VILNIUS UNIVERSITY
INSTITUTE OF ONCOLOGY VILNIUS UNIVERSITY

Aida Laurinaviciene

THE STUDY OF BIOLOGICAL DIVERSITY OF DUCTAL BREAST CARCINOMA BY MOLECULAR AND DIGITAL PATHOLOGY METHODS

Summary of doctoral dissertation

Biomedicine, medicine (06B),
Cytology, oncology, cancerology (B200)

Vilnius, 2012

Doctoral dissertation has been done in 2006 – 2012 Institute of Oncology Vilnius University

Scientific Supervisors:

Prof. Dr. Sonata Jarmalaite (Vilnius University, Biomedical Sciences, Biology – 01B, Cytology, Oncology, Cancerology – B200). 2006 – 2010 academic years.
Habil. Dr. Valerijus Ostapenko (Institute of Oncology Vilnius University, Biomedical Sciences, Medicine – 06B, Cytology, Oncology, Cancerology – B200). 2011 – 2012 academic years.
Scientific Consultant:
Prof. Dr. Dainius Characiejus (Vilnius University, Biomedical Sciences, Medicine – 06B, Cytology, Oncology, Cancerology – B200). 2006 – 2008 academic years.

Doctoral dissertation will be defended on the Council of Medicine Science of Vilnius University:

Chair:

Prof. Dr. Janina Didziapetriene (Institute of Oncology Vilnius University, Biomedical Sciences, Medicine – 06B, Cytology, Oncology, Cancerology – B200)

Members:

Prof. Habil. Dr. Elona Juozaityte (Lithuanian university of Health Sciences, Biomedical Sciences, Medicine – 06B, Cytology, Oncology, Cancerology – B200)
Prof. Dr. Arturas Inciura (Lithuanian university of Health Sciences, Biomedical Sciences, Medicine – 06B, Cytology, Oncology, Cancerology – B200)
Prof. Habil. Dr. Vida Kirveliene (Vilnius University, Biomedical Sciences, Biology – 01B, Cytology, Oncology, Cancerology – B200)
Prof. Dr. Dainius Characiejus (Vilnius University, Biomedical Sciences, Medicine – 06B, Cytology, Oncology, Cancerology –B200)

Opponents:

Prof. Habil. Dr. Dalia Pangonyte (Lithuanian university of Health Sciences, Biomedical Sciences, Medicine – 06B, Cytology, Oncology, Cancerology – B200)
Dr. Zivile Gudleviciene (Institute of Oncology Vilnius University, Biomedical Sciences, Medicine – 06B, Cytology, Oncology, Cancerology – B200)

The doctoral dissertation will be defended at the open session of the Medical Research Council on the 20th of April 2012, at 14.00 in Institute of Oncology Vilnius University, address: P. Baublio str. 3b., Vilnius, Lithuania.

Summary of doctoral dissertation was sent on March _____ 2012.
The doctoral dissertation is available in Vilnius University and Institute of Oncology Vilnius University Librarys.

VILNIAUS UNIVERSITETAS
VILNIAUS UNIVERSITETO ONKOLOGIJOS INSTITUTAS

Aida Laurinavičienė

DUKTALINĖS KRŪTIES KARCINOMOS BIOLOGINĖS ĮVAIROVĖS TYRIMAS MOLEKULINĖS IR SKAITMENINĖS PATOLOGIJOS METODAIS

Daktaro disertacijos santrauka

Biomedicinos mokslai, medicina (06B),
Citologija, onkologija, kancerologija (B200)

Vilnius, 2012

Disertacija rengta 2006 – 2012 metais Vilniaus universiteto Onkologijos institute

Moksliniai vadovai:

Prof. dr. Sonata Jarmalaitė (Vilniaus universitetas, biomedicinos mokslai, biologija – 01B, citologija, onkologija, kancerologija – B200). 2006 – 2010 m.m.
Habil. dr. Valerijus Ostapenko (Vilniaus universiteto Onkologijos institutas, biomedicinos mokslai, medicina – 06B, citologija, onkologija, kancerologija – B200). 2011 – 2012 m.m.

Mokslinis konsultantas:

Prof. dr. Dainius Characiejus (Vilniaus universitetas, biomedicinos mokslai, medicina – 06B, citologija, onkologija, kancerologija – B200). 2006 – 2008 m.m.

Disertacija ginama Vilniaus universiteto Medicinos mokslo krypties taryboje:

Pirmininkė:

Prof. dr. Janina Didžiapetrienė (Vilniaus universiteto Onkologijos institutas, biomedicinos mokslai, medicina – 06B, citologija, onkologija, kancerologija – B200)

Nariai:

Prof. habil. dr. Elona Juozaitytė (Lietuvos sveikatos mokslų universitetas, biomedicinos mokslai, medicina – 06B, citologija, onkologija, kancerologija – B200)
Prof. dr. Artūras Inčiūra (Lietuvos sveikatos mokslų universitetas, biomedicinos mokslai, medicina – 06B, citologija, onkologija, kancerologija – B200)
Prof. dr. Vida Kirvelienė (Vilniaus universitetas, biomedicinos mokslai, biologija – 01B, citologija, onkologija, kancerologija – B200)
Prof. dr. Dainius Characiejus (Vilniaus universitetas, biomedicinos mokslai, medicina – 06B, citologija, onkologija, kancerologija – B200)

Oponentai:

Prof. habil. dr. Dalia Pangonytė (Lietuvos sveikatos mokslų universitetas, biomedicinos mokslai, medicina – 06B, citologija, onkologija, kancerologija – B200)
Dr. Živilė Gudlevičienė (Vilniaus universiteto Onkologijos institutas, biomedicinos mokslai, medicina – 06B, citologija, onkologija, kancerologija – B200)

Disertacija bus ginama viešame Medicinos mokslo krypties tarybos posėdyje 2012 m. balandžio 20 d. 14 val. Vilniaus universiteto Onkologijos instituto, Mokslinių tyrimų centro auditorijoje, adresas: P. Baublio g 3b., Vilnius, Lietuva.
Disertacijos santrauka išsiuntinėta 2012 m. kovo mėn. _____ d.
Disertaciją galima peržiūrėti Vilniaus universiteto ir VU Onkologijos instituto bibliotekose.

CONTENT

1. ABBREVATIONS

AR	– the androgen receptor
DA	– digital image analysis
DCIS	– ductal carcinoma in situ
ER	– the estrogen receptors
FDA	– Food and Drug Administration
FISH	– fluorescence *in situ* hybridization
G	– histological grade
HER2	– the human epidermal growth factor receptor 2
HIF1α	– the hypoxia-inducible factor 1α
HR	– the hormone receptors
i-Grade	– the "immunohistochemical grade"
IHC	– immunohistochemistry
PR	– the progesteron receptors
N	– lymph node involvement
SATB1	– special AT-rich sequence-binding protein 1
T	– tumour stage
TMA/TMAs	– tissue microarray/tissue microarrays
VE	– visual evaluation

2. INTRODUCTION

The last decade was marked by intense molecular studies of breast cancer recognizing significant biological heterogeneity of the disease and leading to definition of the molecular types. This has opened new perspectives for personalized therapy and development of multiple gene expression-based systems to prognosticate the disease outcomes and assist in therapeutic decisions. Despite proven clinical utility of the systems, at least in the context of some categories of breast cancer, they remain relatively expensive, centralized and frequently require fresh frozen tumour specimens.

