Study on breast carcinoma Her2/neu and hormonal receptors status assessed by automated images analysis systems: ACIS III (Dako) and ScanScope (Aperio)

Janina Słodkowska[1], Violetta Filas[2], Elżbieta Buszkiewicz[2], Paweł Trzeciak[2], Michał Wojciechowski[3], Robert Koktysz[1], Wojciech Staniszewski[3], Jan Bręborowicz[2], Marcial Garcia Rojo[4]

[1]Department of Pathology, Military Institute of Medicine, Warsaw,
[2]Department of Tumour Pathology, Wielkopolskie Center of Oncology, Poznan,
[3]Precoptic, Nikon's Distributor, Warsaw, Poland,
[4]Pathology Department, Hospital General de Ciudad Real, Ciudad Real, Spain.

Abstract: Her-2/neu is overexpressed in 20-30% of breast cancer patients and is associated with a more aggressive disease. Identification of Her-2/c-erbB-2-neu overexpression is based on immunohistochemical [ihc] detection of protein and/or gene amplification in fluorescence in situ hybridization test (FISH). Also Estrogen receptors [ER] and Progesterone receptors [PR] are the prognostic and predictive biomarkers, recently analysed by ihc methods. Subjective, manual scoring of the ihc Her-2/neu expression and expression of the ER/PR reported as the percentage of immunopositive cells are the most common mode of interpretation among pathologists. Automated microscopy and computerised processing have provided increased accuracy in quantification and standardisation. The aims of our study were: to evaluate the scoring reproducibility of Her-2 /neu ihc expression tested by two automated systems: ACIS (Dako) and ScanScope (Aperio); to estimate the ER/PR expression in ihc staining methods with different anti-ER/anti-PR antibodies (the monoclonal and the ER/PR pharmDx TM Kit) by the ACIS system. Her-2/neu ihc expression was measured in 114 primary invasive breast carcinomas by the manual and the automated scoring (ACIS and Aperio system). 106 slides stained ihc with two types of anti-ER/anti-PR antibodies entered the quantisation. The results of our investigations showed very high reproducibility of Her-2/neu scores in intra- and interobserver analysis by ACIS evaluation. The major concordance was present in strong 3+ ihc cases; very small discordance was shown by cases with low expression of Her-2/neu. The accuracy of scoring by the Aperio was little lower in comparison to ACIS but it might result from the smaller and variable series of samples analysed by Aperio. The concordance in scoring of two automated systems was 86.5% (p<0.0001; γ=0.887); the discordance was referred only to the lower expression of Her-2/neu. The concordance in manual scoring performed by the single observer and the panel was 84.2% (p<0.0001, γ = 0.99); the discordance comprised a few cases with strong expression (2+ *vs* 3+). Very high intra- and interobserver reproducibility of the ER/PR ihc measurements was present in the readers results (referred to the percentage of immunoreactive carcinomatous cell population in the breast carcinomas acc. to the ACIS algorithm). No differences were disclosed in the percentage of ER-immunoreactive and PR-immunoreactive carcinomatous cell populations when used 2 different type of antibodies, in the ACIS automated method.

Key words: Her-2/neu scoring, automated quantisation, ACIS, APERIO, manual scoring, hormonal receptors, ER/PR expression.

Introduction

Breast cancer is the most common cancer among women in Europe, in the United States and Australia.

It accounts for 22% of all female cancers [1]. *Invasive ductal carcinoma, not otherwise specified (ductal NOS) s. infiltrating ductal carcinoma* comprises the largest group of invasive breast cancers [2]. The histopathological examination of the breast cancer bases on the morphological features but more specific prognostic information about its biology are obtained from the immunohistochemical (ihc) testing of the

Correspondence: J. Słodkowska, Military Institute of Medicine, Department of Pathology, Szaserów 128 Str. 04-141 Warsaw, Poland; tel./fax.: (+4822) 8103892, e-mail: joanslo@wp.pl

human epidermal growth factor receptor Her-2 (Her2/c-erbB-2-neu), Estrogen Receptors (ER) and Progesterone Receptors (PR). Identification of Her-2/c-erbB-2-neu overexpression is based on ihc detection of protein and/or gene amplification in fluorescence *in situ* hybridization test (FISH).

Her-2/neu is overexpressed in 20-30% of breast cancer patients; is associated with a more aggressive disease, a poor clinical prognosis and with the targeted therapy agent trastuzumab (Herceptin). Expression of the ER/PR is the most reliable factor for predicting responsiveness to hormonal therapy. The hormone receptor expression measured ihc is accepted as standard evaluation method all over the world, but the cut off point in ihc evaluation is still controversial. Another controversy in testing ER /PR is a mode of presentation of the receptors quantitation as: positive *vs* negative results or the quantitative results, or both modes. Various scoring systems have been used in ihc evaluation of the ER/PR [3-5].

