



Image analysis in IHC - overview, considerations and applications

NordiQC Workshop in Aalborg

19th - 21st September

Rasmus Røge, MD, NordiQC scheme
organizer

When?

- Time consuming repeatable tasks
- Standardizable
- Output are simple or quantifiable parameter:
 - Count
 - Length
 - Area
 - Volume
 - Regions of Interest with specific characteristics
 - Categorical



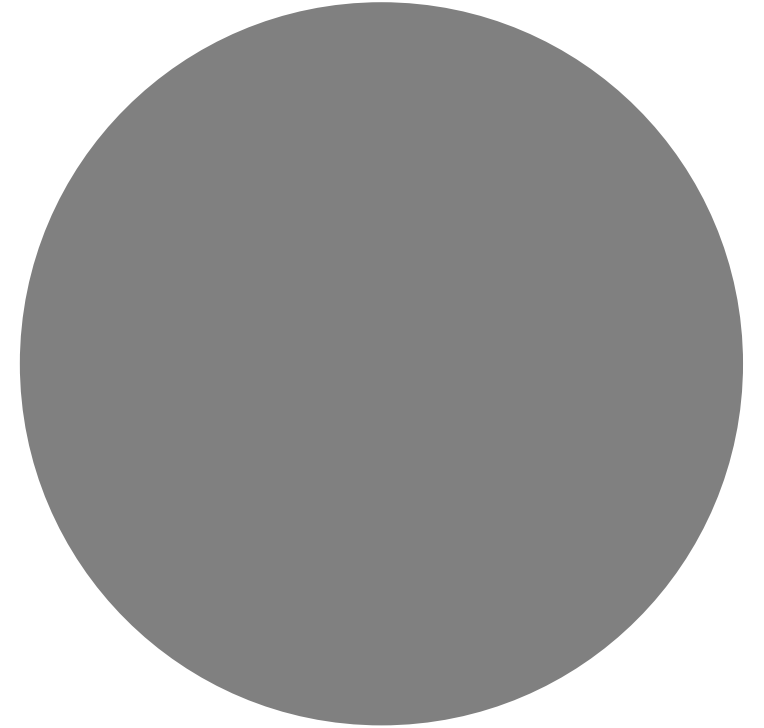
When not?

- “We will just solve that by some image analysis...”
- “Ready by Friday...?”
- Very complex setups that requires (human) interpretation
- Jobs that could easily be solved in another way



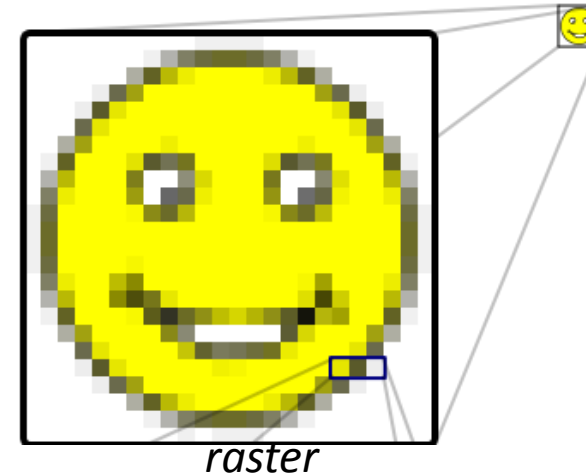
Theory

Image analysis in IHC - overview,
considerations and applications

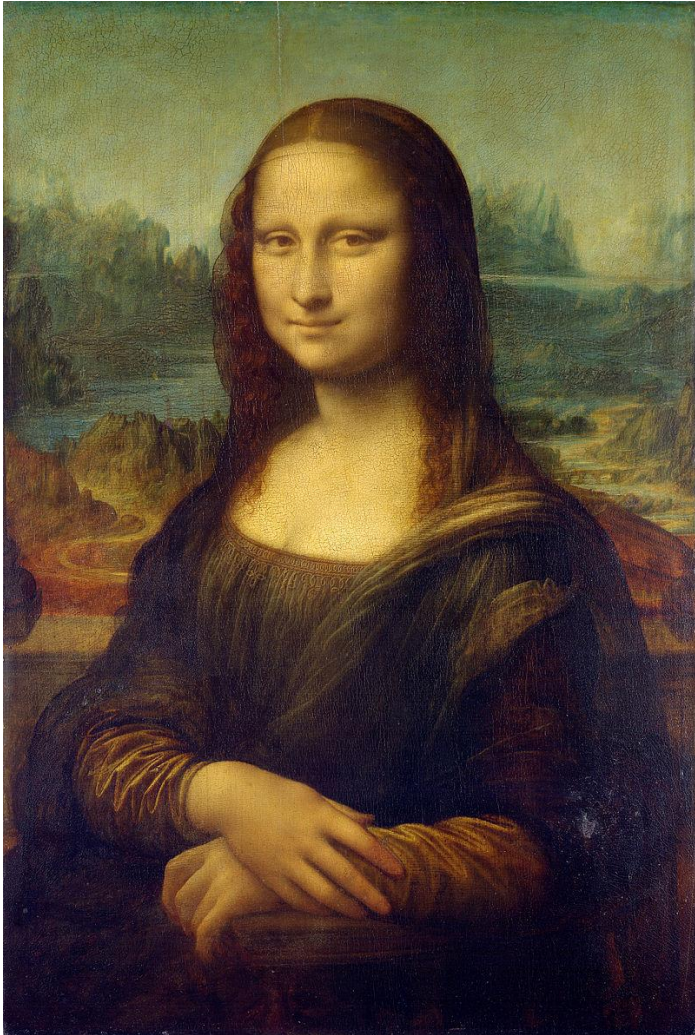


Theory

- Digital Image – numeric representation of two-dimensional image
- Either
 - Raster type: coordinate system of pixels, resolution-fixed (bmp, jpg, gif)
 - Vector type: build from primitive geometrical shapes, not-resolution-fixed (pdf, ps, fonts)

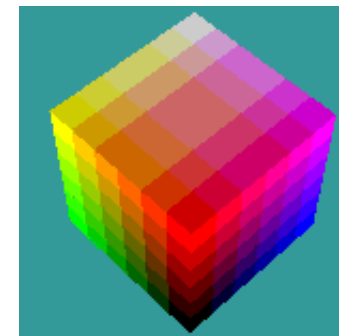
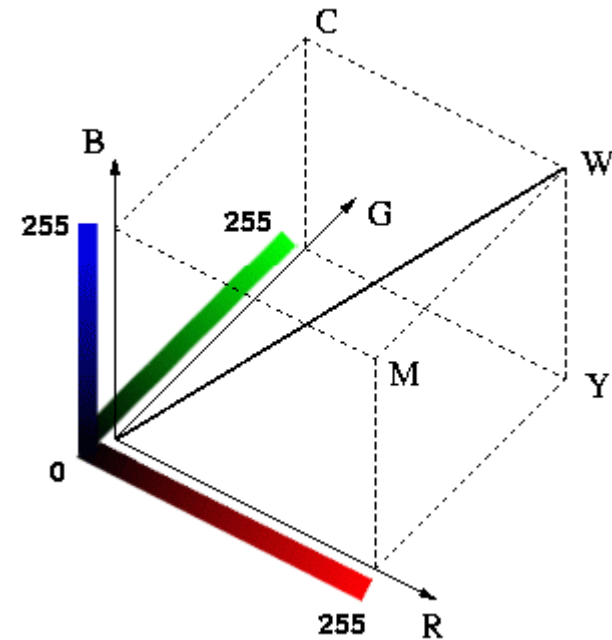


Pixels



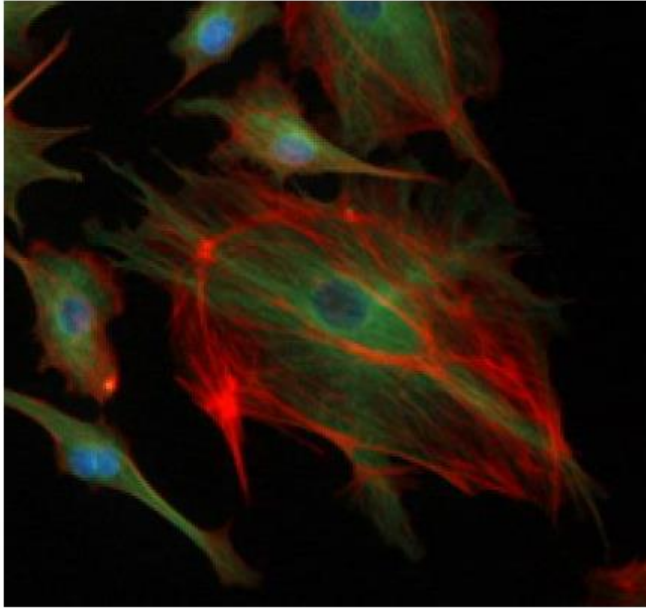
RGB colour model

- Additive colour model
- Red, green and blue light
- System to encode representation of colour

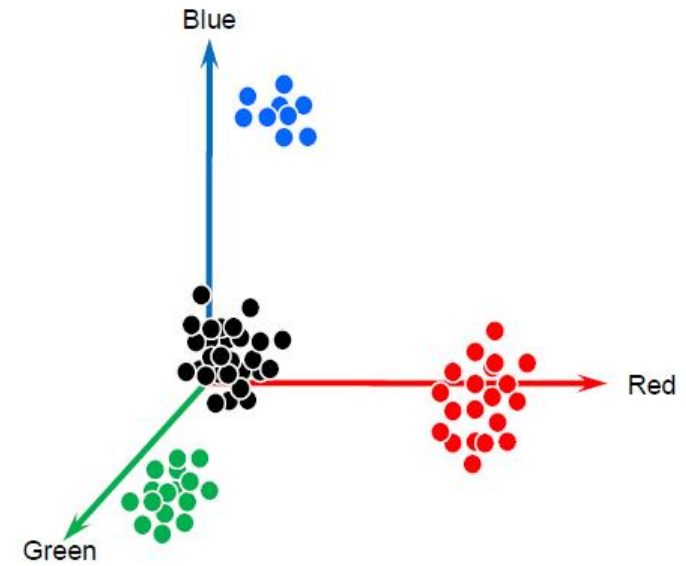


A FEW WORDS ON PIXELS

Grouping in color space



$$\bar{P}(x,y) = \begin{pmatrix} R(x,y) \\ G(x,y) \\ B(x,y) \end{pmatrix}$$



- Pixels which have similar colors will be closely grouped in color space

Color Models

IHS/HSV (Intensity, Hue, Saturation / Hue, Saturation, Value)

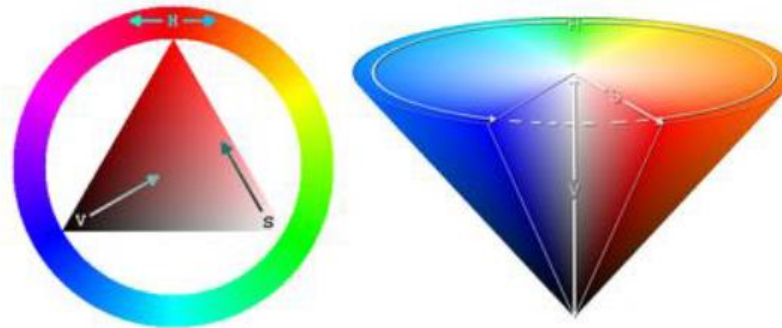
$$I = \frac{1}{3}(R + G + B)$$

$$H = \begin{cases} \theta & \text{if } B \leq G \\ 360 - \theta & \text{if } B > G \end{cases}$$

with

$$\theta = \cos^{-1} \left(\frac{\frac{1}{2}[(R - G) + (R - B)]}{\sqrt{(R - G)^2 + (R - B)(G - B)}} \right)$$

$$S = 1 - \frac{3}{(R + G + B)} \min(R, G, B)$$



From Wikimedia Commons

Color Models

Color chromaticities

- Normalize out intensity, relative amount of each RGB color component

$$r = \frac{R}{R + G + B}$$

$$g = \frac{G}{R + G + B}$$

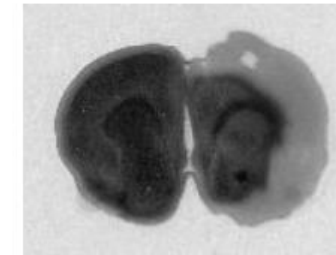
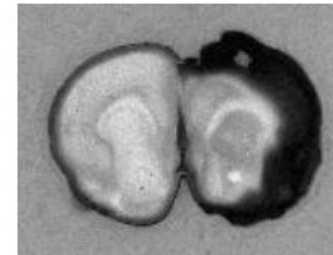
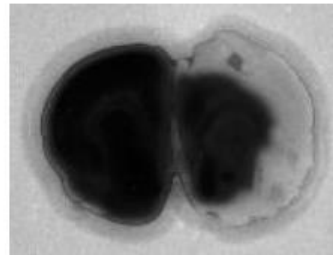
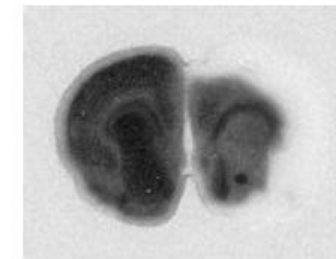
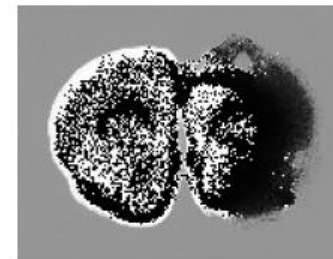
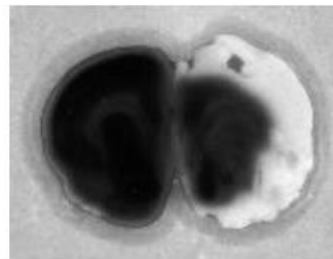
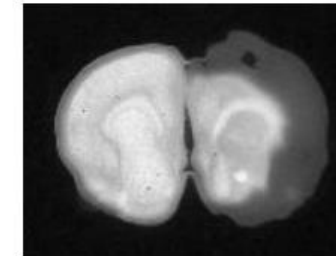
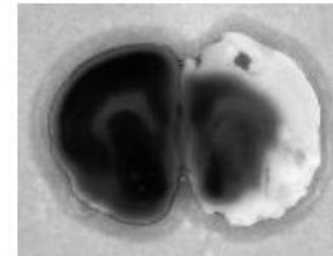
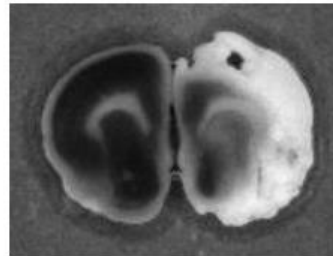
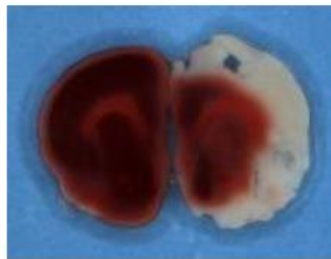
$$b = \frac{B}{R + G + B}$$

Color Models

R,G,B

I,H,S

r,g,b



Digitization – microscope / scanners

Camera mounted on microscope

- Pro
 - Area of interest
 - Quick
- Con
 - Time consuming
 - Not standardizable
 - Area of interest only

Slide scanner

- Pro
 - Standardizable
 - Quality
- Con
 - Price
 - Time
 - File size

Slide scanner

- Single or multi-slide scanner
- Whole experiment on same scanner!
- Whole experiment after calibration

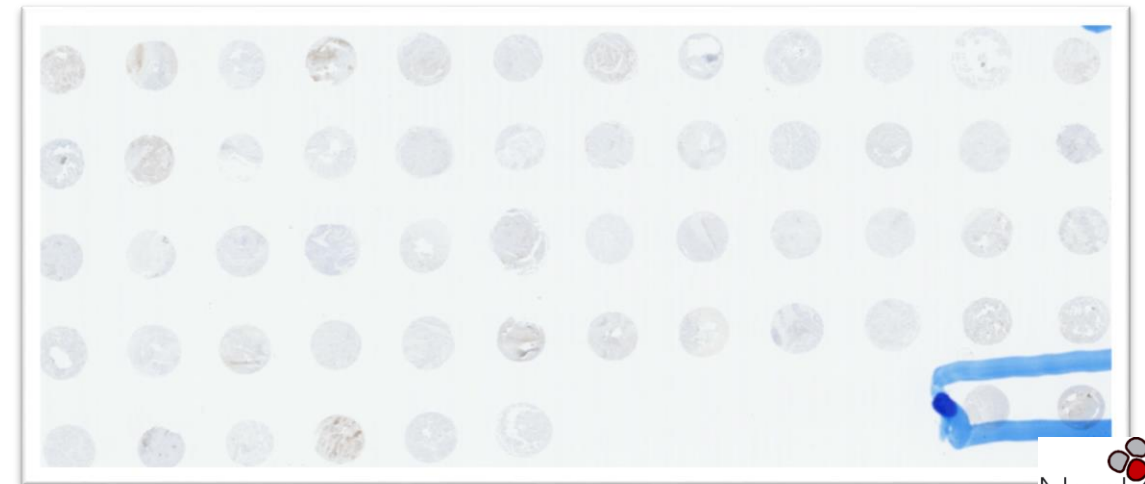
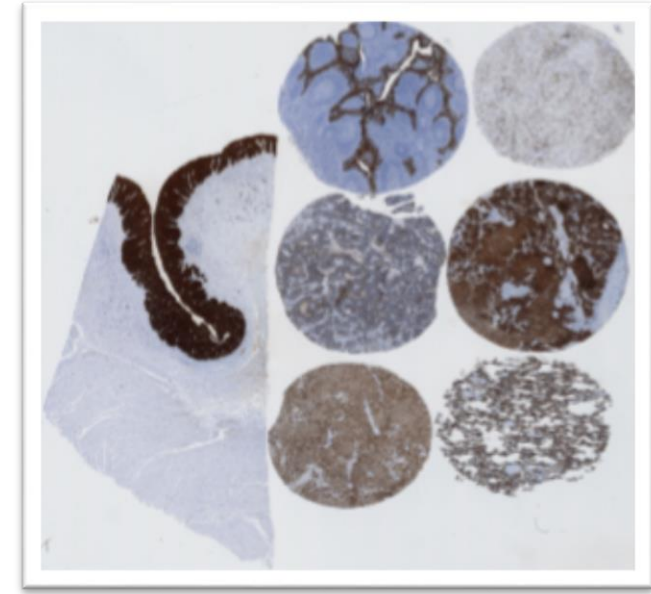


Image analysis

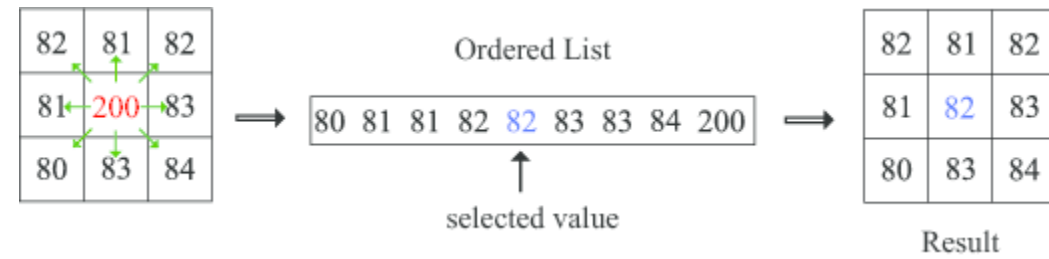
- Selection of filters
- Preprocessing – optimization of image to classification
 - Noise filtering, enhancement
- Classification / Segmentation
- Post processing
- Report of quantitative results

Selection of relevant tissue

- TMA will often contain several irrelevant or less interesting areas
- Algorithm will analyse whole image or ROI (Region of interest)
- Manually or automatic detection of ROI?