Due to the limitations of the molecular systems, current clinical practice of breast cancer therapy is largely based upon conventional clinical and pathologic criteria, including mainly tumour stage (T), lymph node involvement (N), histological grade (G), expression of hormone receptors (HR), and hyper-expression and amplification of human epidermal growth factor receptor 2 (HER2) in the tumour tissue. The gap between the accumulated knowledge on multiple molecular profiles of the breast cancer and common clinical practice remains open and in some way is compensated by intrinsic biological subtypes adopted by St Gallen in 2011. The subtypes may be approximated using clinicopathological rather than gene expression array criteria. Therapy recommendations follow the subtype classification: Luminal A disease generally requires only endocrine therapy, which also forms part of the treatment of the Luminal B subtype. Chemotherapy is considered for most patients with Luminal B, HER2 positive, and Triple-negative (ductal) disease, with the addition of trastuzumab in HER2 positive disease. Distinction between the Luminal A and Luminal B subtypes is based on the estimate of proliferative activity of the tumour, measured by the percentage of Ki67-positive tumour cells by immunohistochemistry (IHC).

Although the proposed approach provides a bridge between the molecular types of the disease and clinical practice, it is still largely based on semi-quantitative evaluation of estrogen receptor (ER), progesteron receptor (PR), HER2, and Ki67 expression visualized by IHC. The latter method is confined to an issue of defining and then following cut-off values which leads to misclassification of some patients, at least in borderline cases. According to the currently accepted standards, the reproducibility of

the IHC tests is suboptimal, the concordance between the methods and laboratories is below expectations for good clinical practice. The improvement in this area could come from standardizing all phases of the IHC (and HER2 FISH) tests along with application of image analysis tools to obtain more accurate, reproducible and quantitative results. In addition, digital image analysis (DA) providing continuous data of the IHC biomarker expression is an important prerequisite to apply more powerful mathematical analysis tools for tissue-based biomarker research.

In the view of urgent need to improve prognostic classifiers in breast cancer, efforts are being made to use a combinatorial approach revealing new aspects of the disease and promising more reliable stratification of the risk based on combined biomarkers rather than single ones. In essence, it corresponds to the multivariate analysis approach used to develop multiple gene expression-based systems. It has been shown that similar information can be obtained by a combination of relevant IHC markers, including the heterogeneity of the disease revealed by cluster analysis. However, combined IHC biomarkers proposed up-to-the-date are mostly based on a combination of several biomarkers evaluated qualitatively or semi-quantitatively. Although clinical utility can be achieved already, it is important to employ multivariate analysis methods to exploit broad dynamic range of the IHC DA data. An important exploratory step of the investigation is delivered by factor analysis, revealing independent factors of variation in the data set of multiple IHC biomarkers. Multidimensional data space reduction and extraction of latent variability factors may uncover true biological meaning and informative value of single biomarkers and provide integrated factor scores as quantitative estimates of the biological processes. This may be the only right approach since one biomarker can reflect several biological processes and have different roles in different disease entities. Furthermore, most robust prognostic factors are likely to come in a form of integrated metamarkers derived by multivariate analysis of multimodal data of various aspects (clinical, pathology, molecular, imaging, etc.) of the disease.

In our study, we performed an automated image analysis on a set of 10 IHC markers, including the conventional ER, PR, HER2, and Ki67 along with less investigated androgen receptor (AR), BCL2, HIF1α, SATB1, p53 and p16 on tissue microarrays (TMAs) of 109 patients with ductal carcinoma of the breast. We present the potential of

factor analysis of the IHC marker expression data set to reveal biologically and clinically meaningful interdependencies of the breast cancer immunophenotype.

THE AIM

To explore biological diversity of the primary ductal breast carcinoma by molecular and digital pathology techniques and associated clinical features of the disease.

OBJECTIVES

1. To apply the digital image analysis for quantitative assessment of protein expression by immunohistochemistry and evaluate the validity of this method compared to a pathologist's visual evaluation.
2. To investigate the expression of ER, PR, HER2, Ki67, p53, AR, p16, BCL2, SATB1, HIF1α proteins, frequencies of *p14*, *p16*, *RARB*, *RASSF1*, *ESR1*, *DAPK1*, *GSTP1* ir *MGMT* gene hypermethylation, *TP53* gene mutation, and *HER2* gene amplification in the tumour tissue of ductal carcinoma of the breast.
3. To associate the ductal breast carcinoma biological profiles, established by molecular and digital pathology methods, with clinical characteristics of the disease.

3. MATERIALS AND METHODS

3.1. STUDY POPULATION AND CLINICAL METHODS

Tumour samples were prospectively collected from 203 patients with an invasive ductal carcinoma of the breast treated at the Institute of Oncology Vilnius University and investigated at the National Center of Pathology during the period of 2007 to 2009. The patients and tumour characteristics are presented in Table 1. Informed consent was obtained and documented in writing before study entry. The study was approved by the Lithuanian Bioethics Committee.

Table 1. Patient and tumour characteristics (all group)

Age group	Frequency	%
Age < 55 year	100	49.3
Age > 55 year	103	50.7
T		
1	107	52.7
2	96	47.3
N		
0	104	52.5
1	64	32.3
2	21	10.6
3	9	4.6
M		
0	203	100
Histological grade		
1	40	19.7
2	79	38.9
3	84	41.4

The tissue microarrays were constructed from 10% buffered formalin-fixed paraffin-embedded tissue blocks. One millimetre-diameter cores were punched from tumour areas randomly selected by pathologist (4 cores per patient), thus producing 11 tissue microarrays constructed using the tissue arraying instrument (3DHISTECH, TMA Master, Budapest, Hungary). Paraffin sections of the tissue microarrays were cut for IHC (3 μm-thick) and HER2 FISH testing (4 μm-thick).

Immunohistochemistry for ER, PR, HER2, AR, Ki67, p53, p16, BCL2, SATB1 and HIF1α was performed.

Digital images were captured using the Aperio ScanScope XT Slide Scanner (Aperio Technologies, Vista, CA, USA) under 20x objective magnification. All TMA spots were evaluated on the monitor visually by the pathologist, providing semi-quantitative estimates of the percentage of positive cells and excluding the spots containing inadequate tumour sample or DCIS from further analyses.

3.2. DIGITAL IMAGE ANALYSIS

The DA was performed on the same images as the visual evaluation. Aperio Genie Classifier was trained to recognize tumour tissue, stroma and background (glass). The Genie classifier was then combined with Aperio Membrane v9 and Aperio Nuclear v9

algorithms. The percentage of tumour cells with complete membranous (HER2 and BCL2) staining and positive nuclear (ER, PR, AR, Ki67, p53, p16, SATB1 and HIF1α) staining was used for further analyses. The examples of IHC and DA analysis output images are presented in Fig. 1. The data from all adequate TMA spots were summarized (positive and total cells in the spots were summed, then the percentage of positive cells calculated) into one estimate per patient with a threshold of total number of tumour cells per patient set at >500. A total of 109 patients with a complete set of 10 IHC markers remained for multivariate analyses.