Several factors might be responsible for the Her-2/neu and ER/PR ihc scoring results: the quality assessment system, heterogeneous distribution of positive cells and staining intensity, lack of uniformity in the manual testing, the mode of quantitation – subjective manual *vs* automated computerised reading, and insufficient training of the observers.

Automated microscopy and computerised processing have provided increased accuracy in quantification and standardisation. Use of automation in ihc interpretation has decreased the number of ambiguous cases and helped reducing the need for further testing by another modality [6,7]. The development of the digital pathology and telepathology has increased the interest in implementation of the automated computerised methods in the routine diagnostics [8,9].

The aims of our study were: *to evaluate the scoring reproducibility of Her-2 /neu ihc expression tested by two automated systems: ACIS (Dako) and ScanScope (Aperio) used as the aids for the pathological diagnostics of breast carcinoma; *to estimate the ER and PR ihc expression resulted from two different ihc staining methods: a/with the monoclonal antibodies anti-ER and anti-PR and b/with the ER/PR pharmDx ™Kit -DAKO, by automated measurement with ACIS system.

Materials and methods

Her-2/neu: 114 primary invasive breast carcinomas (core biopsies and mammotomes material) were selected for the study, from the routine diagnostic and consultative material of the Tumour Pathology Department, at the Wielkopolskie Center of Oncology [WCO] (in Poznan, Poland) (from the period 2006-2008y). Tissue samples were fixed in 10% neutral buffered formalin and processed to paraffin blocks using standard histopathological techniques; 4 μm paraffin sections were stained routinely with H&E. For demonstration of the Her-2 protein expression a commercially available kit

HercepTest™ (cat. no. K5204, Dako, Denmark) was used. The deparaffinised tissue sections were rehydrated and incubated in 10 mmol/L citrate buffer (Epitope Retrieval Solution) in a water bath 95-99°C (for 40'). Next, endogenous peroxidase activity was blocked using 3% hydrogen peroxide solution (for 5'). Subsequently the sections were incubated with the primary antibody, rabbit anti-human Her-2 (for 30', at room temperature) and incubated with the kit Visualization Reagent (for 30' at room temperature). The reaction product was visualized by the chromogen DAB (DAB Substrate-Chromogen Solution). Cell nuclei were counterstained using haematoxylin. Positive controls slides were supplied with the HercepTest kit 9 [10].

The original glass slides of breast carcinoma were examined by single pathologist and by the panel of pathologists of WCO, with the light microscopes Olympus BX50 (Japan). The slides marked for Her-2/neu membrane staining were evaluated by four-point scoring scale: 0, 1+, 2+ and 3+ with cut off level >10% immunopositive tumour cells [11]. For 20 out of 29 breast cancer cases scored manually as 2+ Her2/neu, the test FISH was performed.

The automated quantisation of Her-2/neu ihc expression was done by pairs of pathologists with ACIS (Automated Cellular Imaging System III) (Dako, Denmark), and Aperio system (Aperio Technologies, U.S.A.) [6,12,13].

ACIS: The digital images of 114 glass slides were captured automatically by ACIS scanner at low power magnification and the whole slide images [WSI] were viewed on a monitor. The regions of the highest immune intensity and highest percentage of positive cells were selected by user (or chosen from the displayed "hot spots" formerly identified by ACIS system) for the further automated scoring. The minimum six "40x areas" (a circular field marked automatically by the system) containing only tumour cells were selected for a quantitative evaluation. In the process of a region selection only *the invasive carcinoma* morphology (not *carcinoma in situ* areas) were analysed. WSI were evaluated 3 times by 2 pathologists with algorithms for membrane staining (Her-2) according to the manufacturer's manuals. A total score of Her-2/neu expression was calculated automatically from the percentage of immunopositive cells and immunostaining intensity, and was shown as a mean region score for all studied regions. The final result of Her-2/neu expression was read by the observers from four point scale: from 0 to 3+. A cut off level for negative ihc expression was 10% population of immunopositive carcinoma cells.

APERIO: The scanning of 114 glass slides was done manually with Aperio Scanscope CS under 20x. The score areas were selected by the free form drawings (minimum 6 score areas of various size) and were evaluated twice by 2 pathologists under 40x with Aperio algorithm for membrane staining, according to the manufacturer's manuals. Her-2/neu ihc quantisation was done automatically by the system and the results were displayed in the scale from 0 to 3+.

Estrogen and progesteron receptors. For the ER/PR study, 106 standard quality cases of the invasive breast carcinoma were selected from the archive of the TPD WCO in Poznan (from the period 2006-2008y). The major part of the material covered the cases of Her-2/neu study. The breast carcinoma tissues were fixed and processed according to the standard procedures [14].