Noise filtering



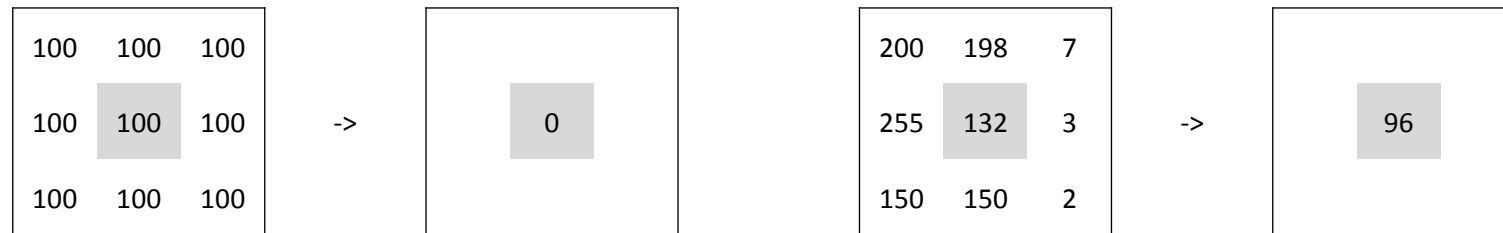
Edge Enhancement

Standard deviation filter



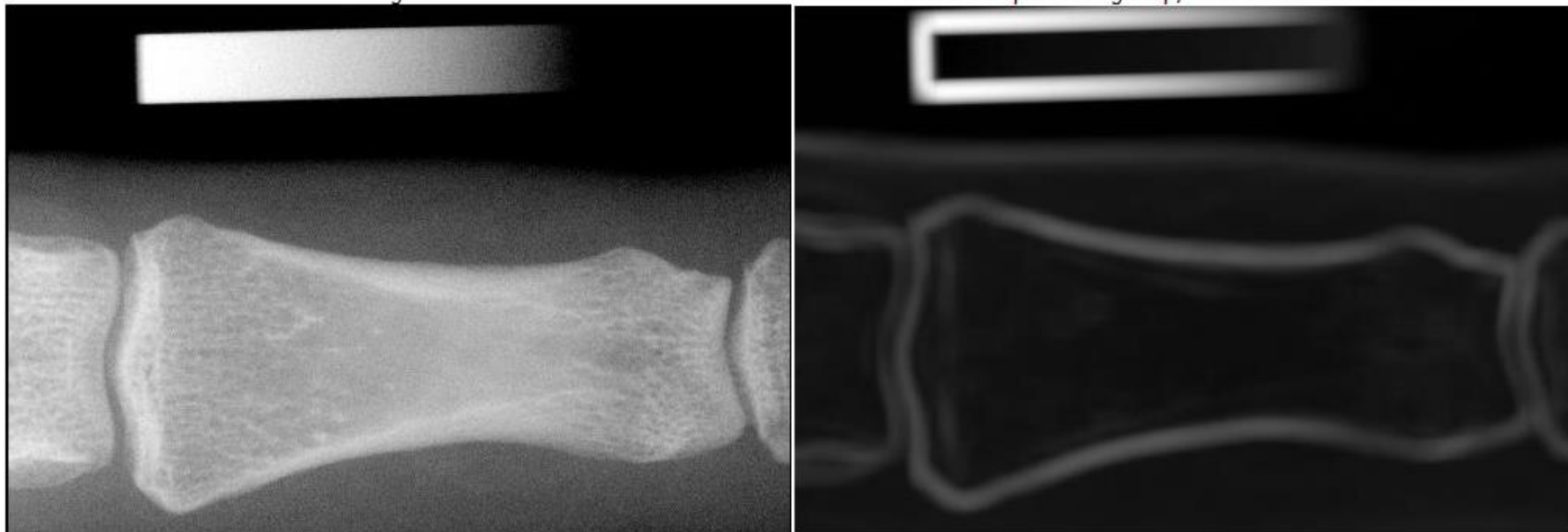
Edge Enhancement

Standard deviation filter



Original

After the processing step, with a filter size of 35x35.

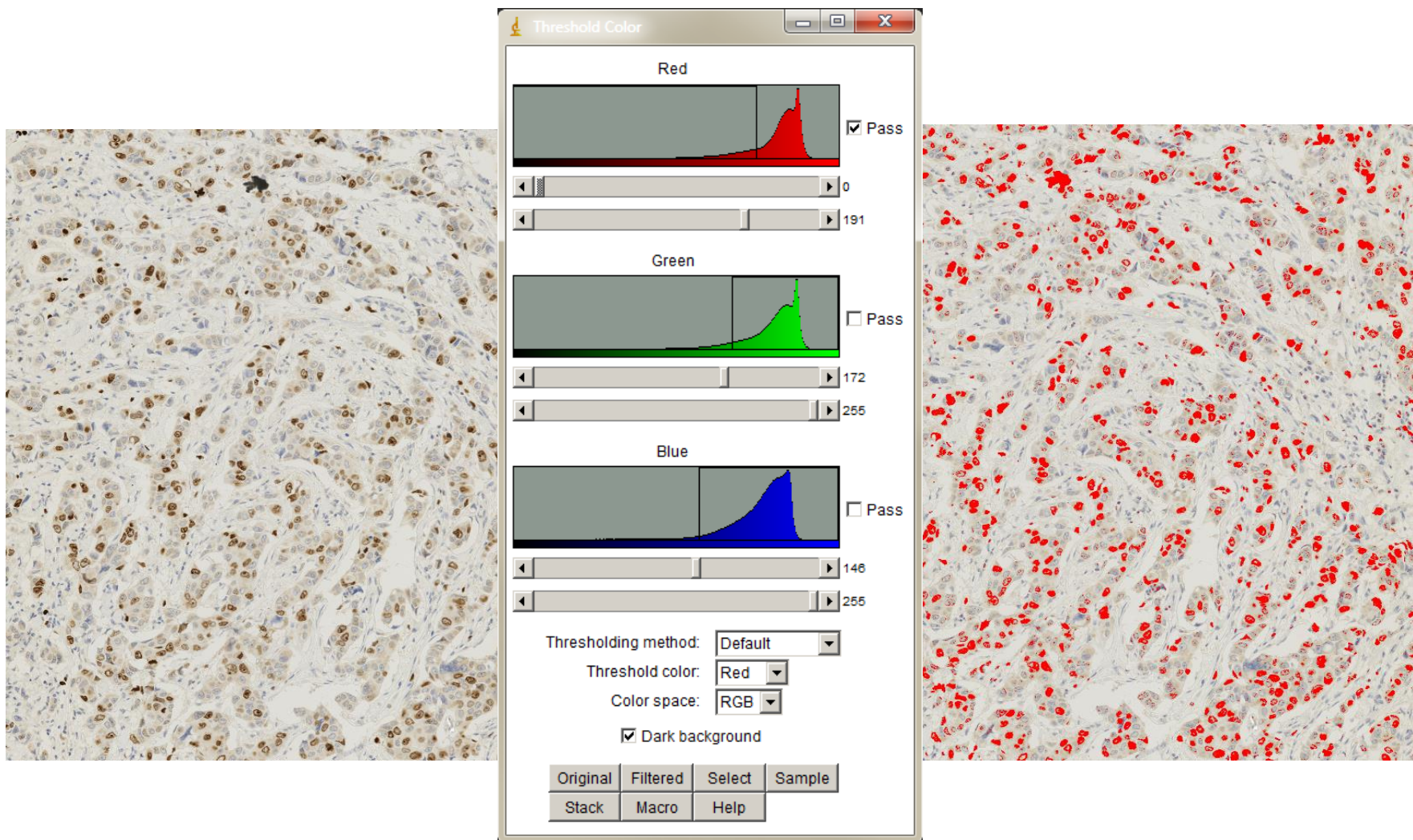


Classification / segmentation

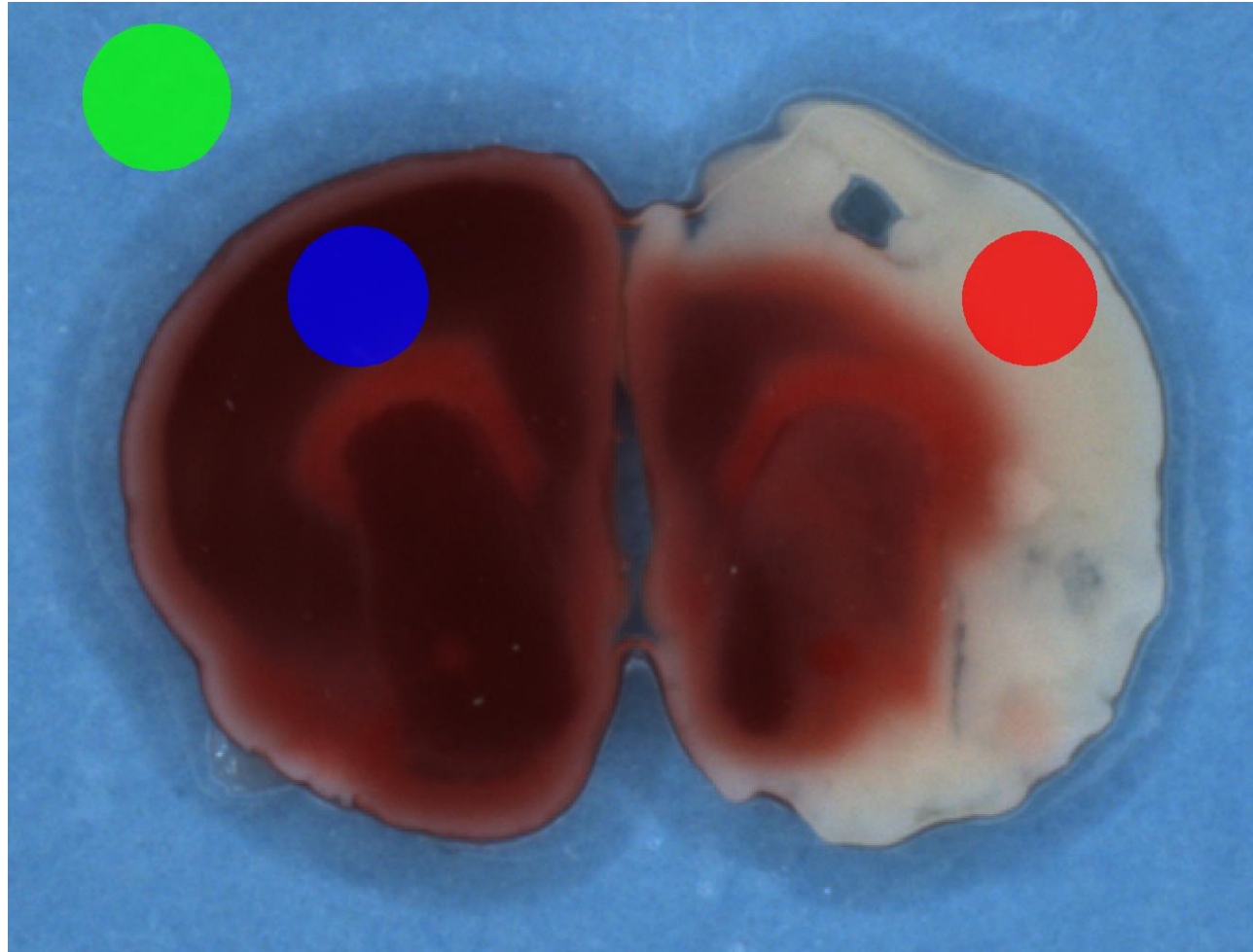
- Algorithms that group every pixels according to defined criteria
- Can be unsupervised or supervised
 - Simple: based on threshold
 - Complex: several thresholds, probabilistic (Bayesian), model-fitting (K-means), texture