We performed our study to test the robustness of HER2 IHC scoring based on a novel membrane connectivity estimate in TMA of breast cancer tissue. The digital analysis results were matched with the data of visual evaluation of HER2 by IHC and HER2 FISH test results on the same TMA. The pathologist intra-observer agreement (VE1 and VE2) on HER2 IHC score was almost perfect: kappa 0.91 (per spot) and 0.88 (per patient). The agreement between visual evaluation and digital image analysis was almost perfect at the spot level (kappa 0.86 and 0.87, with VE1 and VE2 respectively) and at the patient level (kappa 0.80 and 0.86, with VE1max and VE2max, respectively). The DA was more accurate than VE in detection of FISH-positive patients by recruiting 3 or 2 additional FISH-positive patients from the IHC 0/1+ category by VE1max or VE2max, respectively. The DA cont inuous output variable correlated wi th the FISH data (HER2 and CEP17 copy numbers, and HER2/CEP17 ratio).

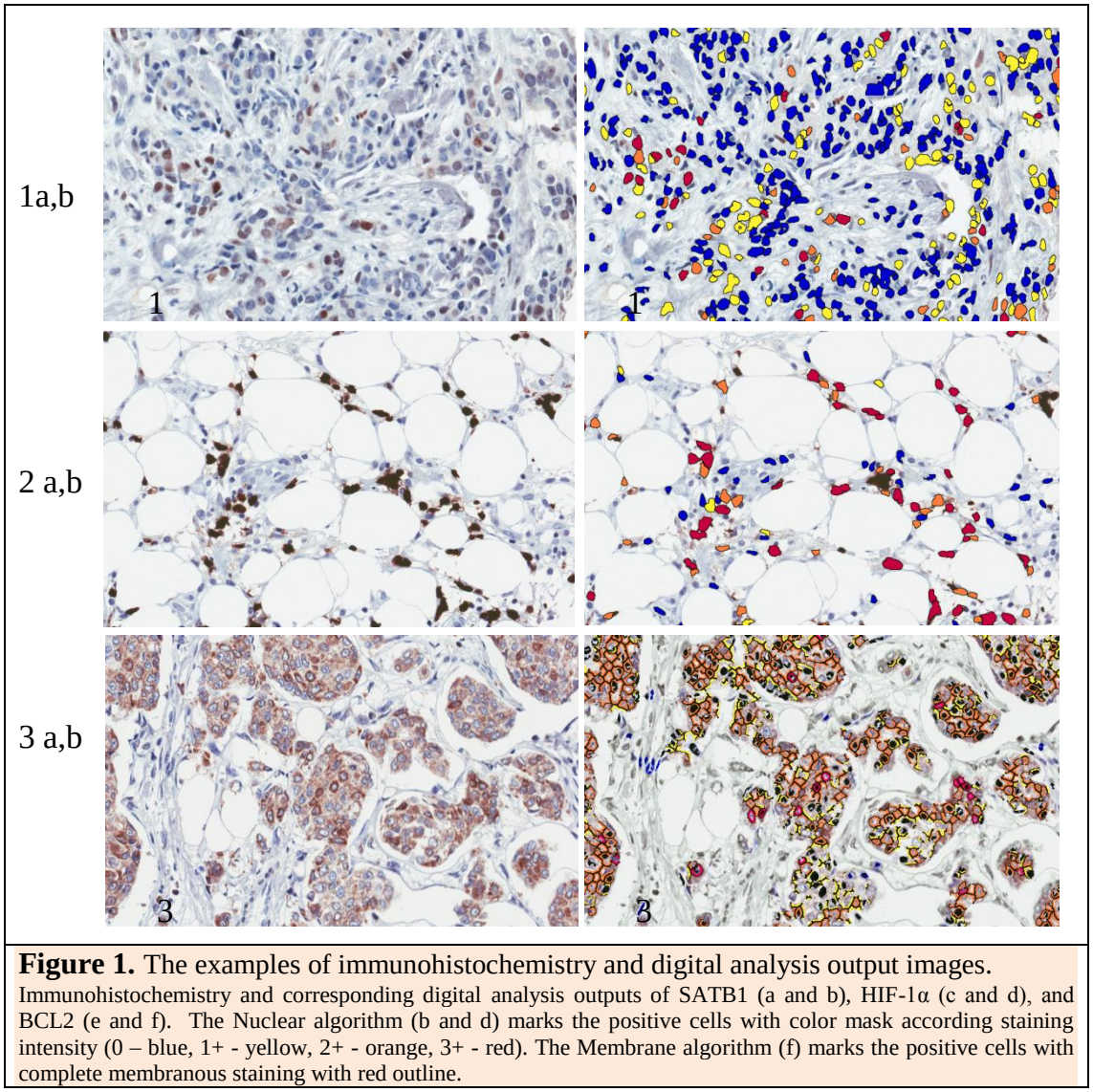

Figure 1. The examples of immunohistochemistry and digital analysis output images.
Immunohistochemistry and corresponding digital analysis outputs of SATB1 (a and b), HIF-1α (c and d), and BCL2 (e and f). The Nuclear algorithm (b and d) marks the positive cells with color mask according staining intensity (0 – blue, 1+ - yellow, 2+ - orange, 3+ - red). The Membrane algorithm (f) marks the positive cells with complete membranous staining with red outline.

3.3. STATISTICAL ANALYSIS

Summary statistics and distribution analyses were performed with significance tests based on one-way ANOVA and Bonferroni (Dunn) t Tests for pairwise comparisons. Since distributions of HER2, Ki67, HIF1α, SATB1, p53, and p16 DA results revealed left asymmetry, logarithm-transformed values were used for parametric statistics. For the sake of readability, the prefix "log" is not used in the text or graphs when referring to these markers. Factor analysis on a DA data set of 10 IHC markers was performed using factoring method of principal component analysis. Five factors were retained based on the threshold of the smallest eigenvalue of 0.84. General orthomax rotation of the initial factors was performed. Factor analyses were performed in two sets of patients: the

whole group of ductal carcinoma (n=109) and HR-positive ductal carcinoma (n=85) including Luminal A, B, and B HER2 positive tumours. Pearson's correlation was performed to test the pairwise linear relationships between the continuous variables as a preparatory step for factor analyses. Chi-square test and Fisher's exact test were used to estimate significant associations in non-parametric statistics. Statistical significance level was set at $p<0.05$. Statistical analysis was performed with SAS 9.2 software.

4. RESULTS

4.1. PATIENT AND TUMOUR CHARACTERISTICS

Patient and tumour characteristics are presented in the Table 2. Since the intrinsic subtypes were subdivided based on the visual evaluation of the IHC images, the DA results on ER, PR, HER2, and Ki67 do not strictly correspond to the conventional cut-off values used for the definition of intrinsic subtypes.