Two types of Dako antibodies were used for the ER/PR ihc staining. A/ The primary monoclonal antibodies anti-ER (clone 1 D5, code M 7047) and anti- PR (clone PGR 636, code M 3569) in two-step immunohistochemistry. Sections were deparaffinized, rehydrated and incubated in Dako Target Retrieval Solution (Code S 1699) in a water bath at 95-99°C (for 40'). Endogenous peroxidase activity was blocked using a solution of 3% hydrogen peroxide (for 10'). Sections were incubated with monoclonal antibodies (anti-ER 1:50 and anti-PR 1:100) (for 60'). Subsequently, the slides

were incubated with Dako's En-Vision Horse Radish Peroxidase complex (Dako, code 4001) (for 30') and with DAB chromogen solution (Dako, code K3466) (for 5'). The cell nuclei were counterstained using haematoxylin. The positive control slides were breast cancer specimens with known positive reactions. B/ The Dako ER/PR pharmDx ™Kit (ER/PR pharmDx K1904 kit for manual use) was applied for the second type of ihc staining. The kit contains the necessary reagents to complete the two stages ihc method according to the manufacturer's instruction. Positive ER/PR pharmDx Control Slides were provided by Dako.

Only good quality ihc slides were included for the quantisation by ACIS system. In the group A: 105 ER slides and 106 PR slides; and in the group B: 90 ER and 93 PR slides – were scanned at low power magnification by the ACIS. The ihc expression of the ER/PR was measured automatically in the regions selected by users (minimum 6 areas, minimum 1000 cells) under magnification 40x. Results of the ER/PR were presented as the percentage of immunopositive carcinomatous cells (the ihc intensity is not included into the algorithm of ACIS system). Each WSI was evaluated 3 times (a, b, c) by 2 pathologists.

The performance of ACIS analysis enclosed many important steps and restrictions provided by the instructions for using the ACIS system in Her-2/neu as well as in ER/PR evaluation. Supplementary to the written instruction a practical training at the Dako Laboratory (Copenhagen, Denmark) and consultations were necessary for an appropriate scoring.

Statistical analysis. The results of the manual and automated counting were analysed by the statistical methods, for each studied receptor separately. Associations between categorical variables were evaluated by the Chi^2 test or Fischer exact test and correlation coefficient (γ). The variance analysis (ANOVA) was performed with the post-hoc RIR-Tukey's test (for the variables with normal distribution). The Mann-Whitney U test or Friedmann ANOVA test were applied for variables with lack of normality.

Results

Her-2/neu

The results of ihc Her-2/neu evaluation are presented according to 3 modes of quantisation: manual reading and the automated assessment by ACIS and APERIO system.

The comparative analysis performed for the manual scores of the single observer and the panel revealed the statistical highly significant concordance (p<0.0001, correlation coefficient $\gamma = 0.99$ by Pearson's Chi^2 test) in 84,2% (96/114) cases. The panel recognised 57 (50%) scores 0/1+, 29 (25,4%) scores 2+ and 28 (24,6%) scores 3+. The single observer recognised 50 (43,9%) scores 0/1+, 35 (30,7%) scores 2+ and 29 (25,4%) scores 3+. Thirteen discordant scores (eight of 2+ *vs* 1+, and five of 2+ vs 3+) were verified by FISH method and/or reviewed by the pathologists.

In ACIS evaluation of 114 WSI, very high reproducibility of two readers results was noticed. The highly statistical significance was present in triple scoring of each reader (E and T): p=0.0001, correlation coefficient $\gamma = 0,999$ and $\gamma = 0,998$ (Pearson's Chi^2 test), respectively. The analysis of variance (ANOVA with post hoc RIR Tukey's) didn't show any significant dif-

Table 1. Her-2/neu ihc measurement by the ACIS: the results of comparative statistical analysis and a list of scores discordant between two readers.

E readings	Discordant cases n (scores)	T readings	Discordant cases n (scores)
	Chi^2 results γ – correlation coefficient		Chi^2 results γ – correlation coefficient
Ea – Eb	3 (0/1+ *vs.* 2+)	Ta - Tb	2 (3+ *vs.* 2+) 3 (2+ *vs.* 1+/0)
	p<0.05 γ=1.0		p<0.05 γ=0.997
Eb – Ec	2 (2+ *vs.* 1+/0)	Tb - Tc	5 (2+ *vs.* 1+/0) 1 (2+ *vs.* 3+)
	p<0.05 γ=1.0		p<0.05 γ=0.997
Ea – Ec	3 (2+ *vs.* 1+/0)	Ta - Tc	3 (2+ *vs.* 1+/0) 1 (2+ *vs.* 3+)
	p<0.05 γ=0.999		p<0.05 γ=0.998

ferences in E and T triple measurements of the mean region scores (Ea-Eb, Eb-Ec, Ea-Ec and Ta-Tb, Tb-Tc, Ta-Tc), p<0.05. The scores discordant in two readers evaluation and the statistical data are presented in Table 1.

