzed results in a more reproducible staining interpretation of arrayed tissue samples as compared to large sections.²⁰ If IHC results are evaluated on large sections, the interpretation bias involves the selection of the 'decisive' tissue area, a step that

Table 2 Associations with phenotype and HER2/ER/PR/cyclin D1 status

	G1	G2	G3	P-value	ER+	ER-	P-value
HER2 neg (fresh)	106	125	86		216	90	
HER2 pos (fresh)	6	19	37	<0.0001	22	39	<0.0001
HER2 neg (old)	126	145	105		230	98	
HER2 pos (old)	3	7	30	<0.0001	6	31	<0.0001
ER neg (fresh)	24	39	66				
ER pos (fresh)	84	105	52	<0.0001			
ER neg (old)	69	63	92				
ER pos (old)	62	86	44	<0.0001			
PR neg (fresh)	63	94	94		128	113	
PR pos (fresh)	52	45	25	0.0004	107	15	<0.0001
PR neg (old)	98	126	112		184	115	
PR pos (old)	32	23	21	0.0903	54	12	0.0017
Cyclin D1 neg (fresh)	32	34	50		53	64	
Cyclin D1pos (fresh)	36	38	20	0.0038	85	10	<0.0001
Cyclin D1neg (old)	49	43	58		62	69	
Cyclin D1pos (old)	30	38	21	0.0287	69	10	<0.0001

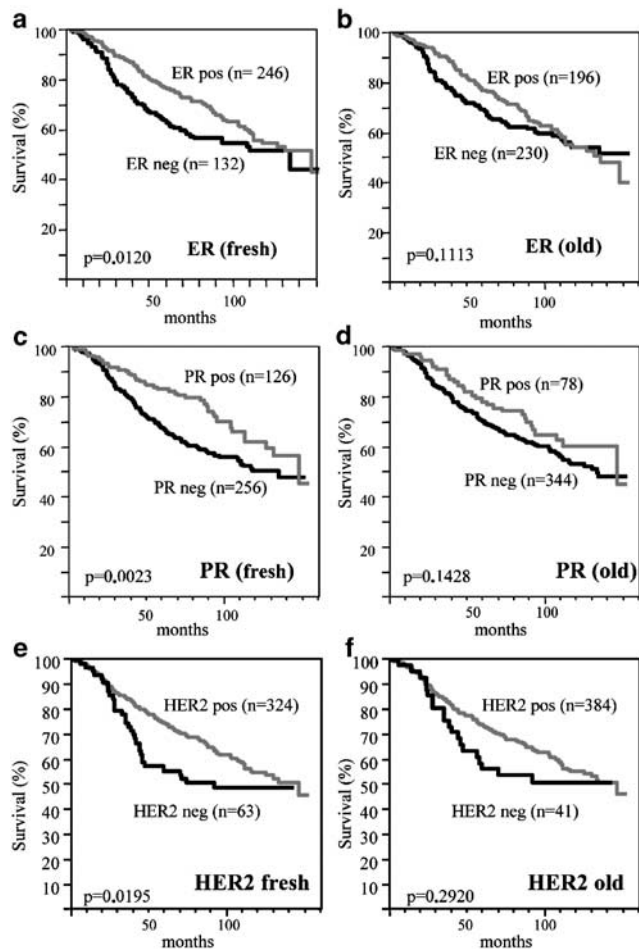


Figure 3 Influence of slide aging on survival analyses. The survival curves show the outcome of survival analyses performed with data obtained on fresh and old sections. (a) ER staining on freshly cut sections. (b) ER staining on old sections. (c) PR staining on freshly cut sections. (d) PR staining on old sections. (e) HER2 staining on freshly cut sections. (f) HER2 staining on old sections.

is excluded in the TMA setting. For example, in a study comparing p53 analysis on TMAs and on large sections of >500 breast cancers, the area selection bias had prevented us from finding prognostic associations on large sections while these were clearly identified on four different TMAs.^{19,20}

In summary, the data of this study confirm a major impact of the age of tissue sections on the outcome of IHC analyses. Importantly, such a deterioration of IHC staining intensity may be acceptable for many research projects since significant association between molecular changes and tumor phenotype can often be detected in case of suboptimal staining conditions. It is hoped that optimized staining procedures may compensate for the slide aging problem, at least in some instances. However, a decreased slide quality after long-time storage will not be acceptable if IHC results are used for clinical decision-making for individual patients. Every effort should be taken to use freshly cut

sections for potentially clinically relevant IHC analyses.