Table 2. Patient and tumour characteristics

	Intrinsic subtype					
	Luminal A	Luminal B	Luminal B HER2+	HER2+	Triple-negative	p
Age group						n.s
Age < 55 year (n=52)	17 (40)	14 (52)	8 (53)	1 (17)	12 (67)	
Age > 55 year (n=57)	26 (60)	13 (48)	7 (47)	5 (83)	6 (33)	
Histological grade						<0.0001
1	16 (37)	2 (8)	1 (7)	0 (0)	0 (0)	
2	26 (61)	6 (22)	6 (40)	3 (50)	3 (17)	
3	1 (2)	19 (70)	8 (53)	3 (50)	15 (83)	
T						n.s.
1	28 (65)	13 (48)	8 (53)	3 (50)	6 (33)	
2	15 (35)	14 (52)	7 (47)	3 (50)	12 (67)	
N						n.s.
0	27 (63)	13 (48)	8 (53)	4 (67)	14 (78)	
1	16 (37)	14 (52)	7 (47)	2 (33)	4 (22)	
% positive cells by immunohistochemistry measured by digital image analysis (mean ± SD)*						
ER	80 ± 13	62 ± 33	52 ± 26	2 ± 1	4 ± 7	<0.0001
PR	53 ± 31	38 ± 36	19 ± 29	2 ± 2	3 ± 4	<0.0001
AR	47 ± 20	32 ± 23	28 ± 21	33 ± 15	10 ± 15	<0.0001
BCL2	56 ± 11	46 ± 24	33 ± 27	7 ± 6	18 ± 15	<0.0001
HER2	7 ± 11	7 ± 12	37 ± 25	64 ± 19	1 ± 3	<0.0001
Ki67	14 ± 7	40 ± 17	22 ± 12	31 ± 15	53 ± 16	<0.0001
p53	13 ± 16	34 ± 32	19 ± 18	17 ± 26	44 ± 35	n.s.
p16	14 ± 8	14 ± 12	14 ± 7	10 ± 4	40 ± 21	<0.0001
HIF1α	9 ± 6	12 ± 10	12 ± 10	18 ± 13	18 ± 10	<0.005
SATB1	12 ± 7	14 ± 10	13 ± 10	10 ± 4	19 ± 18	n.s.

* Statistical significance of variation between the groups tested by one-way ANOVA (logarithm-transformed values of HER2, Ki67, p53, p16, HIF1α, SATB1 were used for the analysis, however, original values are presented in the table).

4.2. FACTOR ANALYSIS OF THE IMMUNOPHENOTYPE OF THE DUCTAL CARCINOMA OF THE BREAST

Factor analysis was performed on 109 patients with a complete set of 10 IHC markers: ER, PR, AR, HER2, BCL2, Ki67, HIF1α, SATB1, p53, and p16. Altogether the five factors explained 80% of the variance in the data set.

Factors 1 and 2 represented major portion of the variance explained by the five factors extracted (43.9 and 15.7%, respectively). Factor loadings of the factors 1 and 2 are plotted on the Fig 2. The factor 1 is characterized by strong positive loadings of HR (ER, PR, AR) and BCL2 as well as strong negative loadings of Ki67 and HIF-1α. Based on the known biological and prognostic information conveyed by these IHC markers in the context of breast marker and their strong association to the histological grade in our study, this factor pattern can be interpreted as representing a spectrum of "the immunohistochemical grade" (i-Grade): from the tumours with predominant expression of HR and BCL2 (i-Grade-Low) to the tumours with predominant expression of Ki67 and HIF-1α (i-Grade-High).

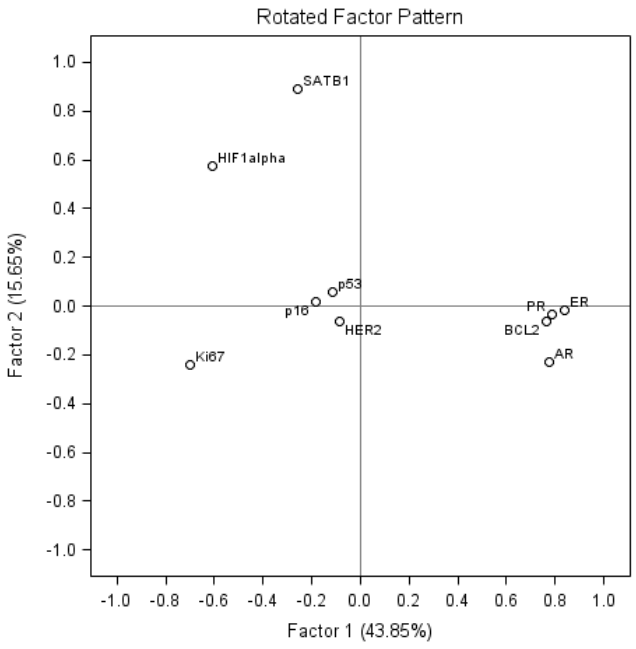

Figure 2. Rotated factor pattern in the patients with breast ductal carcinoma: loadings of the factors 1 and 2 plotted.
The loadings of the factor 1 (i-Grade) and factor 2 (SATB1/HIF1α) plotted.

Factor 2 was characterized by strong positive loadings of SATB1 and HIF1α (factor loading 0.89 and 0.58, respectively) and was labelled as "SATB1/HIF1α" (Fig. 2).

Factors 3, 4 and 5 altogether represented the remaining 40.5% (14.6, 13.5, and 12.4%, respectively) of the variance explained by the five factors extracted. The factors were characterized by positive loadings of single biomarkers and named accordingly: factor 3 (HER2), factor 4 (p53), and factor 5 (p16).

4.3. ASSOCIATIONS BETWEEN THE FACTOR 1 AND 2 SCORES AND THE CONVENTIONAL CATEGORIES OF THE DUCTAL CARCINOMA OF THE BREAST

We explored potential associations between the factor 1 and 2 score categories and the conventional characteristics of the disease: the intrinsic subtype, histological grade (G), tumour stage (T), node status (N), and age group (Table 4).

Table 4. Associations between the factor 1 and 2 scores and the conventional categories of the ductal carcinoma of the breast

	n	i-Grade-High	SATB1/HIF 1α -High	HER2-High	p53-High	p16-High
Number of patients	109					
Age group		n.s.	n.s.	p<0.03	p<0.003	n.s.
Age < 55 year	52	19 (37)	28 (54)	20 (38)	39 (75)	28 (54)
Age > 55 year	57	17 (30)	29 (51)	34 (60)	27 (47)	30 (53)
Histological grade		p<0.00001	n.s.	n.s.	n.s.	n.s.
1	19	0 (0)	12 (63)	10 (53)	7 (37)	12 (63)
2	44	9 (21)	26 (59)	27 (61)	30 (68)	22 (50)
3	46	27 (59)	19 (41)	17 (37)	29 (63)	24 (52)
T		n.s.	n.s.	n.s.	n.s.	n.s.
1	58	16 (28)	31 (53)	33 (57)	37 (64)	27 (47)
2	51	20 (44)	26 (51)	21 (41)	29 (57)	31 (61)
N		n.s.	n.s.	n.s.	n.s.	n.s.
0	66	25 (38)	36 (55)	30 (45)	40 (61)	35 (53)
1	43	11 (26)	21 (49)	24 (56)	26 (60)	23 (53)
Intrinsic subtype		p<0.0001	n.s.	p<0.00001	n.s.	p<0.002
Luminal A	43	0 (0)	29 (67)	25 (58)	23 (53)	26 (60)
Luminal B	27	7 (26)	10 (37)	10 (37)	19 (70)	8 (30)
Luminal B HER2+	15	6 (40)	7 (47)	12 (80)	10 (67)	8 (53)
HER2+	6	6 (100)	2 (33)	6 (100)	2 (33)	1 (17)
Triple negative	18	17 (94)	9 (50)	1 (6)	12 (67)	15 (83)

The factor 1 (i-Grade) was associated with the intrinsic subtypes ($p<0.0001$): all cases of HER2-positive (n=6) and all but one (94%) Triple-negative carcinoma fell into the i-Grade-High category. Significant proportion of Luminal B (26%) and Luminal B HER2 positive (40%) cases but none of Luminal A type (n=43) were i-Grade-High. Association between the i-Grade and the G was highly significant ($p<0.0001$, Fig. 4): all cases of G1 were i-Grade-Low (n=19) whereas G2 and G3 were increasingly i-Grade-High (21%, and 58%, respectively). The i-Grade was not significantly associated with T, N, or patient age group.