In the comparative analysis of the mean measurements obtained by two readers (E and T) the concordance was in 87.7% (100 out of 114) of results. The significant difference was revealed by the analysis of variance (ANOVA p<0.05). The comparative statistical analysis of two readers measurements revealed strong correlation between obtained results and indicate very high interobserver reproducibility for Her-2 results.

The data of the comparative analysis of manual scores obtained by the single reader and the scores obtained by automated method (ACIS) are presented in Table 2. The concordance in the reader E results was 85.1%, 81.6% and 81.6% (p<0.05, correlation coefficient : γ =-0.474, γ =-0.462 and γ =-0.478 [Pearson's Chi^2]); in the reader T results a concordance was 78.1%, 79% and 79% (p<0.05, correlation coefficient were γ =-0.463 and γ =-0.465).

The list of discordant cases between automated and manual evaluation of two readers: for the reader E – **Ea**: 17 scores of 1+ *vs* 2+ ; **Eb**: 15 scores of 1+ *vs* 2+; 1 score of 2+ *vs* 0/1+; 5 scores of 2+ *vs* 3+; **Ec**: 16 scores of 1+ *vs* 2+; 5 scores of 2+ *vs* 3+. And for the reader T – **Ta**: 1 score of 2+ *vs* 1+; 17 scores of 1+ and 1 score of 0 *vs*. 2+; 6 scores of 2+ *vs*. 3+; **Tb**: 16 scores of 1+ *vs* 2+, 6 scores of 2+ *vs* 3+; **Tc**: 18 scores of 1+ *vs*. 2+; 5 scores of 2+ *vs*. 3+.When compare the manual and final ACIS scores the concordance was 72.8%

Table 2. Her-2 scoring: discordant results between manual scoring of the single observer and reading by ACIS.

	ACIS: discordant scores in manual and automated reading					
	Reader E results			Reader T results		
	Ea	Eb	Ec	Ta	Tb	Tc
Number of discordant cases (%)	17 14.9%	21 18.4%	21 18.4%	25 21.9%	24 21.0%	24 21.0%
Statistical data (Pearson's Chi^2)	p<0.05 γ=-0.474	p<0.05 γ=-0.462	p<0.05 γ=-0.478	p<0.05 γ=-0.463	p<0.05 γ=-0.465	p<0.05 γ=-0.465

(83/114) and showed very high statistical significance (p<0.0001, γ = 0.913). 31 discordant results comprise: 7 in the class 0 *vs* 1+, eighteen in 1+ *vs* 2+ and six in 2+ *vs* 3+.

In Aperio study, 90 WSI were evaluated by the reader S and 65 by the reader W; only 40 cases fitted in both readers scoring; 24 WSI were excluded from the analysis due to technical problems or poor quality of images. The interobserver concordance in the scores was 80% (32/40): (p<0.05, γ = 0.996 [Pearson's Chi^2]). The discordant results included: 7 scores 2+ *vs* 1+ and 1 score 3+ *vs* 2+. High intraobserver reproducibility was shown in the reader W scoring (p<0.05, γ = 0.985 [Pearson's Chi^2 test]). Five (13.5%) discordant results included 3 of the class 2+ *vs* 1+ and 2 of the category 2+ *vs* 3+.

The concordance of the automated and manual scoring in two readers evaluation was 75.4% (49/65) and 80% (72 /90); and was statistically significant (p<0.05, correlation coefficient γ =-0.297 and γ =-0.367 [Pearson's Chi^2 test]). The discordant cases included: two of 2+ *vs* 0/1+, twelve of 1+ *vs* 2+; four of 2+ *vs* 3+ (in the reader S study); and ten of 1+ *vs* 2+, six of 2+ *vs* 3+ (of the reader W). When compare the manual and final automated scoring by Aperio the concordance was shown in 70.4% (57/81) results (p<0.0001, γ = 0.794). 24 discordant scores represented the category: 0 *vs* 1+ in six cases; 1+ *vs* 2+ in thirteen cases, 2+ *vs* 3+ in five cases.

The comparative analysis of the final Her-2/neu ihc scores obtained by the ACIS and the Aperio, revealed high concordance in 86.5% (70/81) scores (p<0.0001, γ = 0.887). The discordant cases were present in 2 categories: 0 *vs* 1+ nine cases and 1+ *vs* 2+ four cases. No discrepancy appeared in the class 2+ *vs* 3+.