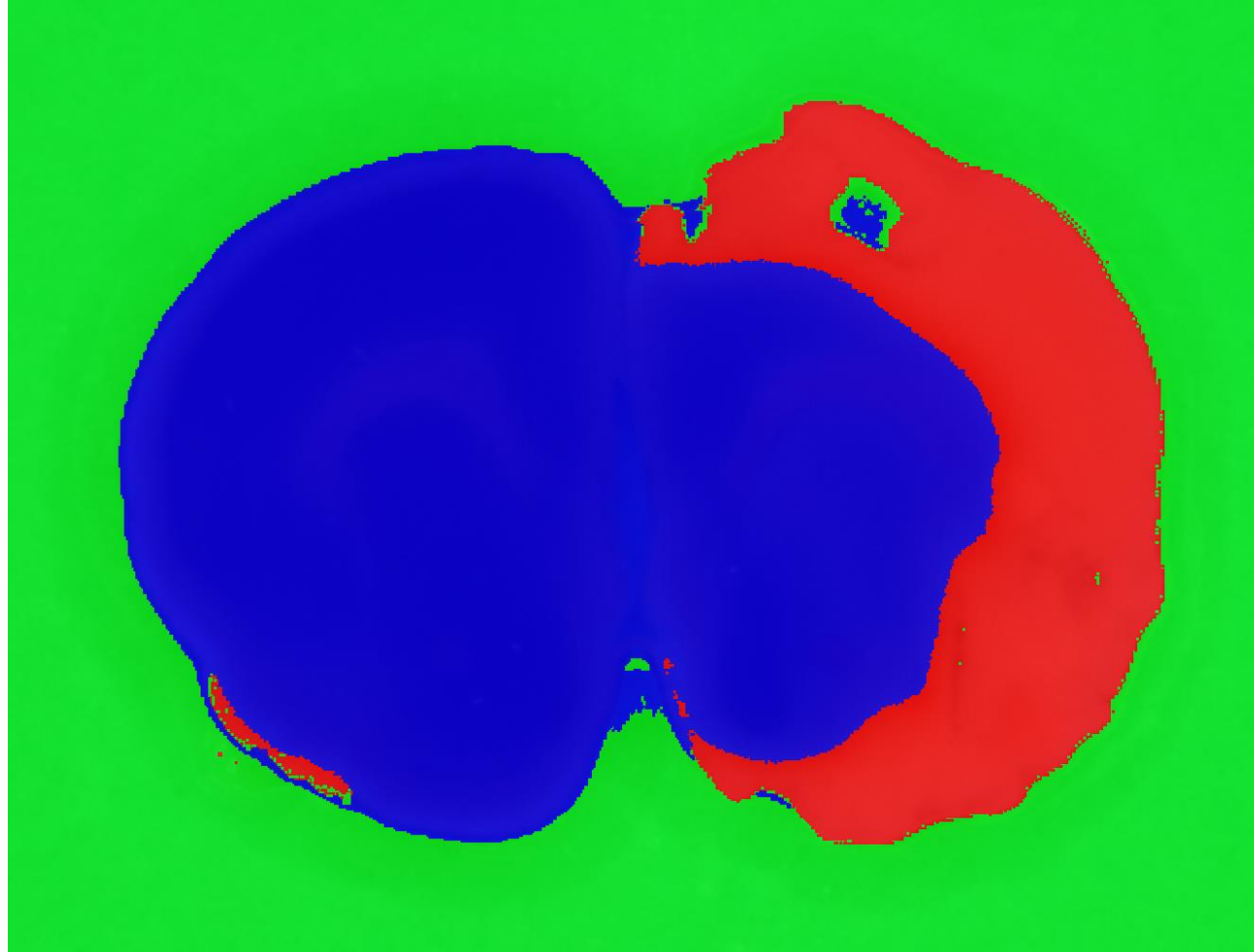
Threshold



Bayesian

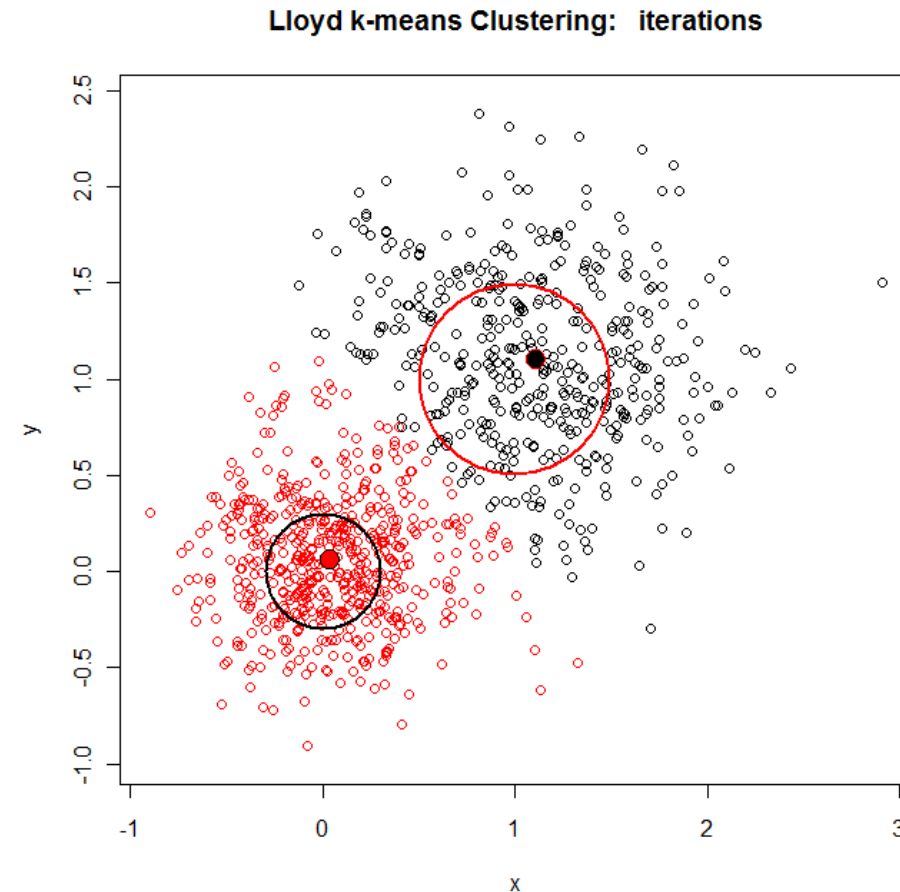


Bayesian

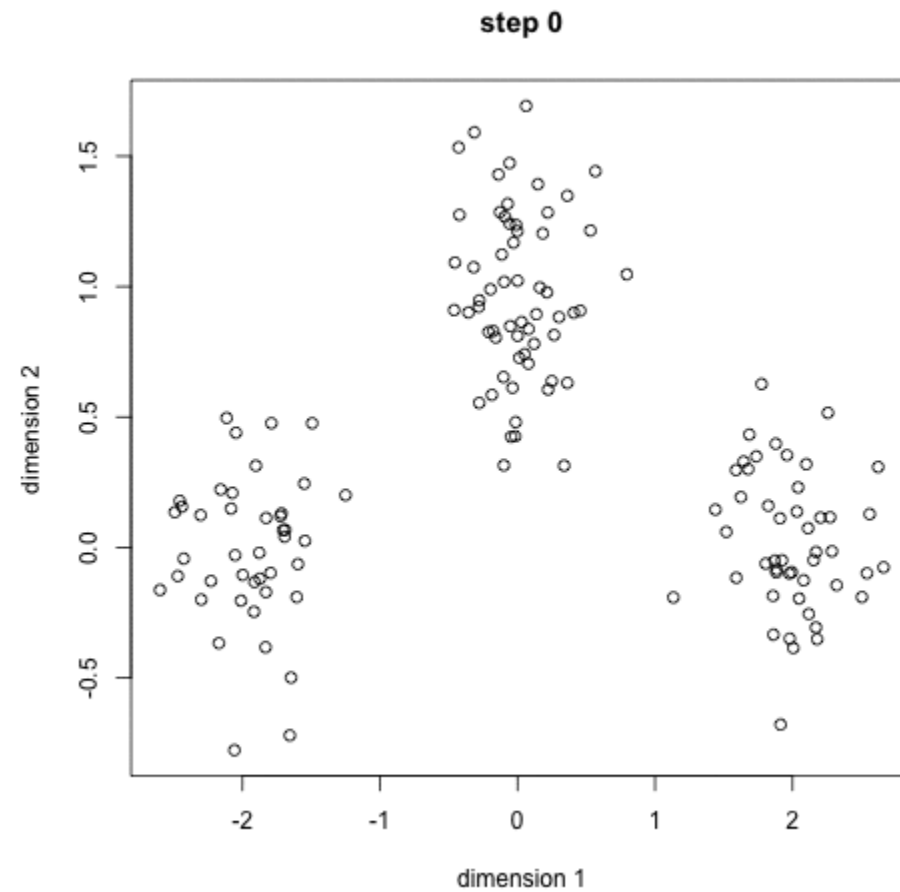


K-means

- Clustering algorithm
- Manually select number of categories (K)
- Randomly select K points (center of groups)
- Assign all point to category according to euclidian distance to center
- Calculate new center
- Repeat as needed



K-means



K-means

$K = 2$



$K = 3$



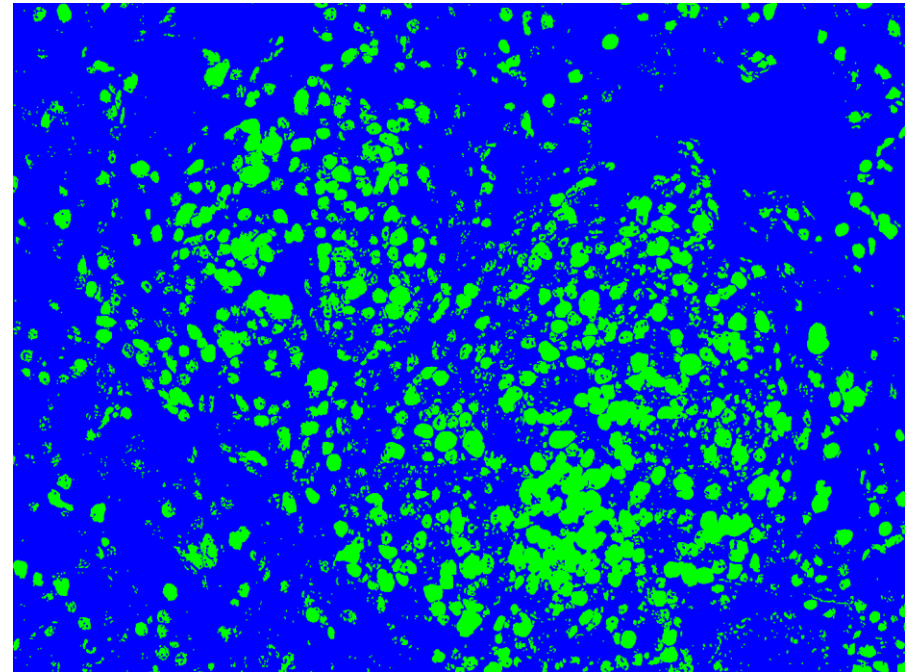
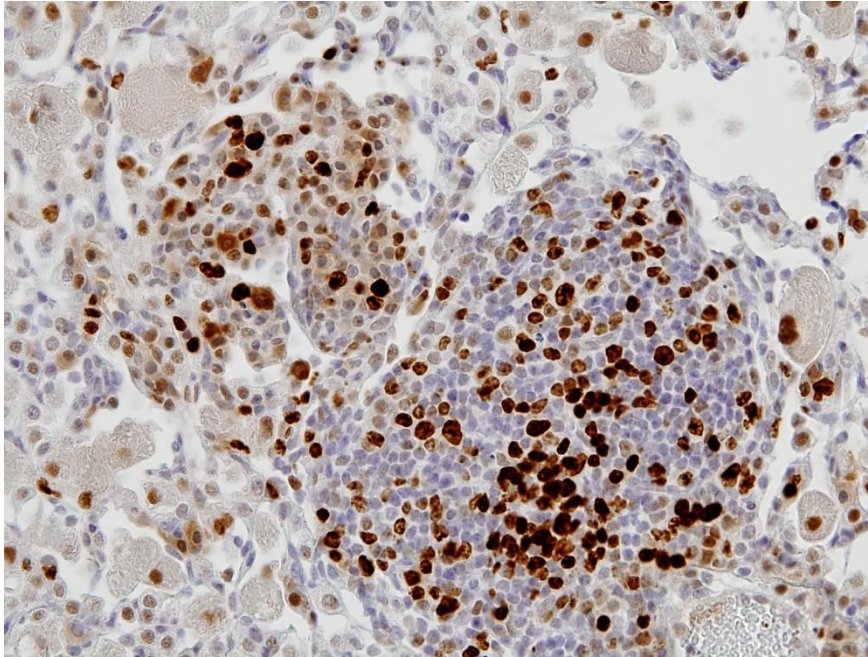
$K = 10$



Original image

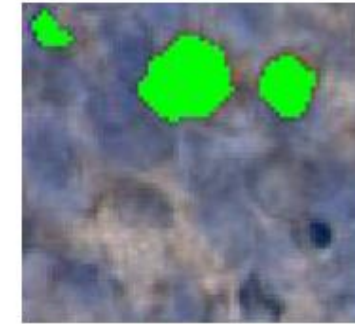
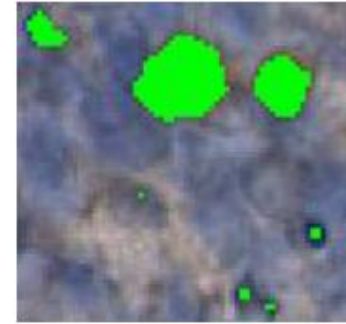


K-means

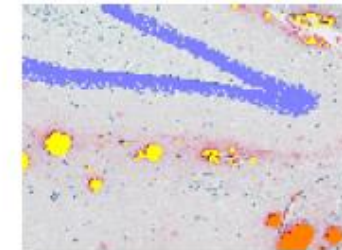
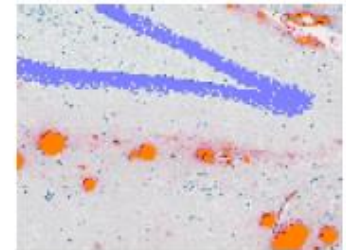


Post processing

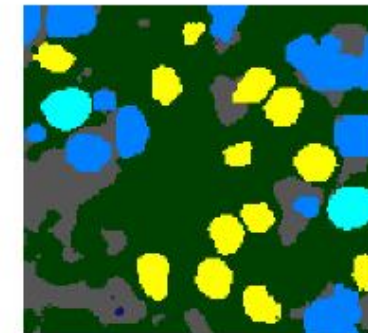
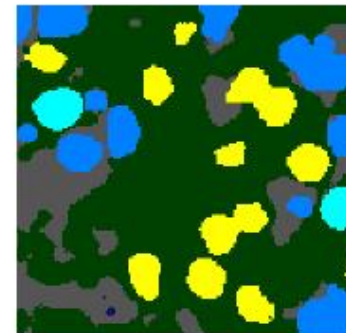
Clean-up / Noise removal:
elimination of small or large
objects



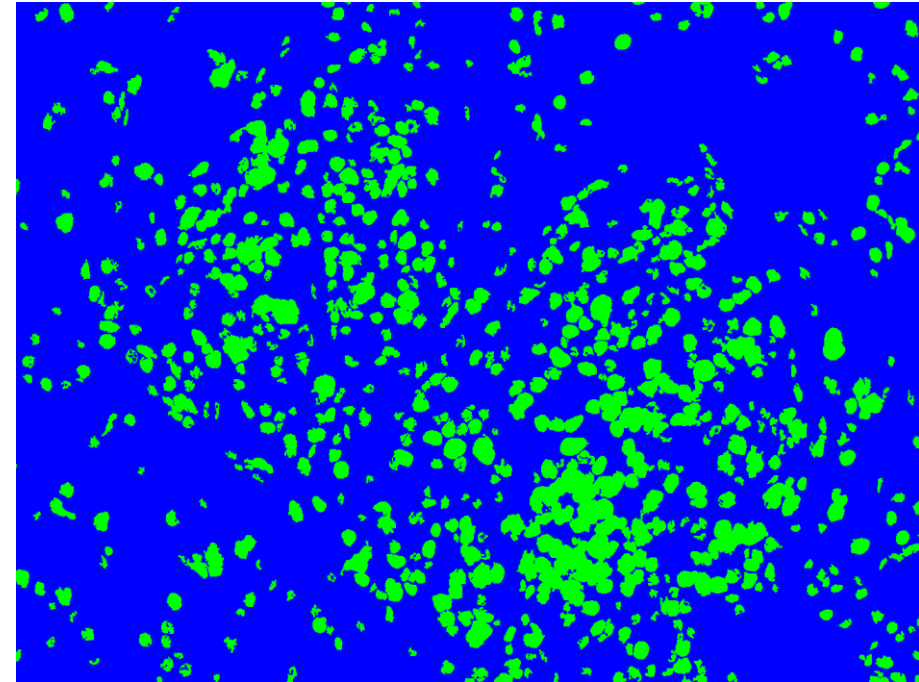
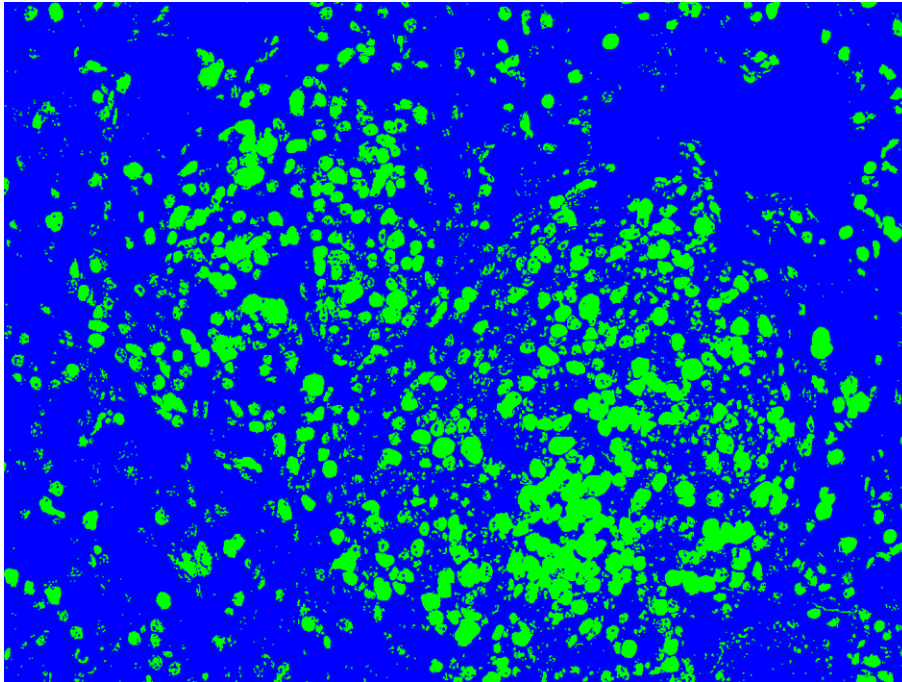
Discriminate objects based on
distance to other objects



Separate objects, change based
on shape or surroundings, erode,
dilate, open, close, skeletonize,
mark maxima, ...