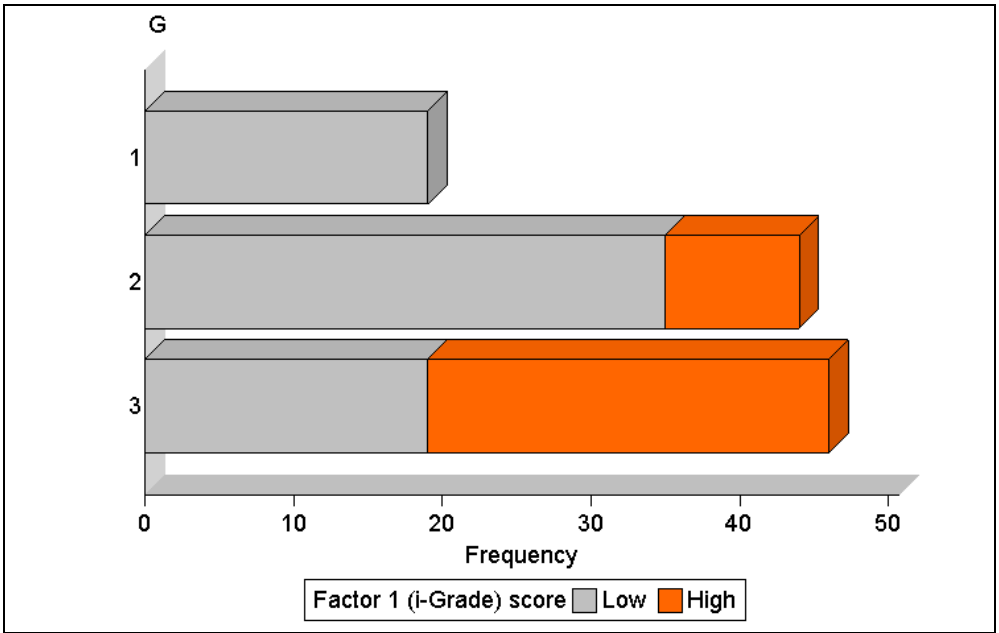

Figure 4. Association between the i-Grade and the histological grade (G). The bar chart represents the distribution of i-Grade-Low (grey) and i-Grade-High (orange) tumours against the histological grade (G1, 2, and 3).

Factor 3 (HER2) was associated with older age and presented relevant associations with the intrinsic subtypes. Factor 5 (p16) was remarkable presenting with high values in the majority (83%) of TN cases. Factor 2 (SATB1/HIF1α) and factor 4 (p53) were not significantly associated with any of the categories tested.

The IHC profile of the intrinsic subtypes was further highlighted by one-way ANOVA with the factor scores used as dependent variables. Summary of the profiles is plotted on Fig. 5.

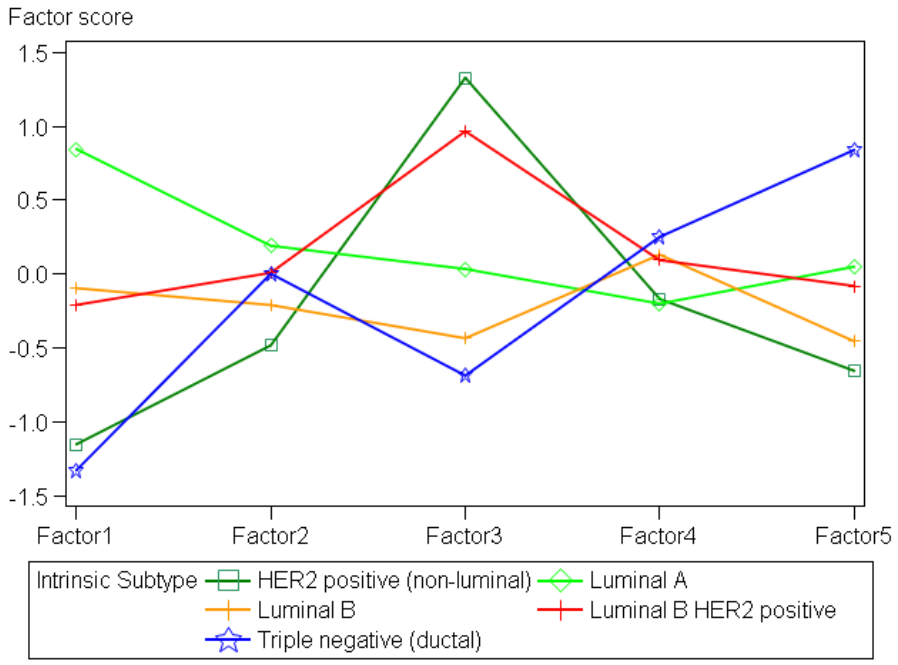

Figure 5. Representation of the factor score profiles of the intrinsic subtypes.
The multiple line chart outlines the factor score profiles of the intrinsic subtypes. The lines connect the mean values of the factor 1-5 scores obtained by one-way ANOVA for each intrinsic subtype as explanatory variable.

4.4. FACTOR ANALYSIS OF THE IMMUNOPHENOTYPE OF THE HR-POSITIVE DUCTAL CARCINOMA OF THE BREAST

To explore the immunophenotype interactions in the group of HR-positive tumours, we performed factor analysis in 85 cases of Luminal A, Luminal B, and Luminal B HER2 positive. Altogether the five factors explained 77% of the variance in the data set.

The factor 1 (31.1% of variation) resembled the factor 2 (SATB1/HIF1α) in the whole group of the patients, however, the factor pattern was different: it was characterized by strong positive loading of HIF1α, identical to that of SATB1, and by moderate negative loadings of AR, ER, and BCL2. This pattern suggests inverse relation between SATB1 and HIF1α co-expression and the co-expression of AR, ER, and BCL2 in the subgroup of HR-positive tumours. Since the biological meaning of this interrelation is not clear, this factor is labelled "SATB1/HIF1α–AR/ER/BCL2".

The factors 2 and 3 contributed a similar proportion of variation (19.8 and 19.7%, respectively) to the data set. Factor 2 was characterized by positive HER2 (0.87) and negative PR (-0.68) loadings, while factor 3 – by positive Ki67 (0.83) and negative ER

(-0.60) loadings. Respectively, these factors were labelled “HER2-PR” and “Ki67-ER”. Factors 4 (p53) and 5 (p16) resembled those already identified in the whole group of patients and contributed 15.7 and 13.8% of variation, respectively.

4.5. ASSOCIATIONS BETWEEN THE FACTOR SCORES AND THE CONVENTIONAL CATEGORIES OF THE HR-POSITIVE DUCTAL CARCINOMA OF THE BREAST

Factor 1 (SATB1/HIF1α–AR/ER/BCL2) score categories were not associated with the patient’s age group, T, N, G, or intrinsic subtype. Factor 2 (HER2-PR) score High category was associated with the older patients’ age group ($p<0.002$) and intrinsic subtype ($p<0.05$).