Estrogen and Progesterone Receptors

All results for the ER and PR ihc measurements of the groups A and B are presented as the mean values (+/- std. dev.) of the percentage of ER/PR-immunopositive carcinomatous cell population.

Estrogen Receptors – Group A [A-ER]. The results of the reader E and T (a, b, c) were as follows: 82.36 (+/- 27.68); 81,70 (+/-28.34); 81.73 (+/-28.26); and 83,47 (+/-28.44); 83.87 (+/-28.11); 84,18 (+/-27.99), respectively; and showed very high intraobserver reproducibility. No significant differences were present between Ea, Eb and Ec readings (Friedman's ANOVA N=105, df=2, p=0.45; Kendall's concordance coefficient =0.034) nor between Ta, Tb and Tc measurements (N=105, df=2, p=0.49; Kendall's concordance coefficient =0.035). There were the significant differences between measured values (p<0.01) when compared two readers results (Ea, Eb, Ec and Ta, Tb, Tc).

Estrogen Receptors – Group B [B-ER]. The results of the observer E and T presented very similar values: 77.07 (+/- 28.21), 76.93 (+/- 28.32), 76,66 (+/- 28.21) and 79.39 (+/- 29.22), 79.27 (+/- 29.09), 79.18 (+/- 29.34), respectively. No significant differences were noticed between Ea, Eb and Ec measurements (Friedman's ANOVA N=93, df=2, p=0.95; Kendall's concordance coefficient =0.0006) nor between Ta, Tb and Tc measurements (N=93, df=2, p=0.66; Kendall's concordance coefficient =0.0045). The data confirm very high intraobserver reproducibility of both readers results. The statistically significant differences p<0.001were shown when compared both readers measurements (Ea, Eb, Ec and Ta, Tb, Tc).

Progesterone Receptors – Group A [A-PR]. The measurements of the reader E and T were: 63.84 (+/- 37.27); 63.60 (+/-37.31); 63.92 (+/-37.39) and 65.98 (+/-38.27); 66.95 (+/-38.06); 66.34 (+/- 38.09), respectively; and showed very high intraobserver as well interobserver reproducibility in the statistical analysis. No significant differences were present between Ea, Eb and Ec measurements (Friedman's ANOVA N=106, df=2, p=0.19; Kendall's concordance coefficient =0.016) nor between Ta, Tb and Tc measurements (N=106, df =2, p=0.55; Kendall's concordance coefficient =0.005). The comparison of both readers results (Ea, Eb, Ec and Ta, Tb, Tc) showed the significant differences between measured values (p <<0.001).

Progesterone Receptors – Group B [B-PR]. The results of the reader E and T showed the following values: 72.65 (+/- 32.81); 72.49 (+/-32.85) and 72.28 (+/-32.91 and 75.02 (+/-33.51), 75.11 (+/-33.69) and 75.40

(+/-33.20), respectively. No significant differences were present between Ea, Eb and Ec measurements (Friedman's ANOVA N=90, df=2, p=0.81; Kendall's concordance coefficient =0.002) nor between Ta, Tb and Tc results (N=90, df=2, p=0.34; Kendall's concordance coefficient =0.012). The statistical data confirmed very high intraobserver reproducibility of E and T results. The comparison of both readers results (Ea, Eb, Ec and Ta, Tb, Tc) indicated the statistically significant differences p<0.001.

Estrogen Receptors in the Group A and B [A-B/ER]. The comparative statistical analysis of ER ihc results for the group A and B, did not reveal any significant difference between the results in both readers evaluation. The "p" values were as follows: p=0.404 (for Ea), p=0.498 (for Eb), p=0.445 (for Ec); p=0.548 (for Ta), p=0.479 (for Tb), p=0.419 (for Tc).

Progesteron Receptors in the Group A and B [A-B/PR]. The analysis of variance (ANOVA) for the A-PR and B-PR results of both readers E and T, didn't show any significant difference between results of both ihc staining methods (A and B). The "p" values for the individual reading were as follows: p=0.196 (for Ea), p=0.198 (for Eb), p=0.231 (for Ec); p=0.178 (for Ta), p=0.266 (for Tb), p=0.191 (for Tc).

Discussion

Semiquantitative ihc technique is of the great importance in the routine breast cancer Her-2/neu and ER/PR evaluation and is dependent on the observers and the applied antibody. The tissue handling, staining procedures and observer estimation can strongly affect the results and cause the inter- and intraobserver diversity [4,15]. For Her-2/neu study we used the validated Dako Hercep Test ihc assay according to the guidelines recommended by the American Society of Clinical Oncology and the College of American Pathologists [4].