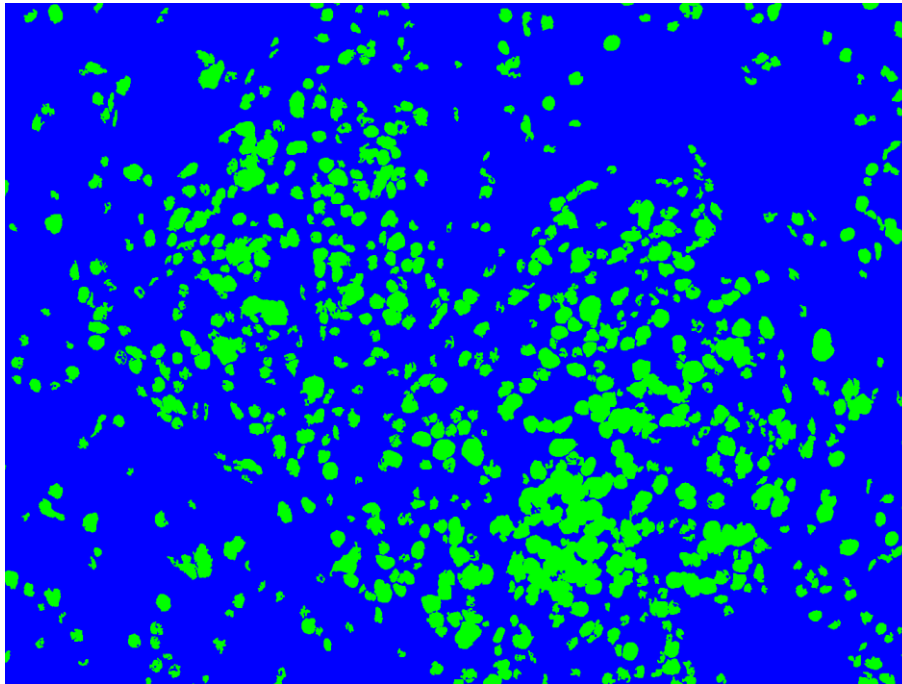


Post processing

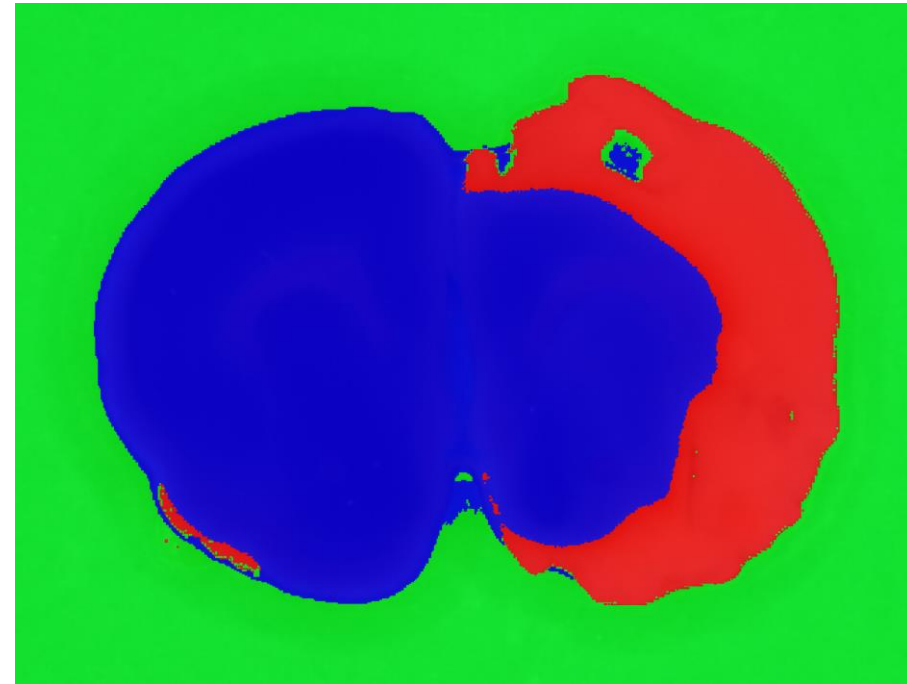


Post-processing:
Small green area, replaced by blue
Small blue area, replaced by green

Report of quantitative results



COUNT:
Typical number or fraction of
objects

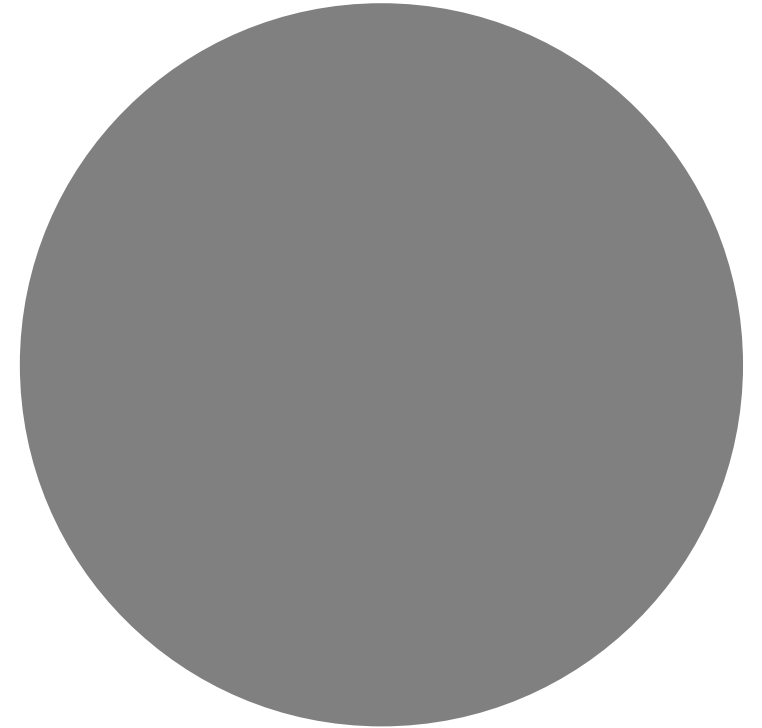


AREA:
Area of each category

Image analysis – example 1

Image analysis in IHC - overview,
considerations and applications

Ki67 & Virtual Double Staining



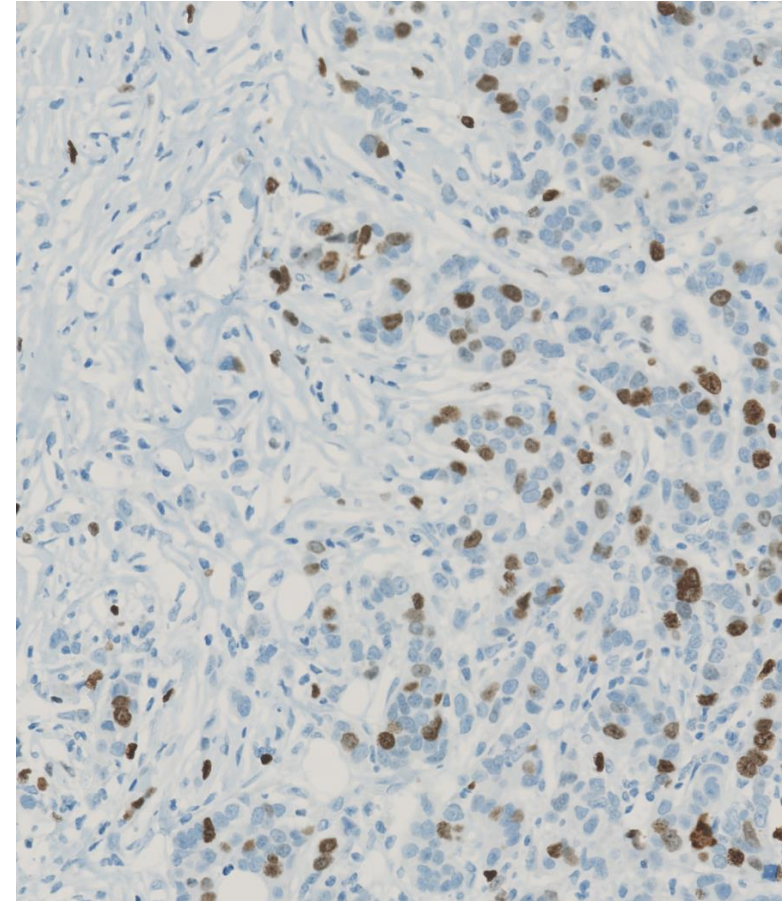
Ki67 – why is it important?

- Breast cancer:
 - Both a prognostic and predictive marker
 - Cut-off points have been suggested
- Neuroendocrine tumours
 - Grading

Digital Image Analysis

Criteria

- Identify nuclei
- Distinguish Ki67 positive and negative nuclei
- Exclude non-tumour cells from analysis



Virtual Double Staining: concept

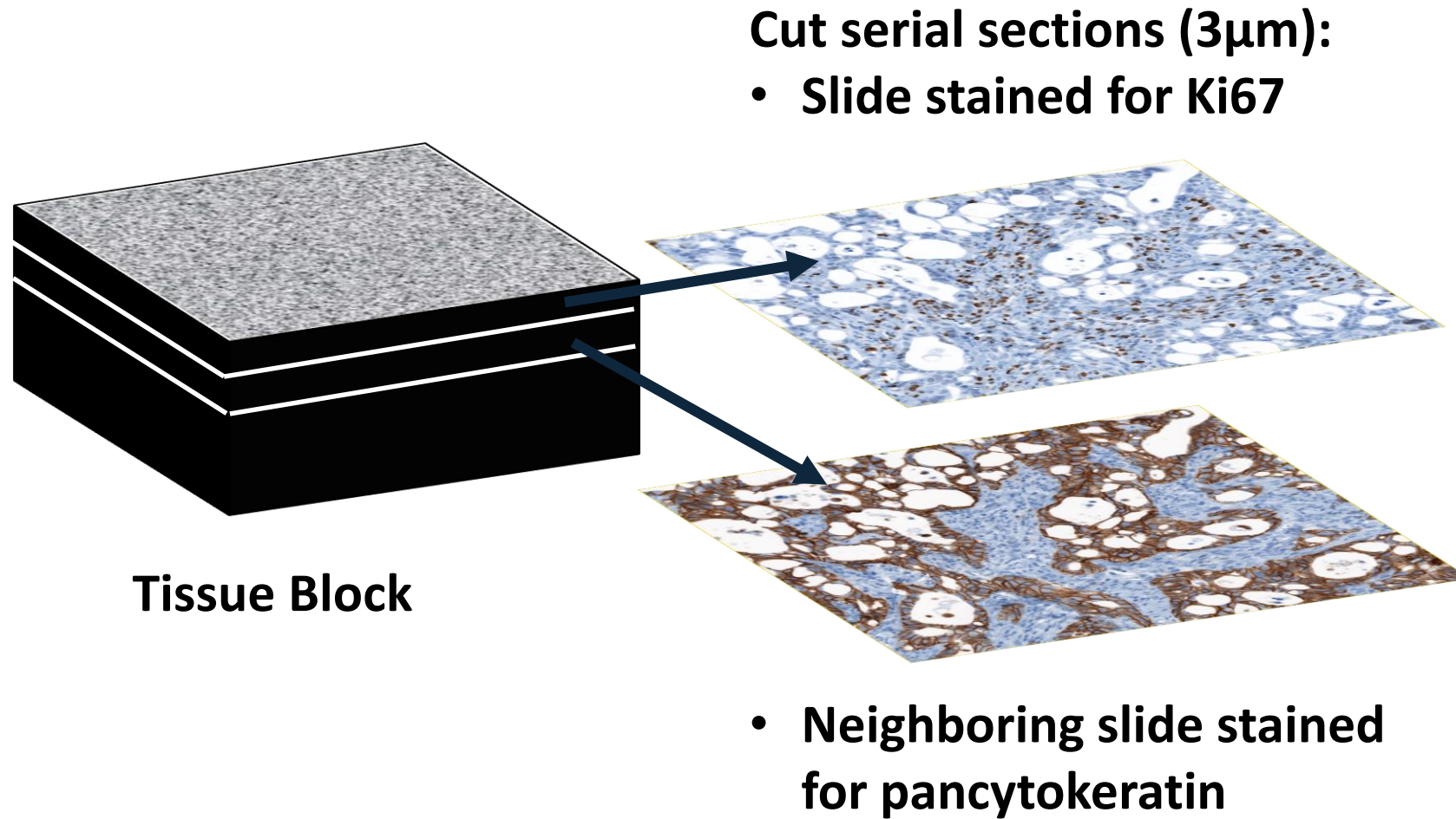
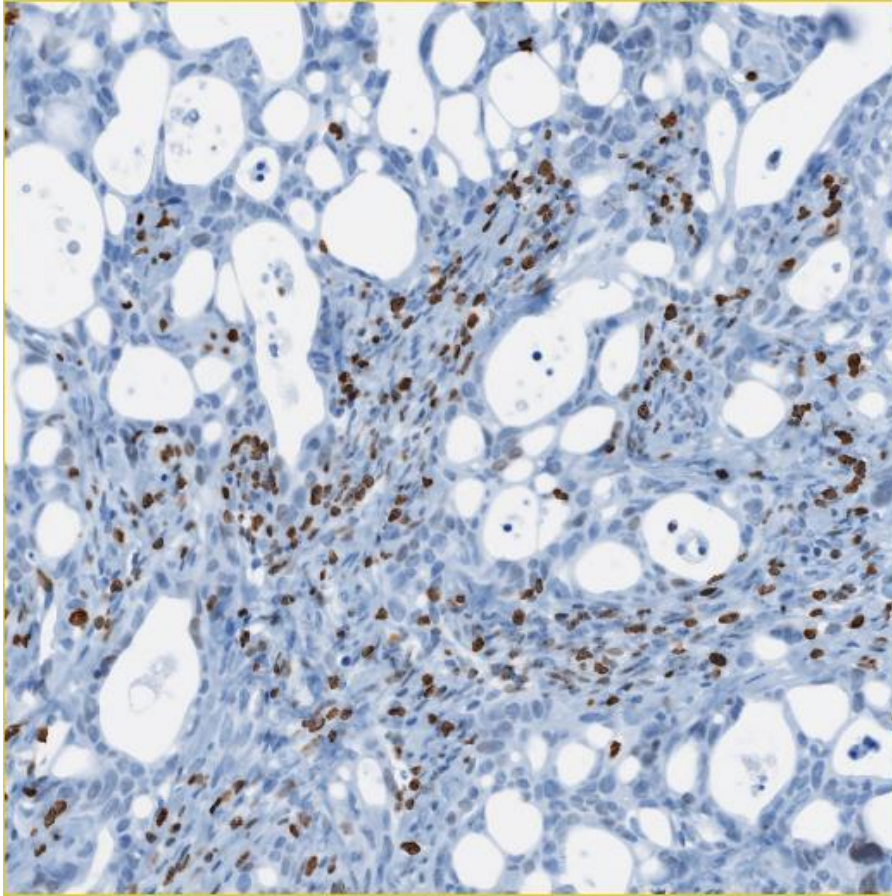
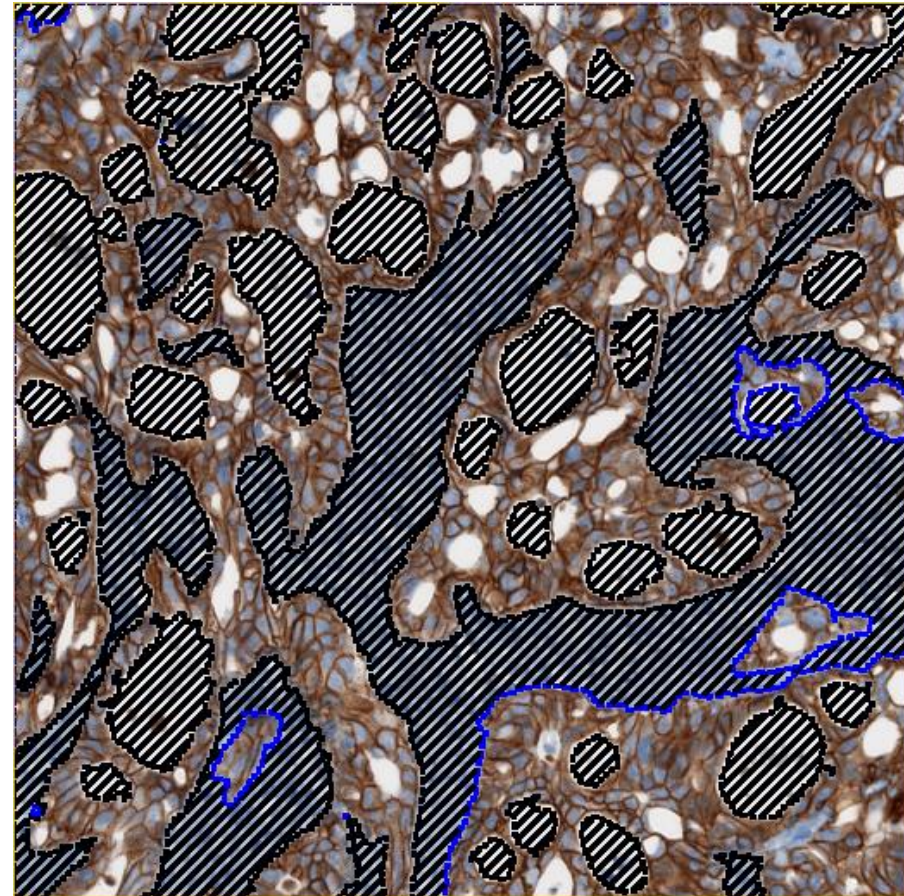