Associations of the factor 3 (Ki67-ER) score categories closely resembled those of the factor 1 (i-Grade) in the whole group of patients: high scores were increasing with the histological grade ($p<0.0001$, Fig. 8) and more frequently found in Luminal B subtypes ($p<0.0001$). Remarkably, some Ki67-ER-High tumours were detected in both G1 (3/19, 16%) and Luminal A (9/43, 21%) categories.

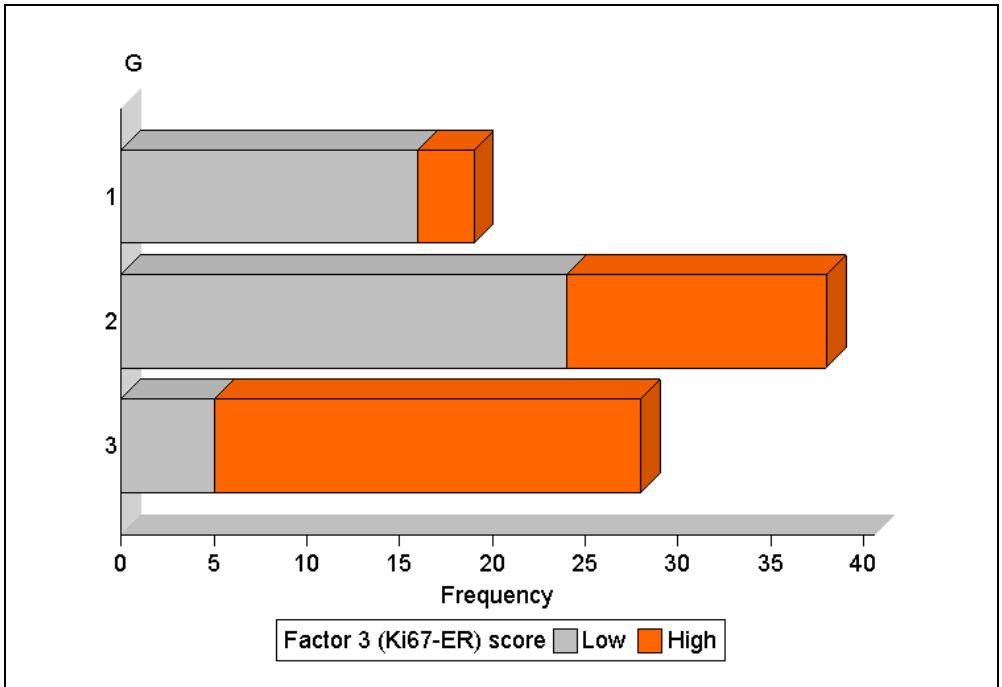

Figure 8. Association between the Ki67-ER and the histological grade (G).
The bar chart represents the distribution of Ki67-ER-Low (grey) and Ki67-ER-High (orange) tumours against the histological grade (G1, 2, and 3).

5. DISCUSSION

Our study proves that important biological interdependencies can be detected at the level of tumour tissue immunophenotype based on the multivariate analysis of DA data. In a cohort of 109 patients with ductal carcinoma of the breast, we were able to detect biologically relevant interdependencies and heterogeneity largely reflecting the main intrinsic subtypes of the disease and providing new data and insights into the breast cancer biology.

The design of our study enabled us to avoid significant human impacts and assumptions while obtaining the results: we performed an automated image analysis with automated detection of tumour tissue of the ductal carcinoma of the breast TMAs stained for 10 IHC markers, followed by factor analysis of the data set. In some sense, our results represent an automated readout of the IHC data in the TMAs. Factor analysis revealed latent factors governing the interdependent variance of the immunophenotype: being orthogonally independent by definition, the factors can be seen as independent biological processes standing behind the IHC profile variability in the disease entities. We then produced integral characteristics (factor scores) for individual patients and tested their associations with main conventional categories of the breast ductal carcinoma.

The factors of the immunophenotype variance, established in our study, are in line with the current knowledge of breast cancer biology, however, new insights emerge.

6. CONCLUSIONS

1. Digital image analysis is a robust tool for the IHC testing and can be useful for high throughput research, quality assurance, and computer-assisted diagnosis:
   1.1. the percentage of Ki67-positive tumor cells detected by digital image analysis

and the pathologist's visual evaluation revealed very strong correlation ($r=0.86$, $p<0.05$), however, digital analysis but not the pathologist's results showed no significant bias when compared to the stereologically obtained reference values;

1.2. validation set for membranous expression of HER2, showed almost perfect agreement (kappa=0.86) between the digital and visual evaluation, whereas digital analysis enabled more accurate detection of the amplified (HER2 FISH) cases.

2. Based on immunohistochemical evaluation of ER, PR, HER2, Ki67 expression, 203 ductal breast carcinoma cases were categorized into 5 intrinsic biological subtypes: 89 (44%) Luminal A, 41 (20%) Luminal B 25 (12%), Luminal B HER2+, 13 (6%) HER2+, and 35 (17%) triple-negative tumors. The subtypes revealed biological variation of the disease:

   2.1. statistically significant associations were found between the biological subtype and stage of disease ($p <0.00002$), histological grade ($p <0.0001$) and regional lymph node metastasis ($p<0.0002$);

   2.2. the biological subtypes were characterized by variable ER, PR, AR, HER2, Ki67, BCL2, HIF1a, SATB1, p53, and p16 expression and unique epigenetic/genetic profile, significantly associated with the subtype.

3. The factor analysis of multiple IHC biomarkers measured by automated DA is an efficient exploratory tool clarifying complex interdependencies in the breast ductal carcinoma IHC profiles:

   3.1. the major factor of the aggressive disease behaviour (i-Grade) was characterized by opposite loadings of ER/PR/AR/BCL2 and Ki67/HIF-1α. The i-Grade factor scores represent integral quantitative characteristics that revealed bimodal distribution and were strongly associated with the histological grade and more aggressive intrinsic subtypes;

   3.2. SATB1 was closely associated with HIF1α in HR-positive and in the whole cohort of ductal carcinoma. This implies that SATB1 and HIF1α may be important markers of the disease, whereas their biological and clinical significance remains to be elucidated. HIF1α and SATB1 may convey important biological messages other than the aggressiveness of the disease reflected by Ki67 expression and histological grade, at least in HR-positive disease;

3.3. factor analysis, performed in the HR-positive subgroup of cases, extracted factor 3 (Ki67-ER) resembling the i-Grade by its associations to the histological grade and the intrinsic subtypes. These indices may provide clinically useful stratification supplementing prognostic groups currently used;

3.4. in HR-positive tumours, factor analysis revealed an inverse relation between HER2 and PR expression which, along with the inverse relation between Ki67 and ER, may shed the light into the differential information conveyed by the ER and PR expression.

7. RECOMMENDATIONS

1. The our study demonstrates the potential of biomarker evaluation in tissue microarrays, based on digital image analysis, enabling better accuracy, precision, and efficiency. We to use these methods for quality assurance and desicion-support in clinical research and practice. Validation of the digital analysis tools should be performed not only based on pathologist‘s evaluation but also using reference values obtained by stereologic measurements.

2. We suggest introducing into clinical practice integrated IHC biomarkers, based on multivariate analysis of the digital analysis data, taking into account intrinsic interdependencies of the biomarker expression and opening new opportunities to prognosticate the disease.

