Validation of digital immunohistochemical evaluation of hormone receptors in breast cancer

Tamás Micsik MD

I[st] Department of Pathology and Experimental Cancer Research, Semmelweis University

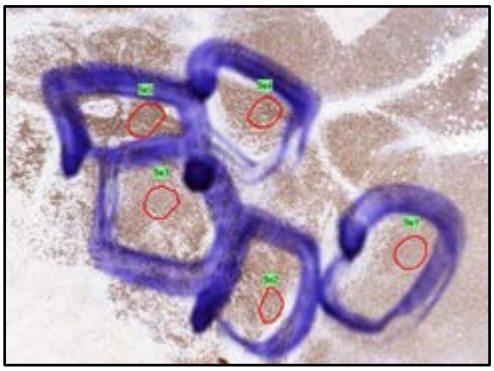
Objective

Estrogen (ER) and progesteron (PR) receptor expression of breast cancers is crucial for hormone therapy. The wide Allred-index (combined from a proportional and an intesity score) assumes lower reproducibility of the scores in histopathological reports influencing patient selection. Thus standardized evaluation and documentation offered by digital pathology can be very beneficial in clinical aspects. Our retrospective study aimed to validate digital evaluation of nuclear-immunohistochemistry.

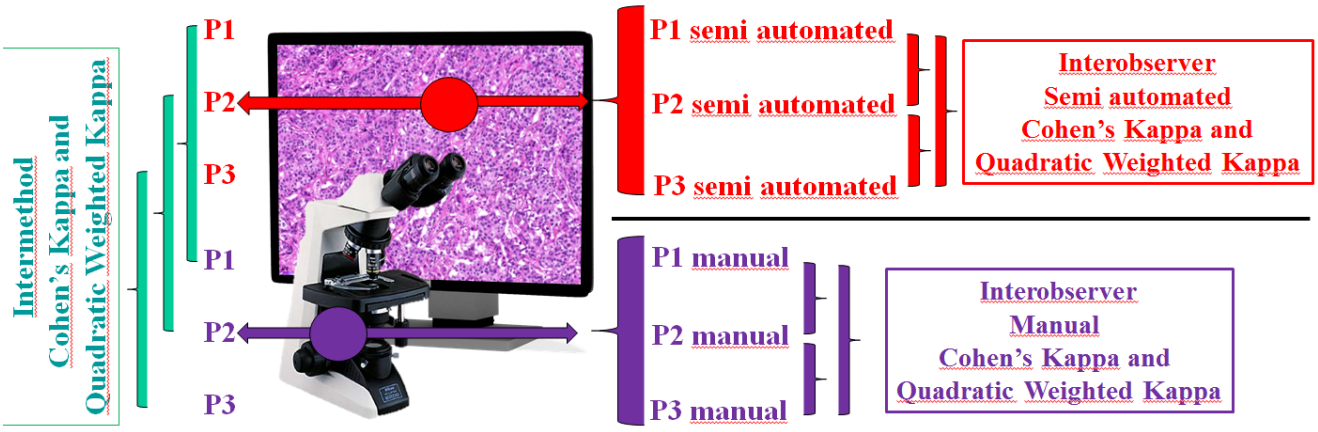
Methods

ER and PR immunostained slides of 186 breast cancer cases were evaluated on glass and on digitized slides by 3 pathologists.

Region of Interests (ROIs) were marked with markers (blue) on glass slides and digital annotations were made within these ROIs (red).



Cohen's kappa (CK) and Quadratic weighted kappa (QWK) were calculated in each interobserver (between pathologists) and intermethod (glass versus digitized) setting.



Results

Manual interobserver CK ranged 0,456-0,645 and QWK ranged 0,917-0,956. Digital interobserver (between pathologists) CK ranged 0,532-0,640 and QWK ranged 0,930-0,962. Intermethod (manual vs semi-automated) CK ranged 0,329-0,781 and QWK ranged 0,894-0,982.

	P1 Manual	P1 Semi-automated	P2 Manual	P2 Semi-automated	P3 Manual	P3 Semi-automated
P1 Manual		0.781 (0.749-0.814)	0.496 (0.459-0.534)	0.650 (0.614-0.687)	0.642 (0.606-0.678)	0.585 (0.549-0.622)
P1 Semi-automated			0.475 (0.438-0.512)	0.638 (0.602-0.675)	0.661 (0.625-0.698)	0.640 (0.604-0.677)
P2 Manual				0.545 (0.517-0.583)	0.501 (0.464-0.538)	0.432 (0.395-0.468)
P2 Semi-automated					0.612 (0.575-0.649)	0.618 (0.580-0.655)
P3 Manual						0.669 (0.633-0.706)
P3 Semi-automated						

Table 23. Comparison of different PgR assessments using Cohen's kappa.
Legend: P=pathologist. Cells highlighted with grey denote agreements between the manual assessments (i.e. gold standard).

	P1 Manual	P1 Semi-automated	P2 Manual	P2 Semi-automated	P3 Manual	P3 Semi-automated
P1 Manual		0.982 (0.977-0.987)	0.924 (0.906-0.942)	0.959 (0.950-0.968)	0.955 (0.945-0.965)	0.934 (0.917-0.951)
P1 Semi-automated			0.926 (0.908-0.943)	0.962 (0.953-0.970)	0.952 (0.942-0.963)	0.943 (0.927-0.959)
P2 Manual				0.939 (0.923-0.955)	0.932 (0.916-0.949)	0.904 (0.882-0.926)
P2 Semi-automated					0.956 (0.947-0.965)	0.945 (0.929-0.960)
P3 Manual						0.942 (0.926-0.958)
P3 Semi-automated						

Table 24. Comparison of different PgR assessments using quadratic weighted kappa.
Legend: P=pathologist. Cells highlighted with grey denote agreements between the manual assessments (i.e. gold standard).

Conclusions

Digitized evaluation of nuclear immunostaining was as precise as the manual evaluation of routine glass slides. Furthermore, digital reading has improved Cohen's Kappa and Quadratic Weighted Kappa-values, especially with PR-immunoslides. Given the wide range of Allred-index, standardization by digital evaluation might improve patient selection for optimal hormone treatment.