Image analysis for identification of tumor

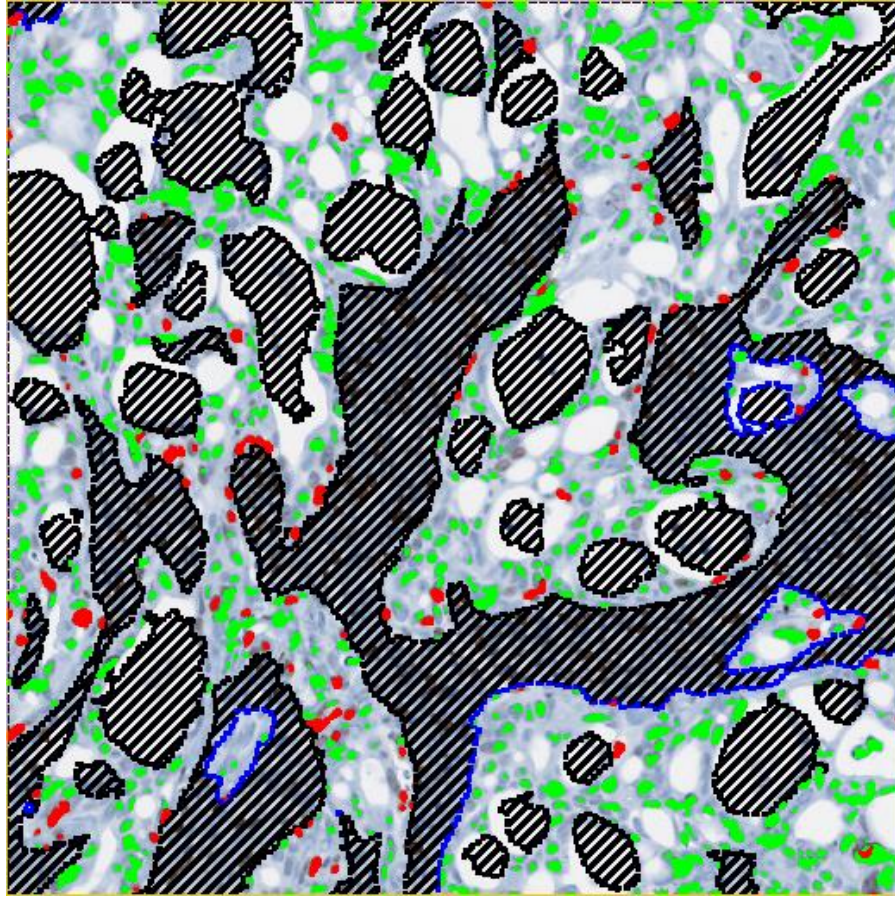


Ki67

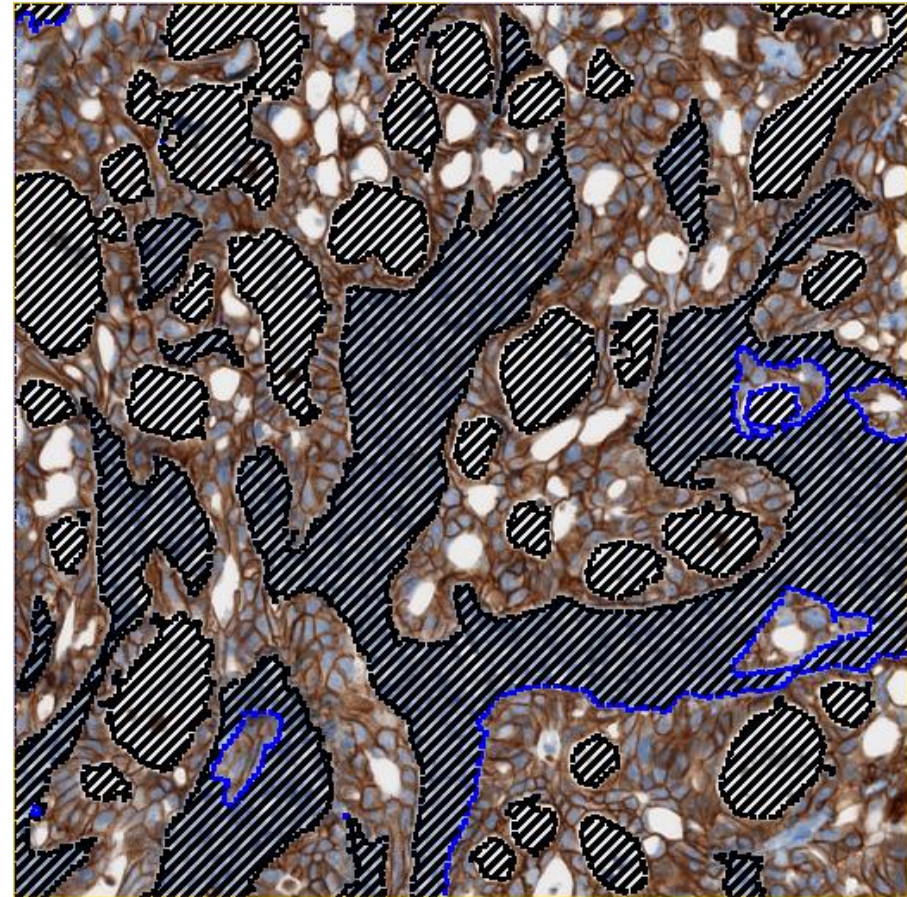


Pancytokeratin

Image analysis for identification of biomarker (Ki67)



Ki67



Pancytokeratin

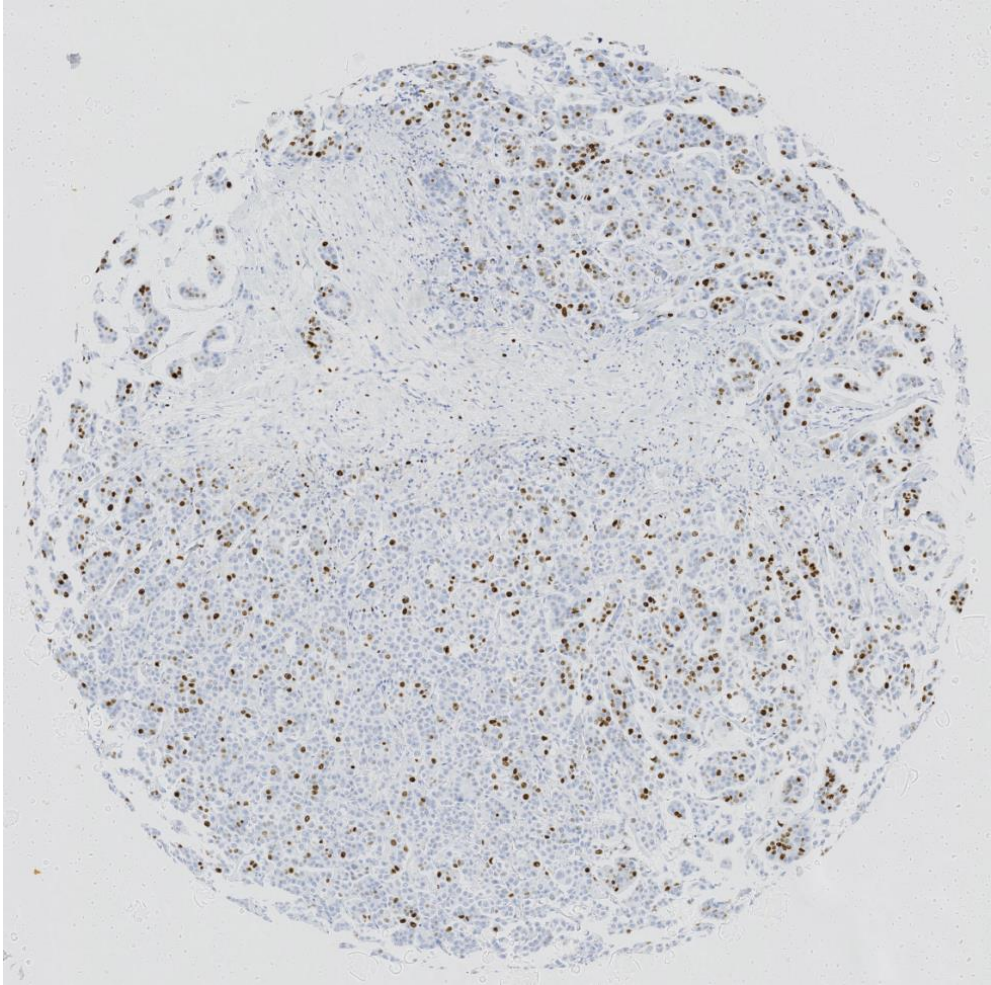
Validation of VDS + Ki67 counting

- Validation of the Nuclear detection and segmentation (number of positive and negative nuclei)
- Validation of the alignment algorithm
 - Overlap/agreement between slides
 - Sensitivity to distance between slides

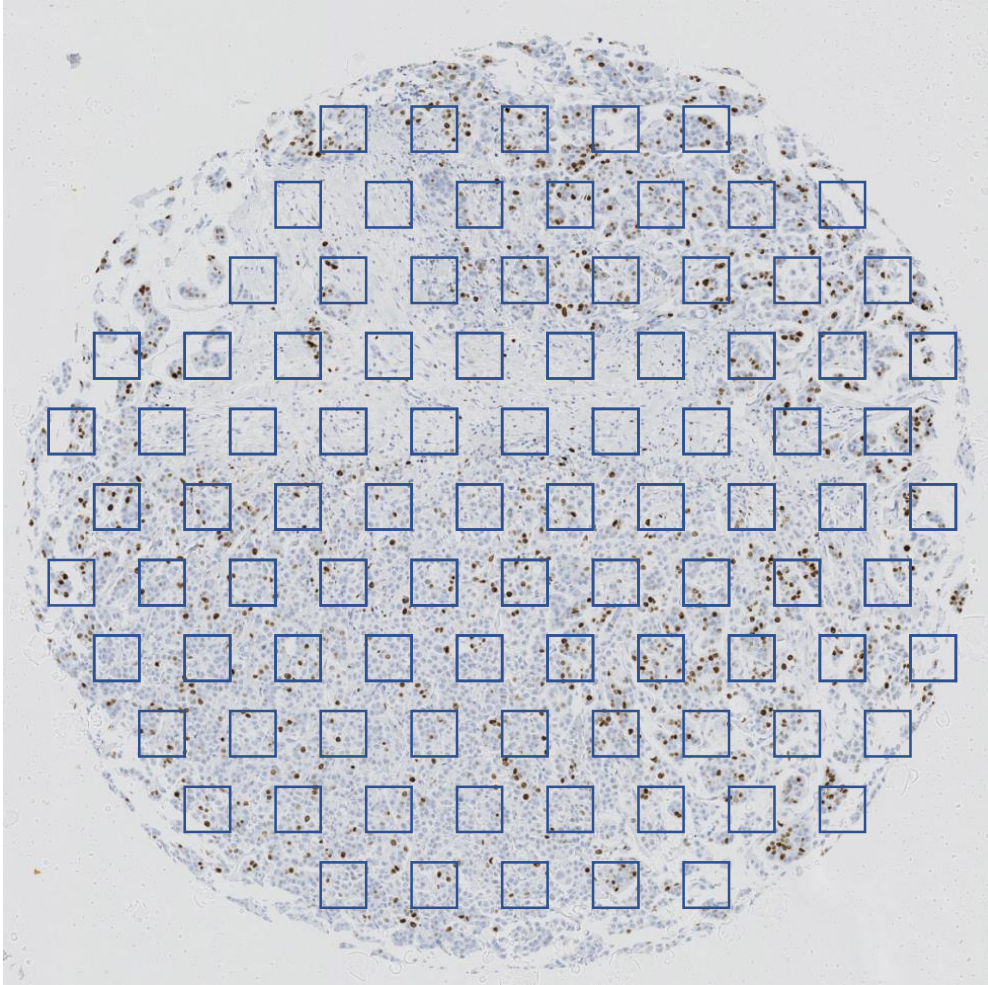
Method

- 3 TMAs containing more than 100 cores of breast carcinomas
- 2 slides were cut from each block, one stained for PCK, one for Ki67
- Areas were sampled from each core using SURS (systematic uniform randomized sampling) for manual counting
- Only a small percentage of total number of cells were counted (200-400)

Systematic Random Sampling

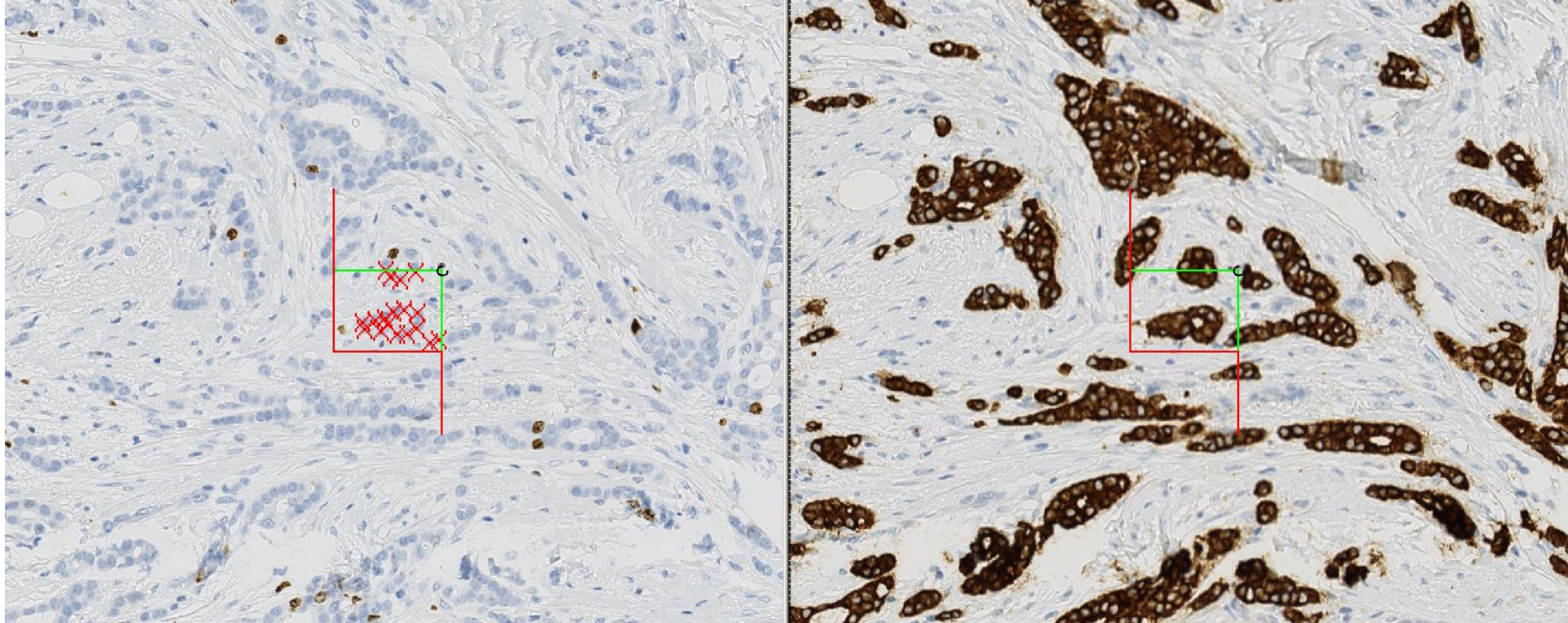


Systematic Random Sampling

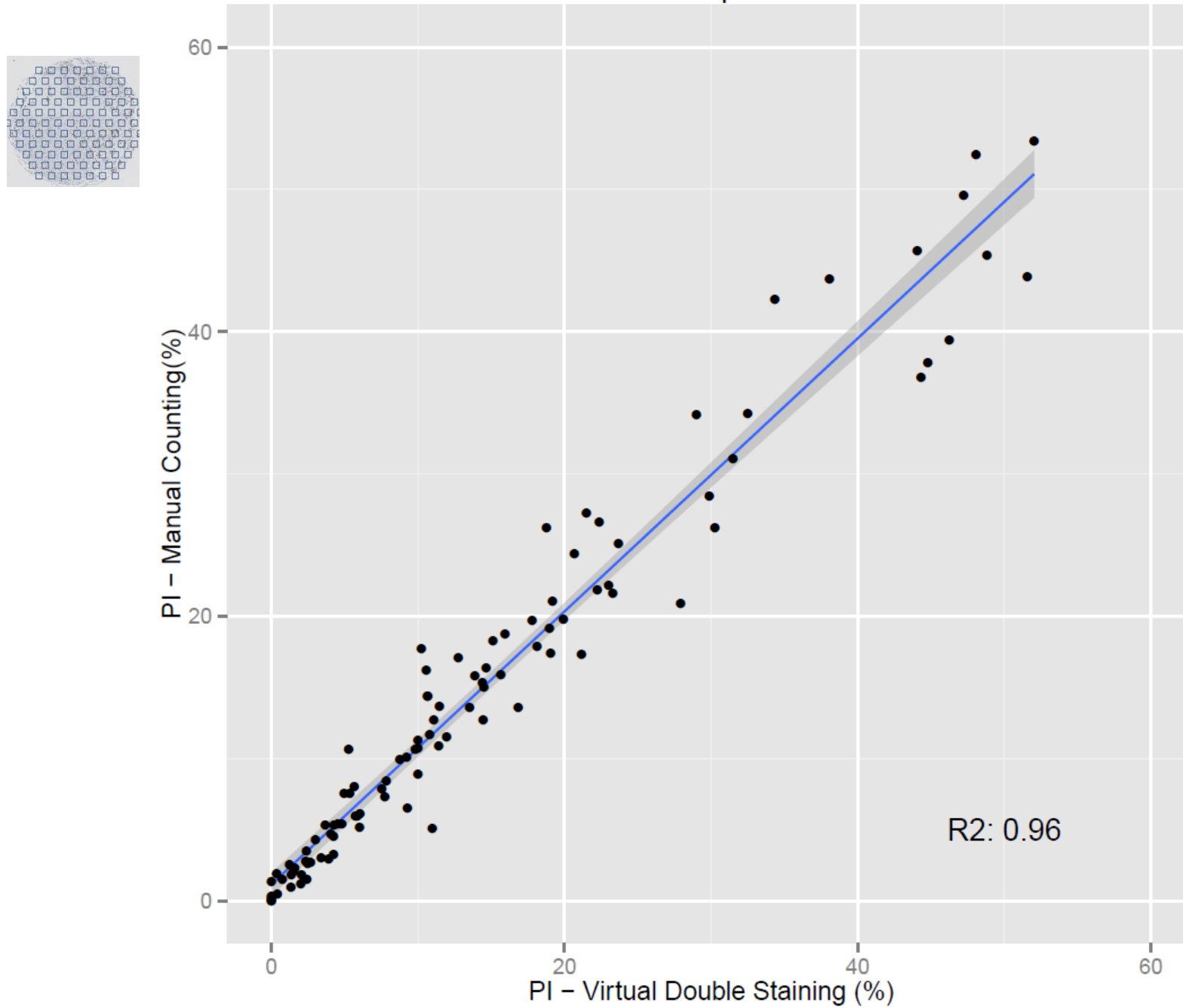


- Grid of frames randomly placed on core
- Positive and negative tumour cells counted manually in each frame
- Each frame extracted as an image for Virtual Double Staining