Various automated image analysis systems offer a possibility of rapid and objective scoring of ihc staining. Her-2 is currently assessed manually by the pathologists with various cut off thresholds (1-30% positive tumour cells) which can lead to significant intra- and interobserver variability. In our study a fully automated membrane algorithm (FDA approved) has been used for Her-2/neu assessment in ACIS III and Aperio system with cut off threshold of 10% positive tumour cells [6]. Both computerised image analysis systems had installed the nuclear algorithms for the ER/PR evaluation.

The quantification by ACIS of Her-2/neu ihc expression in 114 primary breast tumours, showed very high intraobserver reproducibility of the mean region scores (p=0.0001, γ =0.988 and γ =0.944; the statistically significant differences: p>0.05

[ANOVA]); as well as very high interobserver reproducibility of these results (concordance in 87.7%; p<0.05 [ANOVA]).

The concordance between automated and manual scoring of each observer was 79% and 81.6% (p<0.05); the concordance between the final ACIS scoring and manual results was 86.5% (p<0.0001,γ =0.887).

The quantification by Aperio of Her-2/neu ihc expression was performed in the smaller series of breast carcinomas than in ACIS series but the data also showed high and statistically significant intraobserver concordance (86.5%; p<0.05, γ =0.985). The interobserver concordance in Aperio's scores was 80% and showed high statistical significance (p<0.05, γ =0.996). The comparative analysis of the manual and automated scoring by Aperio revealed statistically significant concordance in both reader results: 75.4% and 80% (p<0.05, γ =-0.297 and γ =-0.367). The statistically significant concordance in 70.4% results (p<0.0001, γ =0.794) appeared when compare the final Aperio scores with the manual data. Very high concordance 86.5% (p<0.0001, γ = 0.887) – was shown between automated scoring when compare the final scores of the ACIS and the Aperio system. These results confirm very high accuracy of the automated assessment by ACIS and Aperio of Her-2/neu ihc expression.

The statistical data also indicate the higher accuracy the ACIS system than the Aperio in Her-2/neu ihc assessment. However in the final interpretation of the obtained data, should be considered several factors which could contribute in the scoring results. The quantified series of breast carcinoma were not equal in the ACIS and the Aperio investigations (smaller and variable series were included to the Aperio analysis). Also there was a methodological difference in selection of counting fields: six "40x areas" by ACIS and the free form drawings by Aperio. The third important factor which might influence the results of scoring was the breast carcinoma architecture and samples quality, seen in several cases. In a few biopsies a structure of the intraductal carcinoma dominated over the scanty morphology of the invasive carcinoma, makes difficult the morphological differentiation and proper selection of the representative areas. The second type of the morphological difficulties was related to the microfocal invasion of the carcinomatous cells, enabling a clear separation of the minute carcinomatous foci from a stroma – especially for the automated "six 40x" method, and to some extends for the Aperio method. Less important problem influencing the scoring results seems to be not excellent quality of WSI due to procedure of ihc staining. Uneven thickness of slide and/or staining artefacts caused problem in focusing during scanning

which was easy to overcome by the manual focusing in the Aperio system.

In respect to a validation of the automated scoring results by comparison with the manual reading data, this issue needs more extended and more precious study. In our analysis the concordance between both reading methods (manual and automated) was 72.8% and 70.4% for the ACIS and the Aperio, respectively. The concordance in the manual results of the single observer and the panel scores was 84.2% (p<0.0001, γ = 0.99) with 5 discordant scores within the category of 2+ *vs* 3+. When consider the accuracy of ACIS measurements the interobserver concordance was 87.7% without any discordant case within the category 2+ *vs* 3+. These results are in favour of very high accuracy of measurements by the automated ACIS system. Wang's *et al.* in the comparative study on Her-2/neu ihc expression assessed by manual assay and with ACIS quantitation method showed an improvement of ACIS method over the manual method for objective evaluation of Her-2/neu status for 2+ cases [7]. The interobserver concordance in Aperio measurements was 80% (but in much smaller series than ACIS population) with one discordant case in the category 2+ *vs* 3+. It is impossible to exclude the contribution of the subjectivity and the laboratory standard in the obtained scoring results of all methods.

Among limits of any immunohistochemical study the most important are the specificity and sensitivity of the antibody being used as well as the objective and reproducible quantification. Many approaches to the quantification of the ER/PR described in the literature have used various antibodies [16-19]. In the current comparative study on the ER /PR ihc expression, two types of antibodies and adequate various staining procedures were applied: the monoclonal antibodies anti-ER/anti-PR and the ER/PR pharmDx ^{TM}Kit – both products of Dako.