References

- Jacobs TW, Prioleau JE, Stillman IE, *et al*. Loss of tumor marker-immunostaining intensity on stored paraffin slides of breast cancer. *J Natl Cancer Inst* 1996;88:1054–1059.
- Bertheau P, Cazals-Hatem D, Meignin V, *et al*. Variability of immunohistochemical reactivity on stored paraffin slides. *J Clin Pathol* 1998;51:370–374.
- Manne U, Myers RB, Srivastava S, *et al*. Re: loss of tumor marker-immunostaining intensity on stored paraffin slides of breast cancer. *J Natl Cancer Inst* 1997;89:585–586.
- Henson DE. Loss of p53-immunostaining intensity in breast cancer. *J Natl Cancer Inst* 1996;88:1015–1016.
- Cajulis RS, Frias-Hidvegi D, Yu GH, *et al*. Detection of numerical chromosomal abnormalities by fluorescence *in situ* hybridization of interphase cell nuclei with chromosome-specific probes on archival cytologic samples. *Diagn Cytopathol* 1996;14:178–181.
- Lisowski AR, English ML, Opsahl AC, *et al*. Effect of the storage period of paraffin sections on the detection of mRNAs by *in situ* hybridization. *J Histochem Cytochem* 2001;49:927–928.
- Olapade-Olaopa EO, Ogunbiyi JO, Mac Kay EH, *et al*. Further characterization of storage-related alterations in immunoreactivity of archival tissue sections and its implications for collaborative multicenter immunohistochemical studies. *Appl Immunohistochem Mol Morphol* 2001;9:261–266.
- Prioleau J, Schnitt SJ. p53 antigen loss in stored paraffin slides. *N Engl J Med* 1995;332:1521–1522.
- Kato J, Sakamaki S, Niitsu Y. More on p53 antigen loss in stored paraffin slides. *N Engl J Med* 1995;333:1507–1508.
- Shin HJ, Kalapurakal SK, Lee JJ, *et al*. Comparison of p53 immunoreactivity in fresh-cut versus stored slides with and without microwave heating. *Mod Pathol* 1997;10:224–230.
- Vis AN, Kranse R, Nigg AL, *et al*. Quantitative analysis of the decay of immunoreactivity in stored prostate needle biopsy sections. *Am J Clin Pathol* 2000;113:369–373.
- Olapade-Olaopa EO, MacKay EH, Habib FK. Variability of immunohistochemical reactivity on stored paraffin slides. *J Clin Pathol* 1998;51:943.
- van den Broek LJ, van de Vijver MJ. Assessment of problems in diagnostic and research immunohistochemistry associated with epitope instability in stored paraffin sections. *Appl Immunohistochem Mol Morphol* 2000;8:316–321.
- Wester K, Wahlund E, Sundstrom C, *et al*. Paraffin section storage and immunohistochemistry. Effects of time, temperature, fixation, and retrieval protocol with emphasis on p53 protein and MIB1 antigen. *Appl Immunohistochem Mol Morphol* 2000;8:61–70.
- Kononen J, Bubendorf L, Kallioniemi A, *et al*. Tissue microarrays for high-throughput molecular profiling of hundreds of specimens. *Nat Med* 1998;4:844–847.
- Bubendorf L, Nocito A, Moch H, *et al*. Tissue microarray (TMA) technology: miniaturized pathology

- archives for high-throughput *in situ* studies. *J Pathol* 2001;195:72–79.
- 17 Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* 1991;1991:403–410.
 - 18 Jacobs TW, Prioleau JE, Stillman IE, *et al*. Loss of tumor marker-immunostaining intensity on stored paraffin slides of breast cancer. *J Natl Cancer Inst* 1996;88:1054–1059.
 - 19 Sauter G, Simon R, Hillan K. Tissue microarrays in drug discovery. *Nat Rev Drug Discov* 2003;2:962–972.
 - 20 Torhorst J, Bucher C, Kononen J. Tissue microarrays for rapid linking of molecular changes to clinical endpoints. *Am J Pathol* 2001;159:2249–2256.