Stereological counting



VDS in Sampled Areas



Bland-Altman

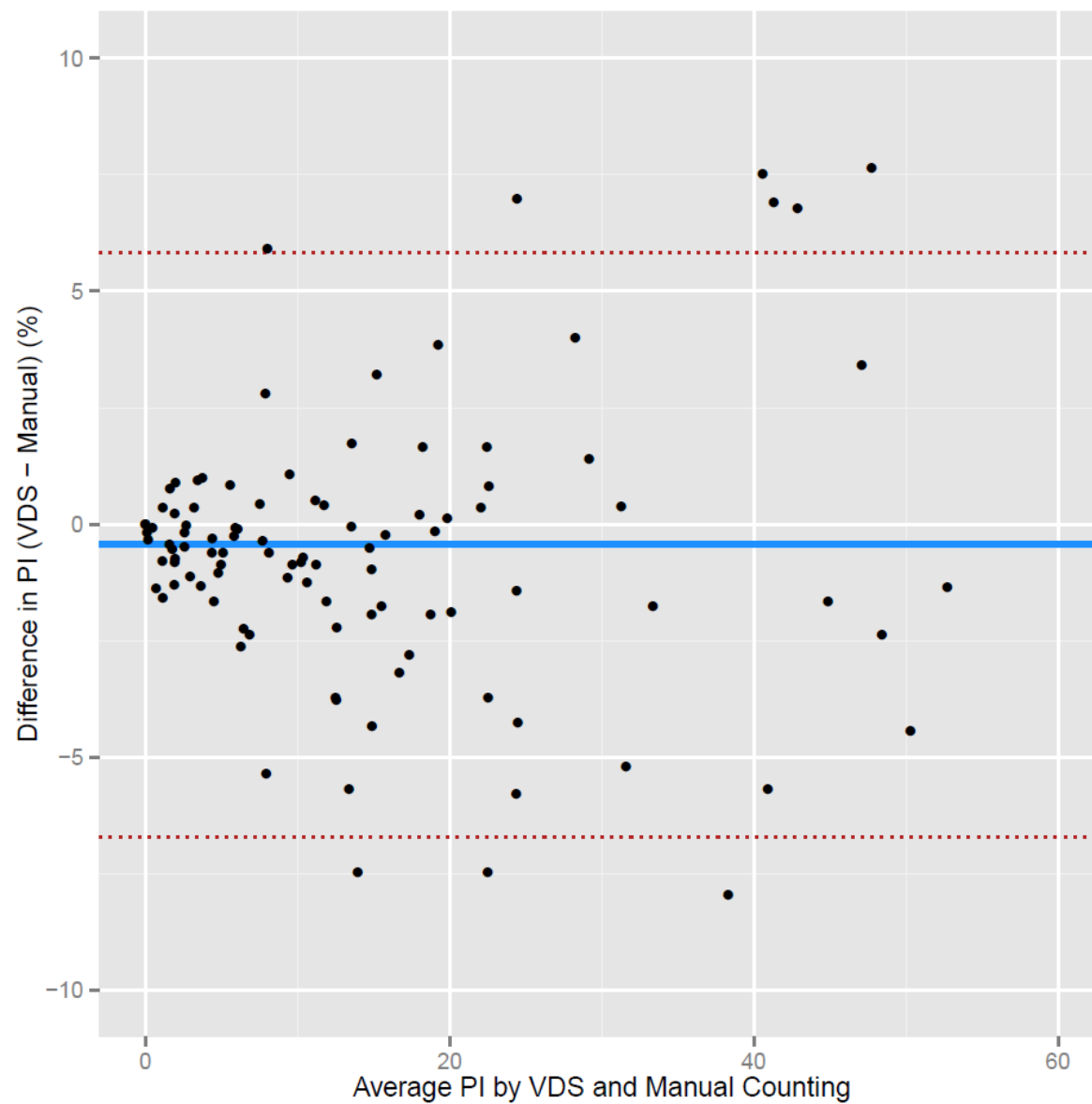
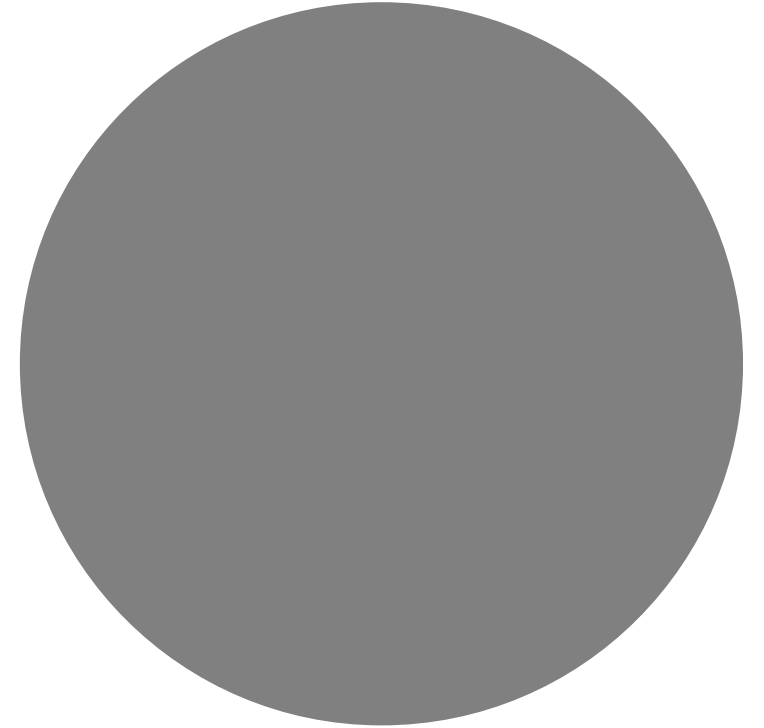


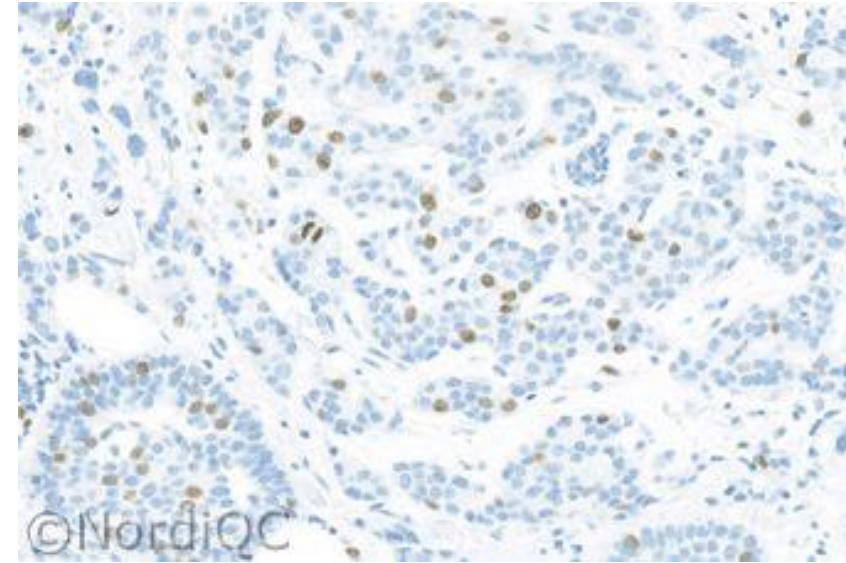
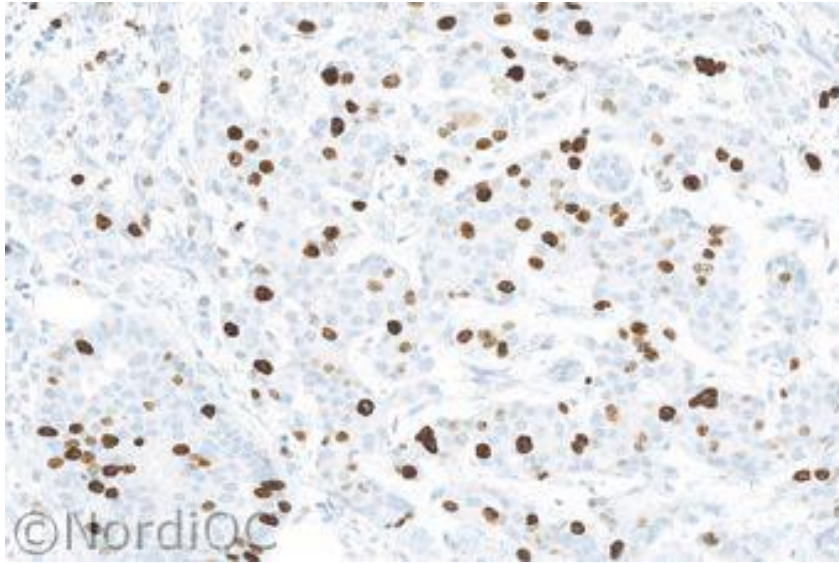
Image analysis – example 2

Image analysis in IHC - overview,
considerations and applications

Ki67 clone comparison



Ki67 – why staining quality is important



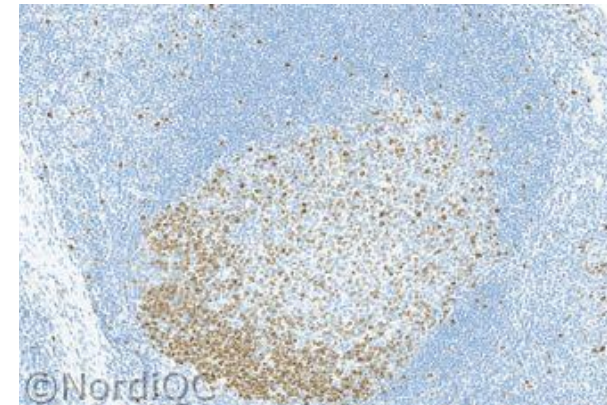
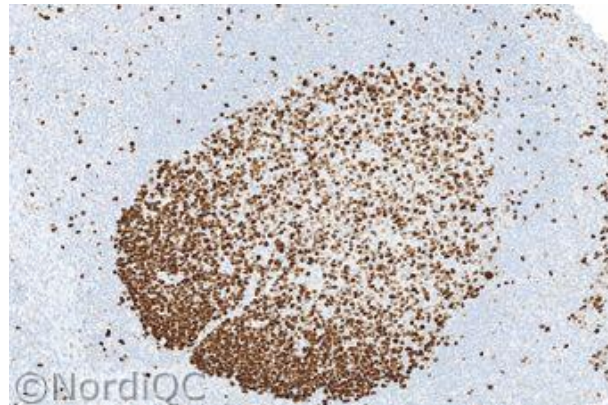
Ki67 - NordiQC

Performance in 4 NordiQC runs

	2001	2007	2009	2012
Participants	42	100	124	229
Sufficient	71%	73%	77%	89%

Performance marks in Run B13 (2012)

	Optimal	Good	Borderline	Poor
Total	166	39	18	6
Proportion	72%	17%	8%	3%



Antibody clone comparison

Immunohistochemical assessment of Ki67 with antibodies SP6 and MIB1 in primary breast cancer: a comparison of prognostic value and reproducibility

Maria Ekholm,^{1,2} Sanda Beglerbegovic,³ Dorthe Grabau,^{2,4} Kristina Lövgren,²
Per Malmström,^{2,5} Linda Hartman^{2,6} & Mårten Fernö²

Conclusions: SP6 was not superior to MIB1, but the two antibodies were comparable in the assessment of Ki67. Both MIB1 and SP6 could therefore be considered for prognostic use in primary breast cancer.

Comparative Validation of the SP6 and MIB1 
Antibodies to Ki67 and Their Use in Tissue Microarray (TMA) and Image Analysis for Breast Cancer.

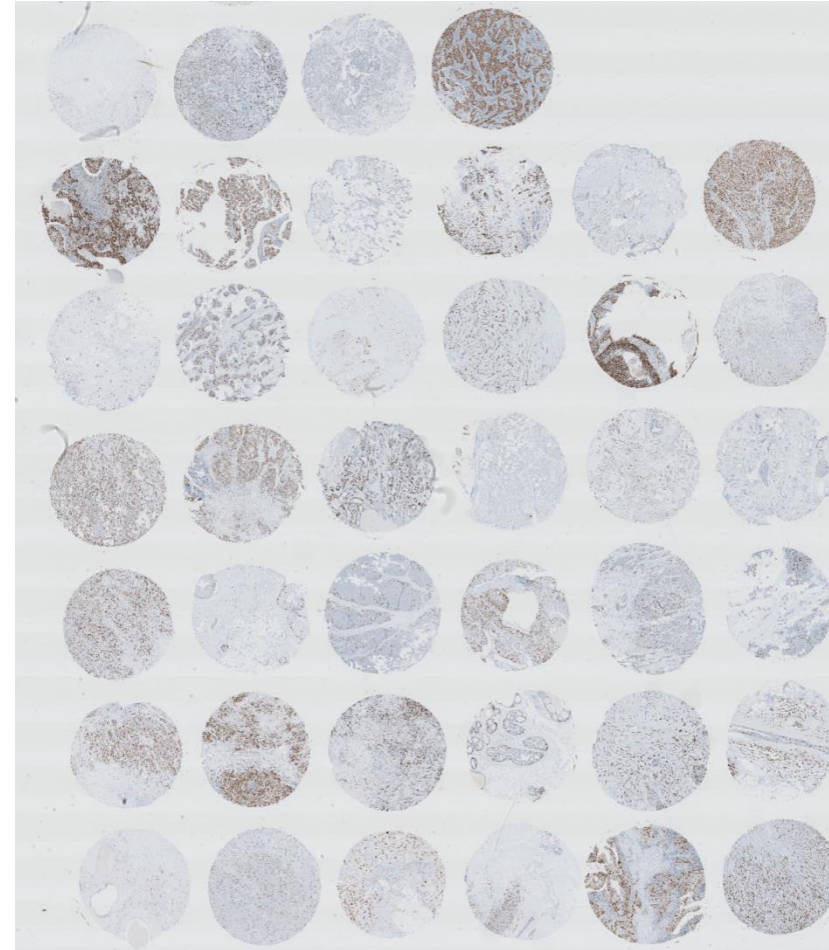
L. Zabaglo¹, L. Zabaglo², J. Salter¹, J. Salter², H. Anderson¹, H. Anderson², M. Hills¹,
R. A'Hern³, M. Dowsett¹, and M. Dowsett²

Conclusions: SP6 and MIB1

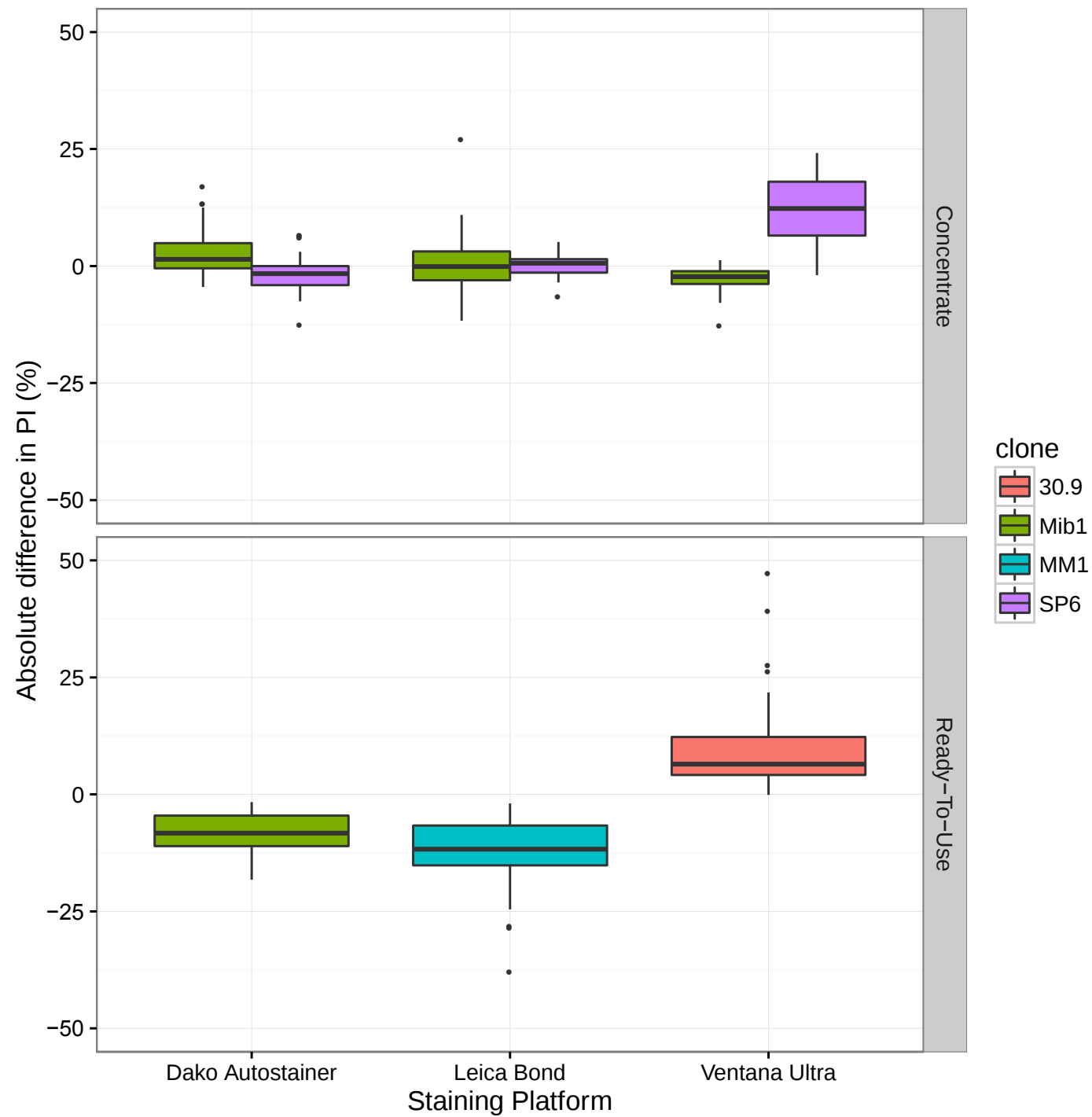
provide highly comparable measures of Ki67 that predict progression of advanced disease similarly. SP6 is substantially better suited than MIB1 to image analysis, and is now our preferred antibody for future studies.