The automated algorithms that precisely quantitate the percentage of staining and measure the intensity on a continuous scale are better suited to provide objective data than traditional subjective methods. A number of groups have published data on the automated assessment of the ER/PR in breast cancer [16,20-22] describing an excellent correlation between manual and automated analysis. Our study focused on the quantitative analysis of the ER/PR ihc expression when two different types of antibodies against ER and against PR were used for the breast carcinoma samples. The quantitative measurement was performed by the automated system ACIS. The measurements of the percentage of the ER-immunoreactive cell population in the breast carcinoma showed very high intraobserver reproducibility for each type of anti-ER antibody (A and B) in two readers evaluation. The statistically significant concordance was present in the measurements

of each observer in the group A stained with the primary monoclonal antibody (p=0.027, Kendall's concordance coefficient =0.034 and p=0.025; Kendall's concordance coefficient =0.035 by Friedman's ANOVA) and in the group B stained with ER pharmDx ^{TM} antibody (p=0.95; Kendall's concordance coefficient =0.0005 and p=0.66; Kendall's concordance coefficient =0.0045). Very high interobserver reproducibility of the ER-A and ER-B results was confirmed by the statistical comparative analysis. No significant differences were shown between the percentage of ER-immunoreactive populations stained by both antibodies ("p" values ranged between p=0.404 and p=0.548).

The measurements of the percentage of the PR-immunoreactive carcinomatous cell population of each type of anti-PR antibody showed very high concordance in two observers evaluation. The intraobserver concordance of each reader data for the monoclonal anti-PR antibody results was statistically significant (p=0.19, Kendall's concordance coefficient =0.016 and p=0.55, Kendall's concordance coefficient =0.005 by Friedman's ANOVA). Also the analogical measurements collected for the group B (staining with the PR pharmDx ^{TM}) showed very high intraobserver concordance (p=0.81, Kendall's concordance coefficient=0.002 and p=0.34, Kendall's concordance coefficient=0.012 for each observer respectively). The statistically significant differences were present between both readers results for the group A-PR and the group B-PR (p<<0.001).

The percentage of ER-immunopositive carcinomatous cell populations in both ihc staining methods was similar, no statistically significant differences between measured values were shown (p <0.548). Also the percentage of PR-immonoreactive carcinomatous cell populations didn't differ in both type of used anti-PR antibodies (p<0.266). These results seems to be contradictive to the general opinion about the Dako ER/PR pharmDx ^{TM} Kit. Although various factors can be assumed as the cause for this discrepancy, such as the effect of activation by autoclaving, our small experience in this method of staining; one important factor should be consider – the ihc intensity in quantitative analysis of immunoreaction [17,18] . The ER/PR algorithm used by ACIS system did not include the parameter of ihc intensity therefore the presented results are referred only to the percentage of immunopositive carcinomatous cell population. Indirectly our results support the opinions about importance of ihc intensity measurements in quantitative evaluation of immunoreactivity.

In conclusion, our findings indicate the following: 1) The grading scores for Her-2/neu evaluated by the ACIS showed very high reproducibility of the results in intra- and interobserver analysis. The major concordance is present in strong 3+ ihc cases; small discordance occurs in the group with lower expression of

Her-2/neu. 2) The accuracy of Her-2/neu ihc scoring by the Aperio seems to be lower in comparison to the ACIS but it might result from the smaller and variable series analysed by Aperio. 3) The concordance in scoring by two automated systems (ACIS and Aperio) was 86.5% (p<0.0001; γ =0.887); the discordance was referred only to the lower expression of Her-2/neu. 4) The concordance in manual scoring performed by the single observer and the panel was 84.2% (p<0.0001, γ =0.99), however the discordance comprised five cases with strong expression (2+ *vs* 3+). 5) The scoring for Her-2/neu measured manually by the single observer and by the automated methods showed the concordance in 72.8% (p<0.0001; γ =0.913) by the ACIS and in 70.4% (p<0.0001; γ =0.794) by the Aperio. 6) Very high intraobserver and interobserver reproducibility of the ER/PR ihc measurements was present in both readers results which were referred to the percentage of immunoreactive carcinomatous cell population in the breast carcinomas. 7) No differences were disclosed in the percentage of ER-immunoreactive and PR-immunoreactive carcinomatous cell populations when used 2 different type of antibodies, by the ACIS automated method (which does not include the parameter of ihc intensity). 8) The used automated methods show very high quantitative accuracy especially ACIS system however the interpretation of the results should include the contribution of other factors.

References

[1] Parkin DM, Bray F, Ferlay J, Pisani P. Estimating the world cancer burden: Globocan 2000. *Int J Cancer*. 2001;94:153-156.

[2] 2.World Health Organisation Classification of tumours: Pathology and genetics of tumours of the breast and female genital organs. Tavassoli FA&Devilee P. IARCPress, Lyon, 2003:13.

[3] Allred DC, Harvey JM, Berardo M, Clark GM. Prognostic and predictive factors in breast cancer by immunohistochemical analysis. *Mod Pathol* .1998;11:155-168.