8. PUBLICATIONS

1. Laurinavicius A., **Laurinaviciene A**, Ostapenko V, Dasevicius D, Jarmalaite S, Lazutka J. Immunohistochemistry profiles of breast ductal carcinoma: factor analysis of digital image analysis data. *Diagn Pathol* 2012, *in press*.
2. Laurinavicius A, **Laurinaviciene A**, Dasevicius D, Elie N, Plancoulaine B, Bor C, Herlin P. Digital image analysis in pathology: Benefits and obligations. *Anal Cell Pathol* 34 (2011) 1–4.
3. **Laurinaviciene A**, Dasevicius D, Ostapenko V, Jarmalaite S, Lazutka J, Laurinavicius A. Membrane connectivity estimated by digital image analysis of HER2 immunohistochemistry is concordant with visual scoring and fluorescence in situ hybridization results: algorithm evaluation on breast cancer tissue microarrays. *Diagn Pathol* 2011, 6(1):87.
4. Jarmalaite S., **A. Laurinaviciene**, J. Tverkuviene, N. Kalinauskaite, D. Petroska, T. Bohling, K. Husgafvel-Pursiainen: Tumor suppressor gene ZAC/PLAGL1: altered expression and loss of the nonimprinted allele in pheochromocytomas. *Cancer Genetics* 2011, 204(7):398-404.
5. J. Tverkuviene, **A. Laurinaviciene**, K. Daniunaite, A. Scesnaite, S. Jarmalaite: Frequent aberrant DNA methylation of CDKN2A locus in capillary hemangioblastomas, pheochromocytomas and gliomas. *Acta Medica Lithuanica*, 2011. Vol. 18. No. 1. P. 4–11.
6. **A. Laurinaviciene**, D. Petroska, J. Tverkuviene, K. Daniunaite, A. Scesnaite, S. Jarmalaite: Frequent aberrant expression of p53 protein in gliomas, but not in capillary hemangioblastomas and pheochromocytomas. *Acta Medica Lithuanica*, 2009 Vol. 16. No. 3–4. P. 130–138.

ORAL PRESENTATIONS

1. PhD students internship project "The promotion of student's research activities": The study of tissue heterogeneity in early breast cancer, 2011-03-17.
2. Molecular pathology for diagnosis and treatment: The use of tissue microarrays for breast cancer research – experience at the National Center of Pathology, 2010-02- 26.
3. Pathology diagnosis and treatment of early breast cancer: Molecular characterization of early breast cancer: preliminary results of ongoing collaborative research project, 2009-11-27.
4. Treatment of tyrosine kinase *ErbB-2* (*HER2*) receptor positive breast cancer: *ErbB* receptors and intracellular signaling biology, 2009-02-06.
5. Precise diagnosis of breast cancer before the treatment: prognostic and predictive factors. Molecular markers in breast cancer tissue, 2009-04-08.

POSTER PRESENTATIONS

1. S. Jarmalaite, **A. Laurinaviciene**, J. Dimonaite, A. Scesnaite, D. Dasevicius, V. Ostapenko, J. Lazutka: Epigenetic and genetic biomarkers, of early-stage breast cancer. *Cellular Oncology*, 30 (3): 259-260, 2008.
2. Dimonaite, **A. Laurinaviciene**, D. Petroska, N. Kalinauskaite, A. Scesnaite, S. Jarmalaite: Epigenetic changes in pathogenesis of VHL-related tumours and primary gliomas. *Cellular Oncology*, 30 (3): 254-255, 2008.

9. SUMMARY IN LITHUANIAN

DARBO TIKSLAS

Taikant molekulinės ir skaitmeninės patologijos tyrimo metodus ištirti pirminių krūties duktalinių karcinomų biologinių pokyčių spektrą, nustatytus pokyčius susieti su klinikinėmis-patologinėmis ligos charakteristikomis.

DARBO UŽDAVINIAI

1. Pritaikyti skaitmeninės imunohistocheminių vaizdų analizės metodą kiekybinei baltymų raiškos analizei atlikti ir įvertinti metodo patikimumą rezultatus lyginant su patologo vizualaus vertinimo rezultatais.
2. Duktalinėse krūties karcinomose ištirti baltymų ER, PR, HER2, Ki67, p53, AR, BCL2, p16, SATB1, HIF1α raišką, genų *ESR1*, *DAPK1*, *GSTP1*, *MGMT, p14, p16, RARB ir RASSF1* hipermetilinimo, geno *HER2* amplifikacijos bei geno *TP53* mutacijos dažnį.
3. Molekulinės ir skaitmeninės patologijos metodais nustatytą duktalinės krūties karcinomos biologinius profilius susieti su klinikiniais rodikliais ir nustatyti tirtų žymenų prognozinę vertę.

TYRIMO METODIKA

Darbe naudojant skaitmeninį vaizdo analizės metodą tirta 10 imunohistocheminių žymenų, apimant įprastinius, klinikinėje praktikoje naudojamus žymenis – ER, PR, HER2 ir Ki67 bei mažiau tirtus – AR, BCL2, HIF1α, SATB1, p53 ir p16, žymenis. Siekiant geriau suprasti krūties navikų molekulinius patogenezės mechanizmus darbe buvo tiriami ir genetiniai – *TP53, HER2* bei epigenetiniai pakitimai – *p14, p16, ESR1, RARB, RASSF1, DAPK1, GSTP1, MGMT* genuose, susijusiuose su ląstelės atsparumu vėžiui, su ląstelės ciklo ir kitų ląstelės atsakų valdymu.

Skaitmeninei IHC analizei atlikti naudotas *AperioGenie* membraninis ir *AperioGenie* branduolinis algoritmai. Siekdami kalibruoti pasirinktus analizės algoritmus, jie lyginti su kitomis to paties objekto matavimo sistemomis. Skaitmeninės branduolinių IHC

žymenų analizės rezultatai lyginti su stereologinės skaitmeninio vaizdo analizės rezultatais, membraninių IHC žymenų analizė – su kito gamintojo skaitmeninės analizės bei FISH rezultatais. Visų analizių rezultatai palyginti tarpusavyje bei su gydytojo patologo vizualaus vertinimo rezultatais.

REZULTATAI

Atsižvelgiant į patologo ER, PR, HER2 ir Ki67 IHC vizualaus vertinimo rezultatus, tiriamieji mėginiai suskirstyti į biologinius tipus. Apie du trečdaliai navikų buvo luminalinio A (LA; 44%) ir luminalinio B (LB; 20%) tipo navikai. Kitą duktalinių karcinomų grupę sudarė HER2 teigiami navikai: luminalinio B tipo (LB HER2+; 12%) ir ne luminalinio tipo (HER2+; 6%) navikai. Trejopai neigiamų (TN; ER-, PR-, HER2-) navikų kategoriją sudarė 35 (17%) pacienčių mėginiai. Nustatytas statistiškai reikšmingas skirtumas tarp biologinio naviko tipo ir ligos stadijos (p=0,0001). Blogai diferencijuotų navikų (G3) daugiausia nustatyta tarp trejopai neigiamų navikų (86%). G3 dominuoja ir tarp ne luminalinio HER2+ (69%) bei luminalinio B tipo navikų (64%). Tarp luminalinio A tipo navikų G3 nustatytas tik 6 (7%) pacientėms.