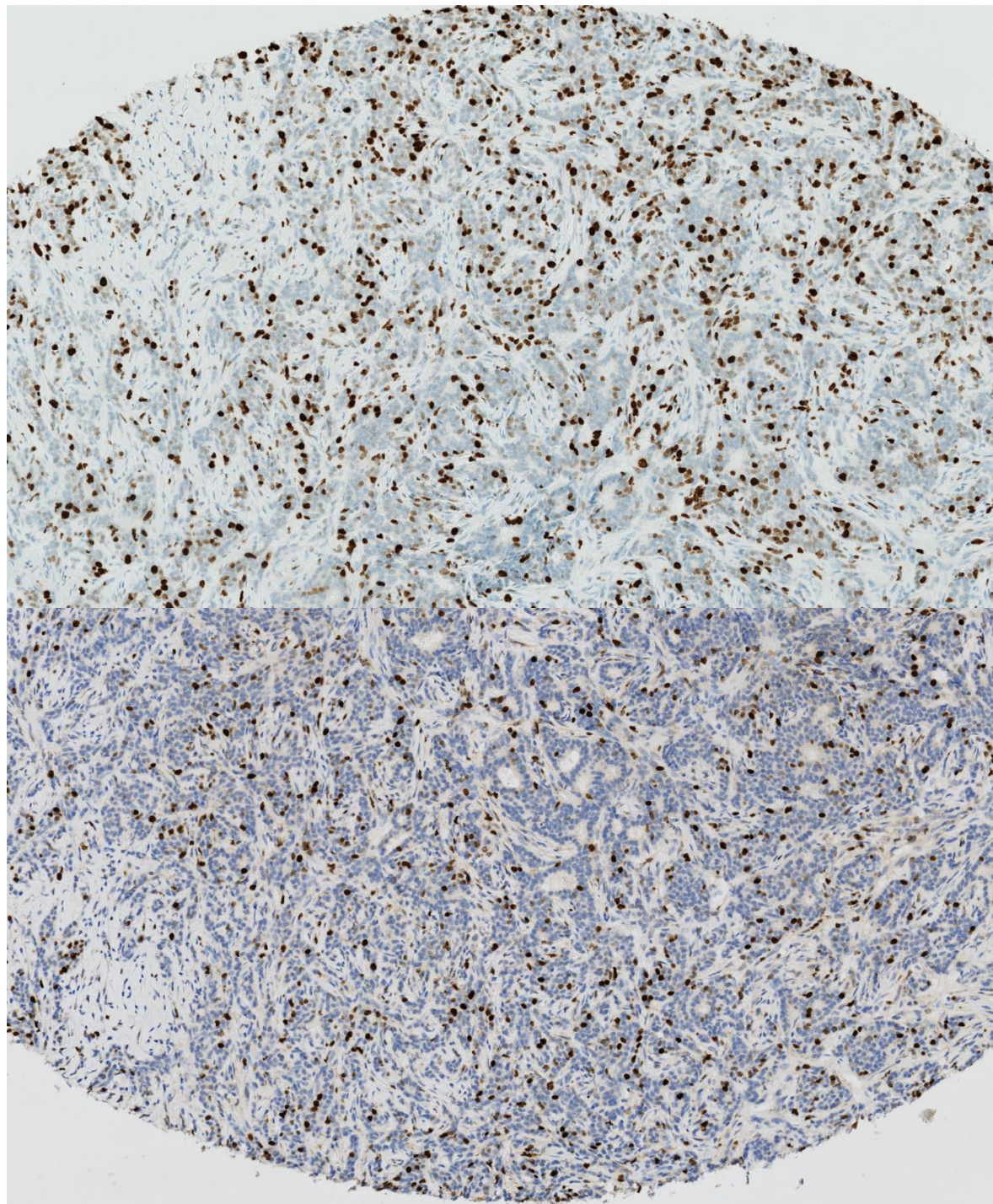
Experimental setup

- TMA with 40 breast cancers
- Stained using most commonly used mAb: Mib1, SP6, 30.9, MM1
- Stained using both (if available) Ready-To-Use format and concentrated format (In-House optimized protocol)
- Stained on all major staining platforms
- Parallel slide stained for PCK
- Proliferation Index calculated using Virtual Double Staining



Results





SP6 concentrate,
Ventana platform

Proliferation Index:
38 %

MM1 RTU,
Leica platform

Proliferation Index:
12 %

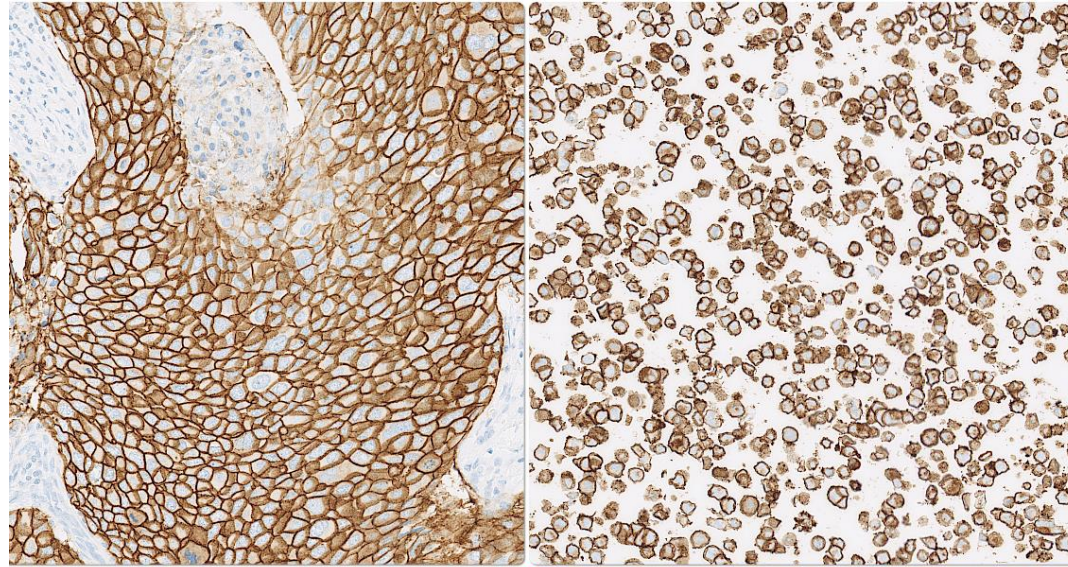
Image analysis – example 3

Image analysis in IHC - overview,
considerations and applications

HER2 connectivity and cell lines

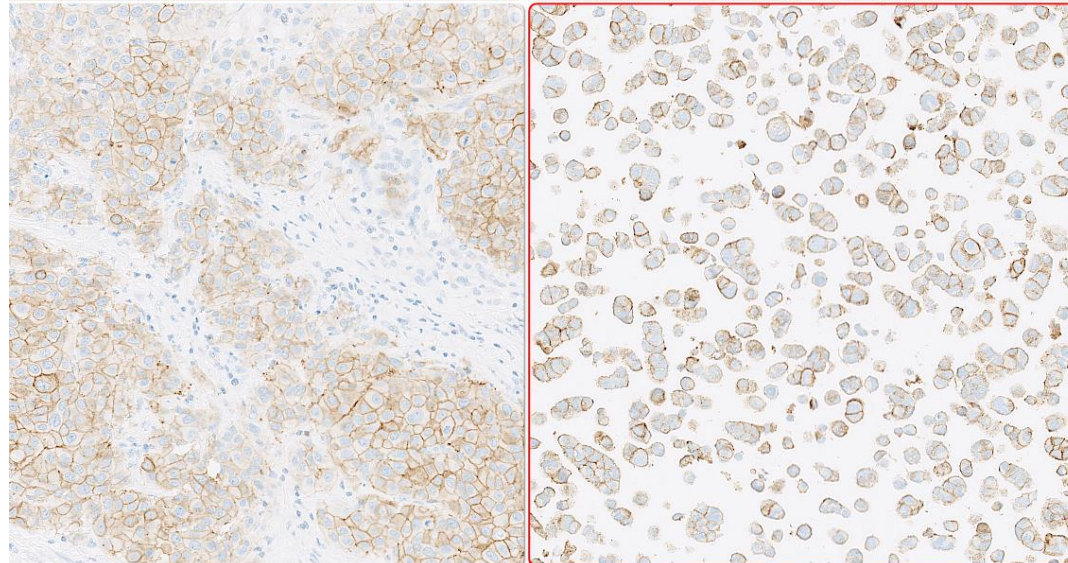
Control material for HER2 IHC: performance control / consistency

Histology:
3+ tumour



Cell lines:
3+

2+ tumour



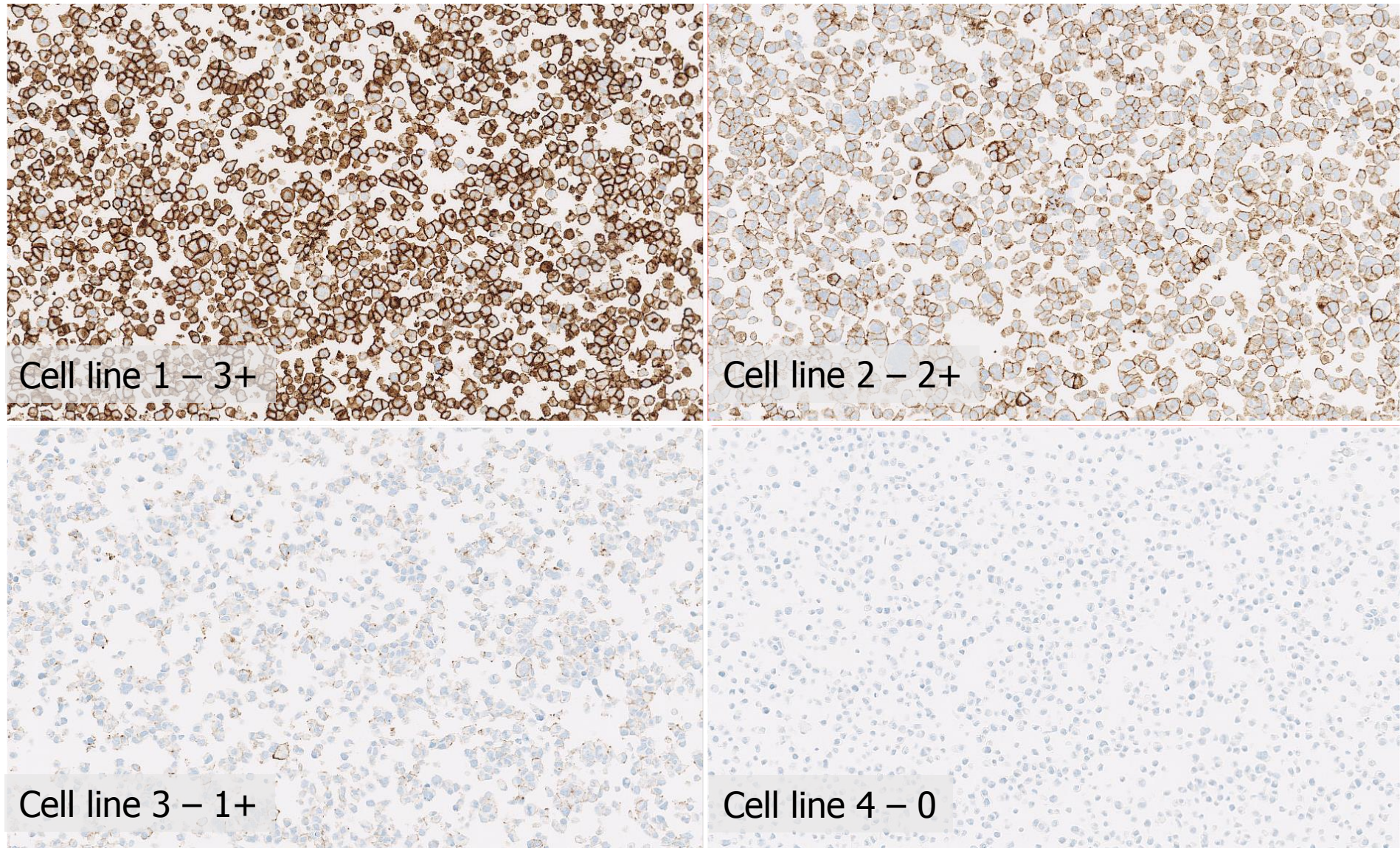
2+

Applicable
for DIA &
ref data
comparing
run-to-run

Courtesy of S. Nielsen

Control material for HER2 IHC: performace control / consistency

Histocyte cell lines HER2: PATHWAY IHC

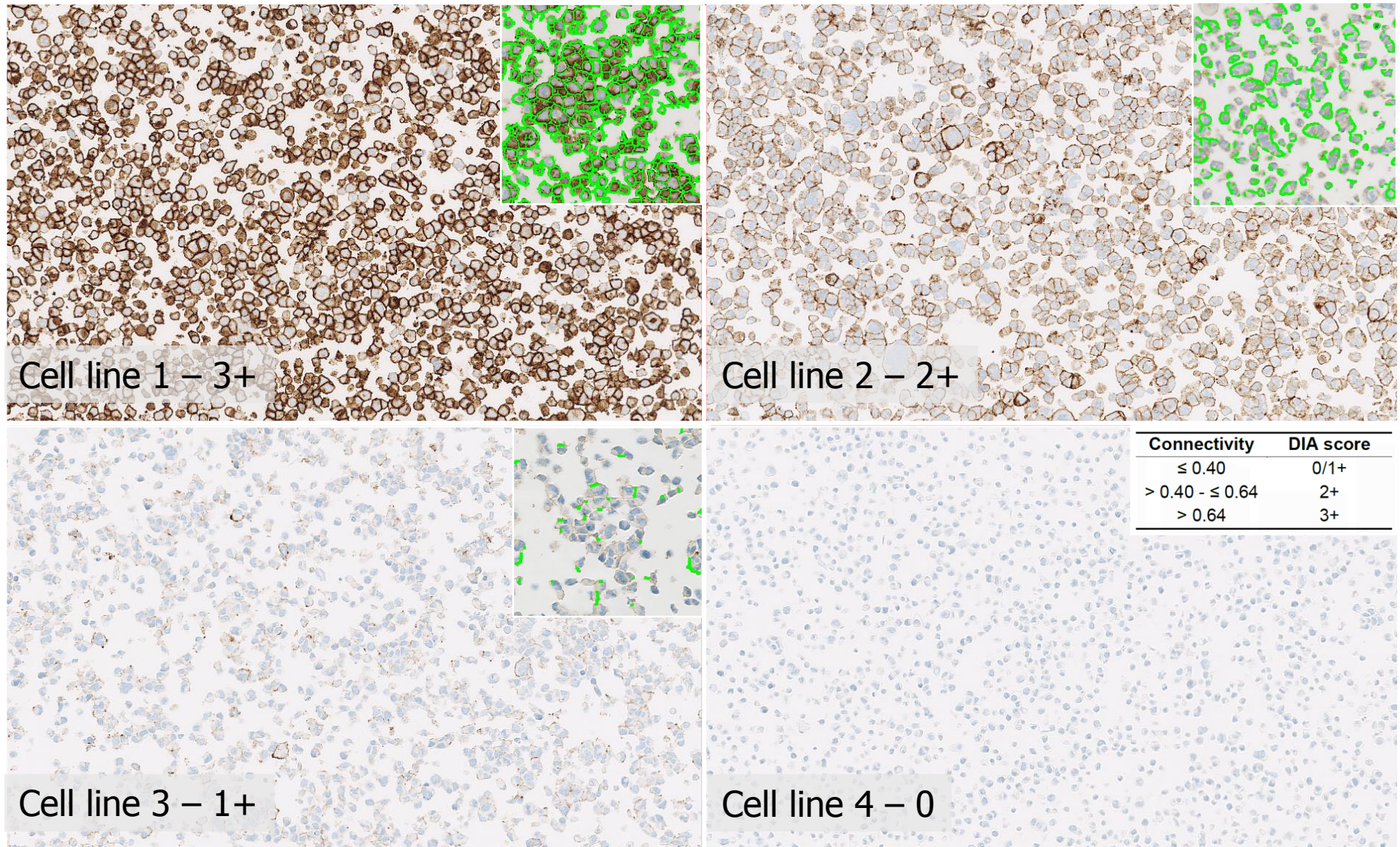


Courtesy of S. Nielsen

Control material for HER2 IHC: performace control / consistency

Histocyte cell lines HER2:

PATHWAY IHC



Courtesy of S. Nielsen

Software

Image analysis in IHC - overview, considerations and applications

Software

- ImageJ (<http://imagej.nih.gov/ij/>): Open-source, FREE, platform-independent, large community, Requires programming-skills
- VIS (<http://www.visiopharm.com/>): fully developed apps, expensive, database-handling of data and images, scanner independent
- Definiens
- INCA
- Aperio (Leica)
- PathXL / Philips
- Matlab

Thank you for your attention!