[4] Umemura S, Kurosumi M, Moriya T *et al* Immunohistochemical Evaluation for Hormone Receptors in Breast Cancer: A Practically Useful Evaluation System and Handling Protocol. *Breast Cancer*. 2006;13: 232-235.

[5] Kurosumi M. Immunohistochemical Assessment of Hormone Receptor Status Using a New Scoring System (J-Score) in Breast Cancer. *Breast Cancer*. 2007;14:189-193.

[6] Tawfik OW, Kimler BF, Davis M *et al.* Comparison of immunohistochemistry by automated cellular imaging system (ACIS) versus fluorescence *in-situ* hybridization in the evaluation of HER-2/neu expression in primary breast carcinoma. *Histopathology*. 2006;48:258-267.

[7] Wang S, Saboorian MH, Frenkel EP *et al.* Assessment of HER-2/neu Status in Breast Cancer. Automated Cellular Imaging System (ACIS)-Assisted Quantitation of Immunohistochemical Assay Achieves High Accuracy in Comparison With Fluorescence *In Situ* Hybridization Assay as the Standard. *Am J Clin Pathol*. 2001;116:495-503.

[8] Peces C, Garcia-Rojo M, Sacristan J, Gallardo AJ, Rodriguez A. Serendipia: Castilla-La Mancha telepathology network. *Diagn Pathol*. 2008;3(suppl.1): S5.

[9] Gilbertson J, Yagi Y. Histology, imaging and new diagnostic work-flow in pathology. *Diagn Pathol*. 2008;3 (suppl.1):S14.

[10] Miller K. Immunocytochemical Techniques. In: Bancroft JD. Gamble M.: *Theory and Practice of Histological Techniques*. Churchill Livingstone; 2002, pp. 421-464.

[11] Briner P, Oberhuber G, Stani J *et al*. Evaluation of the United Sates Food and Drug Administration – approved scoring and test system of HER-2 protein expression in breast cancer. *Clin Cancer Res*. 2001;7:1669-1675.

[12] Bloom K, Harrington D. Enhanced Accuracy and Reliability of Her-2/neu Immunohistochemical Scoring Using Digital Micoroscopy. *Am J Clin Pathol*. 2004;121:620-630.

[13] Fornier M, Risio M, Poznak C, Seidman A. Her-2 Testing and Correlation with Efficacy of Transtuzumab Therapy. *Oncology*. 2002;16:1340-1358.

[14] Goldstein NS, Hewitt SM, Taylor CR *et al*, and Members of Ad-Hoc Committee on Immunohistochemistry. Recommendations for improved standardization of immunohistochemistry. *Appl Immunohistochem Mol Morphol*. 2007;15:124-133.

[15] Oyama T, Ishikawa Y, Hayashi M, Arihiro K, Horiguchi J. The Effects of Fixation, Processing and Evaluation Criteria on Immunohistochemical detection of Hormone receptors in Breast Cancer. *Breast Cancer*. 2007;14:182-188.

[16] Sharangpani GM, Joshi AS, Porter K *et al*. Semi-automated imaging system to quantitate estrogren and progesterone receptor immunoreactivity in human breast cancer. *J Microscopy*. 2007;226:244-255.

[17] Press M, Spaulding B, Groshen S *et al*. Comparison of different antibodies for detection of progesterone receptor in breast cancer. *Steroid*. 2002;67:799-813.

[18] Fischer ER, Anderson S, Dean S *et al*: Solving the dilemma of the immunohistochemical and other methods used for scoring estrogen and progesterone receptor in patients with invasive breast carcinoma. *Cancer*. 2005;103:164-173.

[19] Rocha R, Nunes C, Rocha G *et al*.: Rabbit monoclonal antibodies show higher sensitivity than mouse monoclonals for estrogen and progesterone receptor evaluation in breast cancer by immunohistochemistry. *Pathol Res Pract*. 2008;204: 655-662.

[20] Rexhepaj E, Brennan DJ, Holloway P *et al*. Novel image analysis approach for quantifying expression of nuclear proteins assessed by immunohistochemistry: application to measurement of oestrogen and progesterone receptor levels in breast cancer. *Breast Cancer Res*. 2008;10(5):R89.

[21] Gokhale S, Rosen D, Sneige N *et al*. Assessment of two automated imaging systems in evaluating estrogen receptor status in breast carcinoma. *Appl Immunohistochem Mol Morphol*. 2007;15:451-455.

[22] Diaz LK, Sahin A, Sneige N. Interobserver agreement for estrogen receptor immunohistochemical analysis in breast cancer: a comparison of manual and computer-assisted scoring methods. *Ann Diagn Pathol*. 2004;8:23-27.