TN ir HER2+ navikai dažniau buvo T2 stadijos (atitinkamai 63% ir 62%), o luminalinio A ir luminalinio B HER2+ tipo navikai, turintys hormonų receptorių raišką – T1 stadijos (atitinkamai 61% ir 60%). Nustatytas statistiškai reikšmingas skirtumas tarp naviko biologinio tipo ir T (p=0,00002).

Luminalinis A ir luminalinis B tipai išsiskyrė dažna BCL2 (95% ir 73%, 11 lentelė) bei reta HIF1α (8% ir 20%) žymenų raiška (p<0,0001), o luminalinis A tipas dar pasižymėjo ir dažna AR raiška (72%, p<0,0001). HER2 teigiamų navikų grupėje (LB HER2+ ir HER2+) p53 baltymo raiška konstatuota atitinkamai 74% ir 64%, tačiau statistiškai reikšmingo p53 baltymo raiškos skirtumo pagal naviko biologinius tipus nenustatyta. Pažymėtina, kad HER2+ navikų grupėje skirtinga BCL2 ir HIF1α žymenų raiška: luminalinio B HER2+ navikų grupėje BCL2 raiška yra dažna (50%), HIF1α - reta (13%), HER2+ grupėje BCL2 raiška – reta (8%), HIF1α – dažna (73%). Nustatyti skirtumai yra statistiškai reikšmingi (p<0,0001). Trejopai neigiamų navikų grupėje

nustatyta dažna p16 (79%, p<0,0001) žymens raiška. Šioje navikų kategorijoje AR (10%) ir BCL2 (23%) žymų raiška yra reta.

Dažniausiai hipermetilinami genai duktalinėse krūties karcinomose buvo *RASSF1* (69%) ir *ESR1* (pirmas fragmentas – 50%, antras – 64%). Kiek rečiau hipermetilinami buvo genų *MGMT* (25%), *RARB* (20%), *GSTP1* (19%) ir *DAPK*1 (15%) promotoriai. Ypač retai epigenetinės pažaidos vyko genų *p16* (5%) ir *p14* (2%) promotoriuose. Nustatytos 28% navikų geno *TP53* mutacijos. Nustatytos statistiškai reikšmingos sąsajos tarp krūties navikų biologinių tipų ir geno *ESR1-4*, *p16* hipermetilinimo bei geno *TP53* mutacijos.

Dažniausiai hipermetilintų genų *ESR1* ir *RASSF1* pakitimų dažnis siekė 86%, o kartu su geno *RARB* hipermetilinimu – 90%.

IŠVADOS

1. Skaitmeninė vaizdo analizė yra patikimas būdas žymenų raiškos, analizuotos imunohistocheminiu tyrimo metodu, įvertinimui krūties vėžio audinyje. Ki67 raiškos skaitmeninės vaizdo analizės rezultatai stipriai koreliavo (r=0,86) su patologo vertinimu ir turėjo mažesnį sisteminį nuokrypį nuo stereologiškai nustatytų referentinių verčių. HER2 raiškos skaitmeninės vaizdo analizės rezultatai atitiko patologo vertinimą (kappa=0,86) ir leido tiksliau nustatyti amplifikuotus (HER2 FISH) krūties vėžio atvejus.

2. Biologinių žymenų – ER, PR, HER2, Ki67 – imunohistocheminio tyrimo analize pagrįstas krūties duktalinių karcinomų skirstymas į biologinius tipus padeda atskleisti krūties vėžio biologinę įvairovę:

2.1. duktalinių krūties karcinomų grupėje (n=203) nustatytos statistiškai patikimos sąsajos tarp biologinio naviko tipo ir ligos stadijos (p<0,00002), naviko diferenciacijos laipsnio (p<0,0001), metastazių sritiniuose limfmazgiuose (p<0,0002);

2.2. genų *ESR1-4* ir *p16* hipermetilinimas bei geno *TP53* mutacijos susijęs (p<0,05) su biologiniu krūties vėžio tipu;

2.3. biologiniams krūties vėžio tipams būdingas savitas baltymų ER, PR, AR, HER2, Ki67, BCL2, HIF1α, SATB1, p53, p16 raiškos intensyvumo profilis: visų, išskyrus

p53, imunohistocheminių žymenų raiškos intensyvumas buvo susijęs ($p<0,05$) su biologiniu krūties vėžio tipu.

3. Skaitmeninės vaizdo analizės būdu įvertinto imunohistocheminio žymenų rinkinio daugiamatė analizė yra efektyvus tyrimo metodas, leidžiantis atskleisti baltymų raiškos profilių dėsningumus:

3.1. duktalinių krūties karcinomų grupėje (n=109) faktorinės daugiamatės analizės būdu identifikuotas ligos agresyvumą atitinkantis veiksnys (i-G) pasireiškia priešingos krypties ER/PR/AR/BCL2 ir Ki67/HIF1α raiška. Šio veiksnio vertės stipriai susijusios su naviko diferenciacijos laipsniu ir agresyviais biologiniais tipais;

3.2. SATB1 kartu su HIF1α formuoja antrą reikšmingą variacijos veiksnį krūties duktalinės karcinomos grupėje. Šie žymenys gali suteikti svarbios biologinės informacijos apie krūties vėžio navikus nepriklausomai nuo Ki67 raiškos ir naviko diferenciacijos laipsnio;

3.3. hormonų receptoriams teigiamų navikų grupėje (n=85) naviko agresyvumą geriausiai atspindi Ki67 ir ER žymenų derinys. Šis indeksas gali suteikti papildomos prognostinės informacijos greta jau naudojamų prognostinių grupių;

3.4. hormonų receptoriams teigiamų navikų grupėje nustatytas atvirkštinis ryšys tarp HER2 ir PR raiškos bei tarp Ki67 ir ER raiškos atskleidžia ER ir PR disocijuotos raiškos biologinę reikšmę.

REKOMENDACIJOS

1. Mūsų darbo rezultatai rodo, kad biologinių žymenų nustatymas audinių mikrogardelėse, panaudojant skaitmeninės vaizdo analizės technologijas, leidžia padidinti šių tyrimų tikslumą ir efektyvumą. Siūlome skaitmeninio vaizdo IHC analizę naudoti kokybei užtikrinti bei sprendimams palaikyti klinikinėje praktikoje, šias priemones validuojant ne tik pagal patologo vertinimo rezultatus, bet ir stereologiniais tyrimo metodais.

2. Siūlome įdiegti į klinikinę praktiką integruotus IHC žymenų rinkinius (kompleksinius žymenis), kurie atskleidžia vidinius biologinių žymenų tarpusavio sąsajų dėsningumus ir atveria naujas galimybes prognozuoti krūties duktalinės karcinomos ligos eigą.

10. CURRICULUM VITAE

First name:	Aida
Surname:	Laurinaviciene
Address:	National Center of Pathology, Affiliate of Vilnius University Hospital Santariskiu Klinikos, P Baublio 5, LT-08406 Vilnius, Lithuania
Date of birth:	May 5, 1968
Education:	Vilnius University, Faculty of of Natural Sciences, biologist (1995) Vilnius University, Faculty of Medicine, medical biologist (2002)
Current position:	Head of Laboratory at the National Center of Pathology, Affiliate of Vilnius University Hospital Santariskiu Klinikos
Pedagogical work:	Lectures and seminars for medical students at University of Applied Sciences, Faculty of Health Care
Memberships:	President of Lithuanian Society of Pathology technologists