Collaborators

Søren Nielsen

Rikke Riber-Hansen

Alex Skovsbo Jørgensen

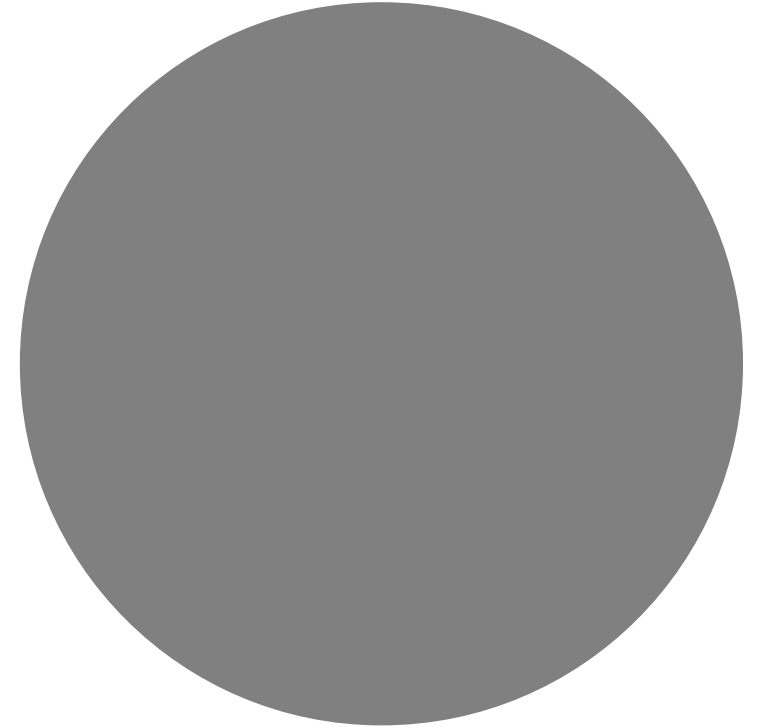
Lasse Riis Østergaard

Mogens Vyberg

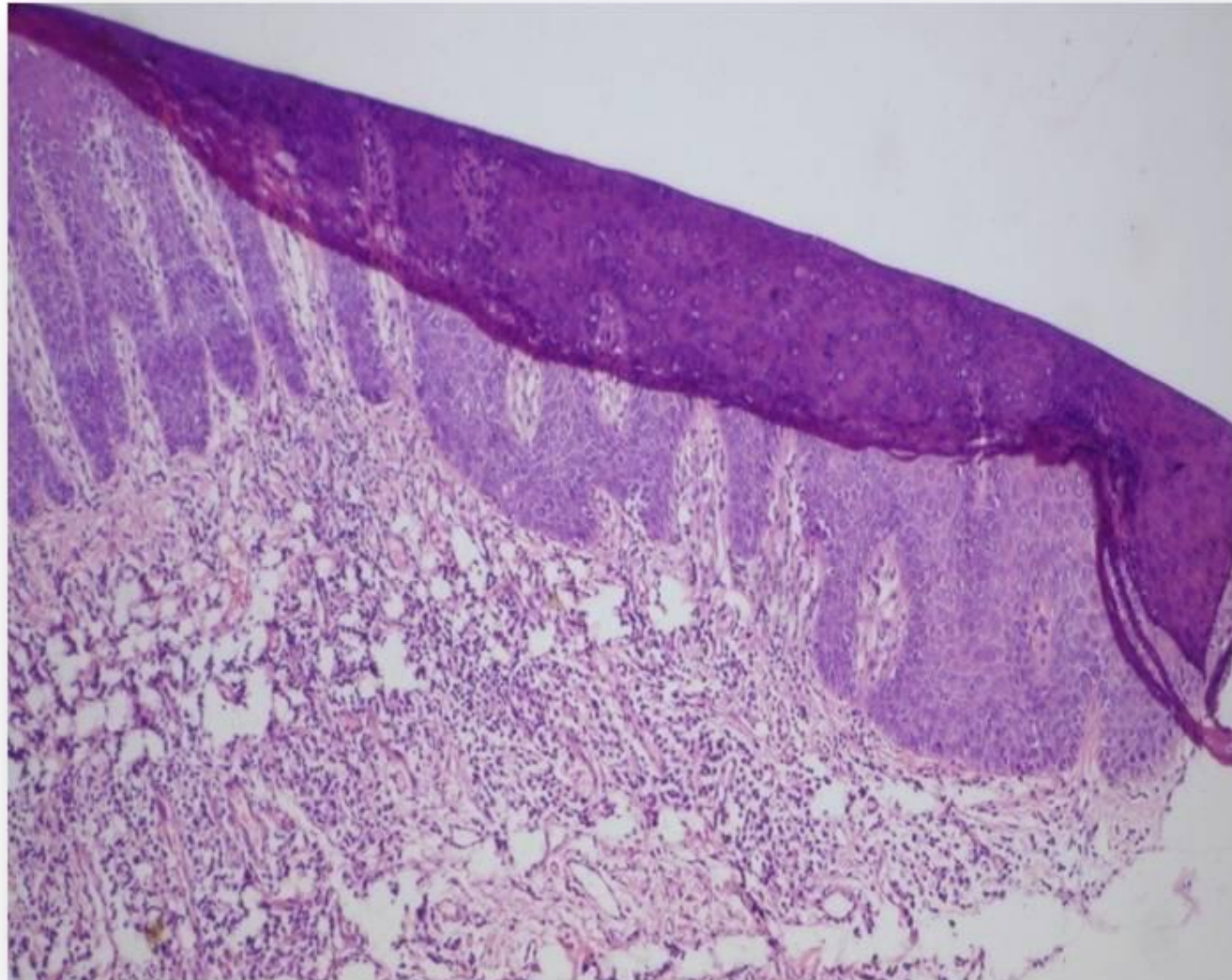


Pitfalls

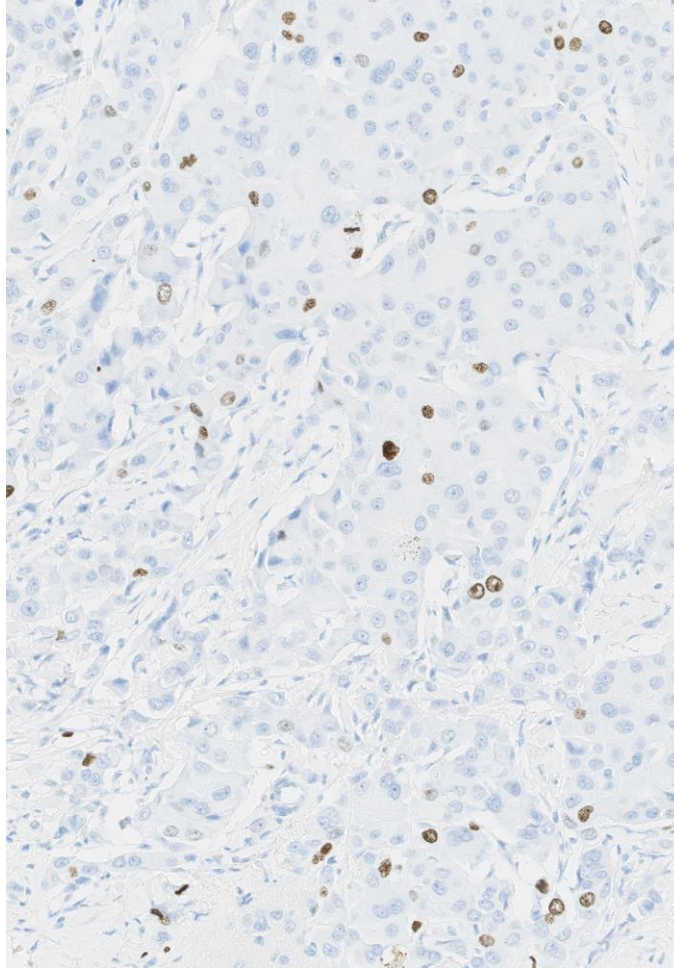
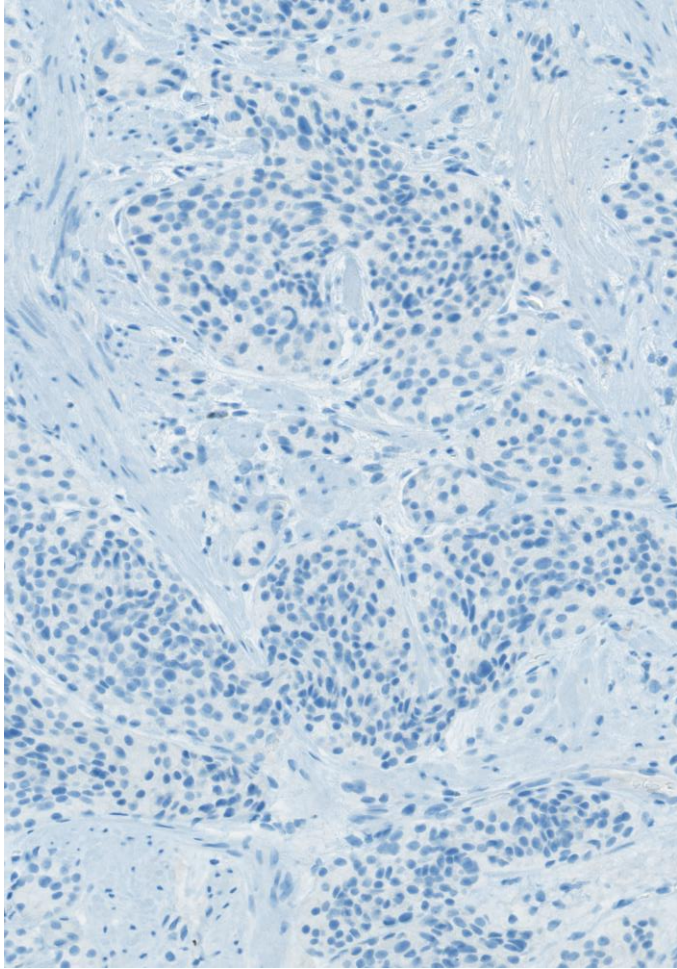
Image analysis in IHC - overview,
considerations and applications



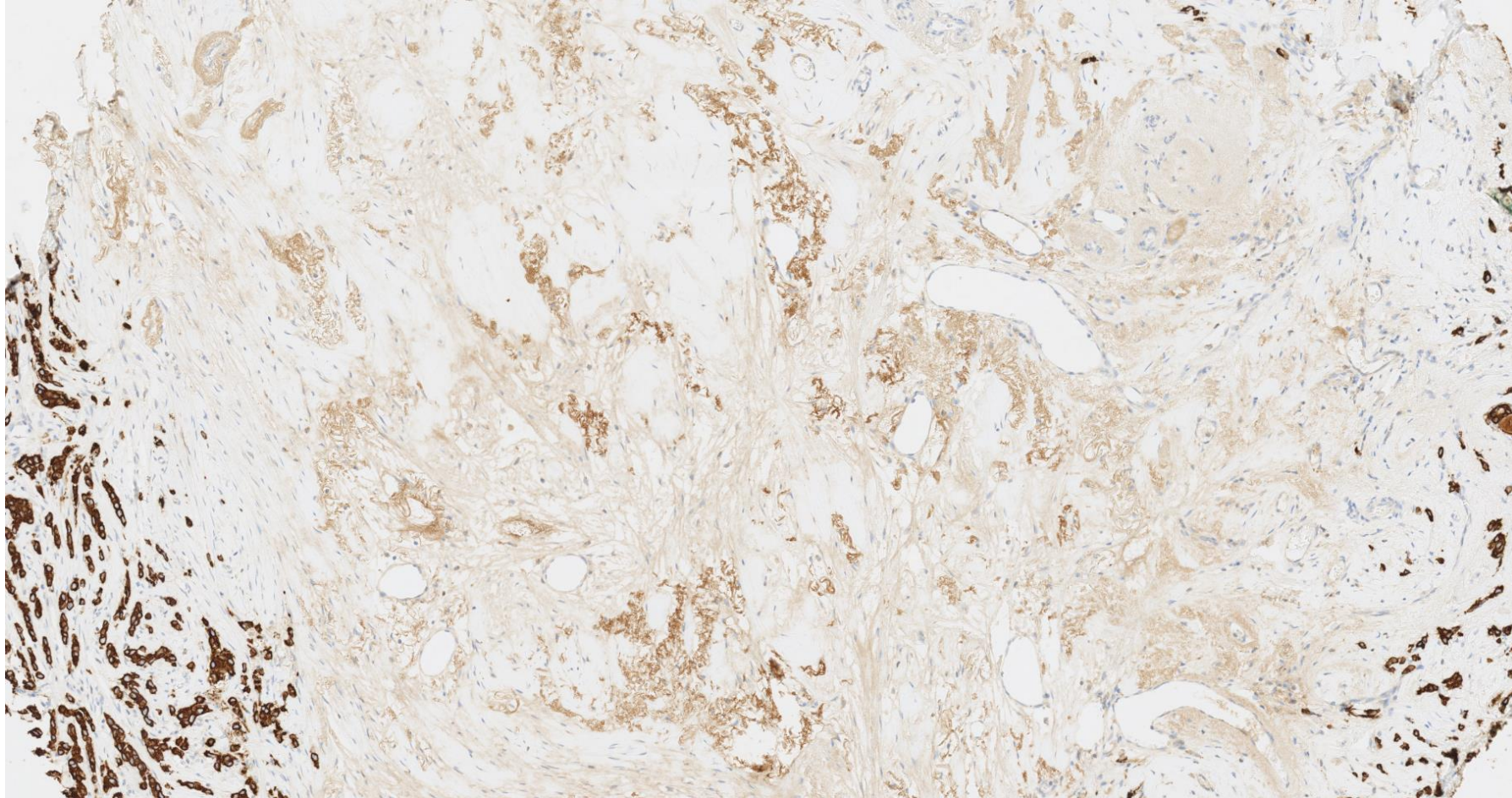
Pitfalls - artefacts



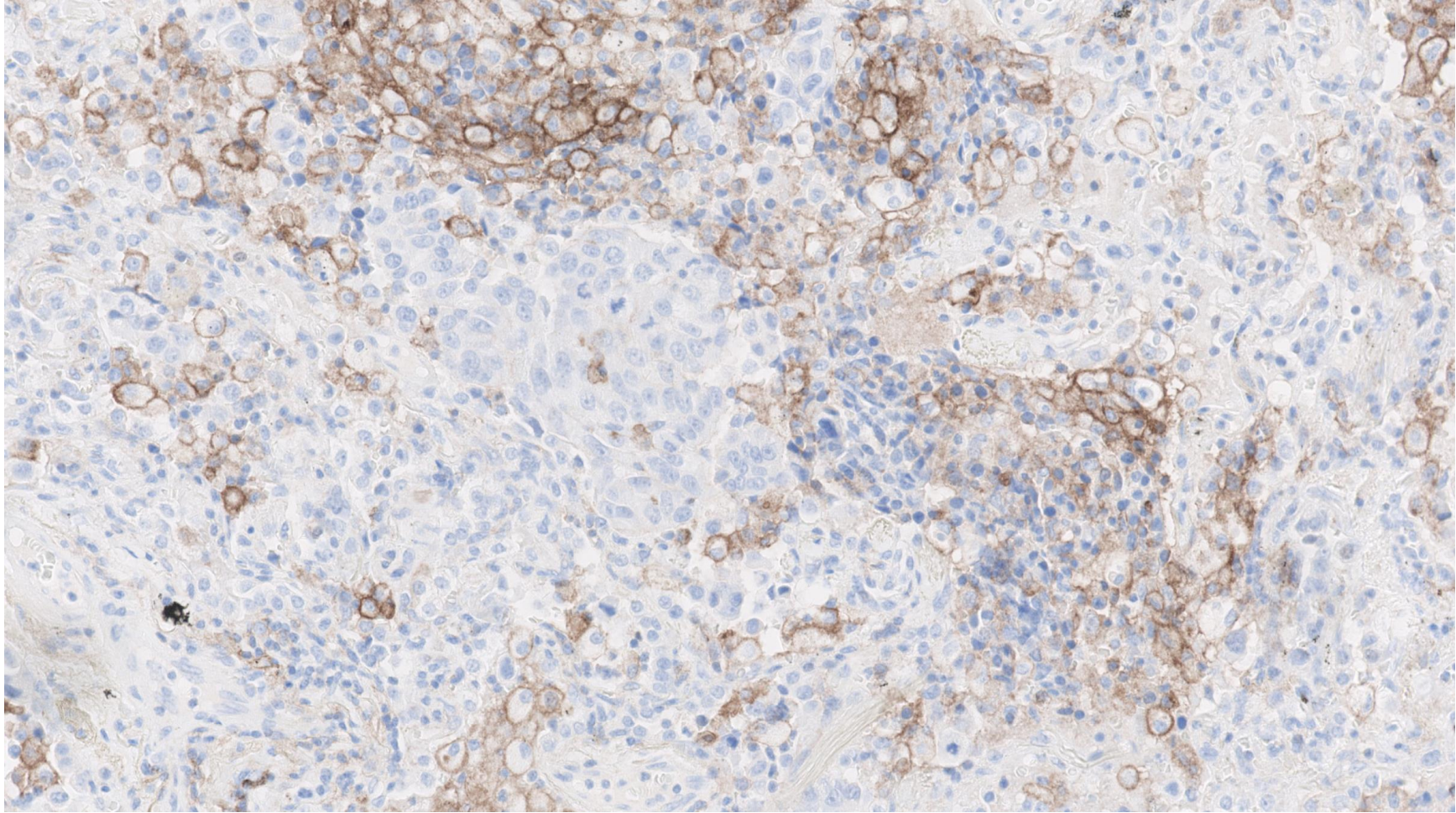
Counter staining



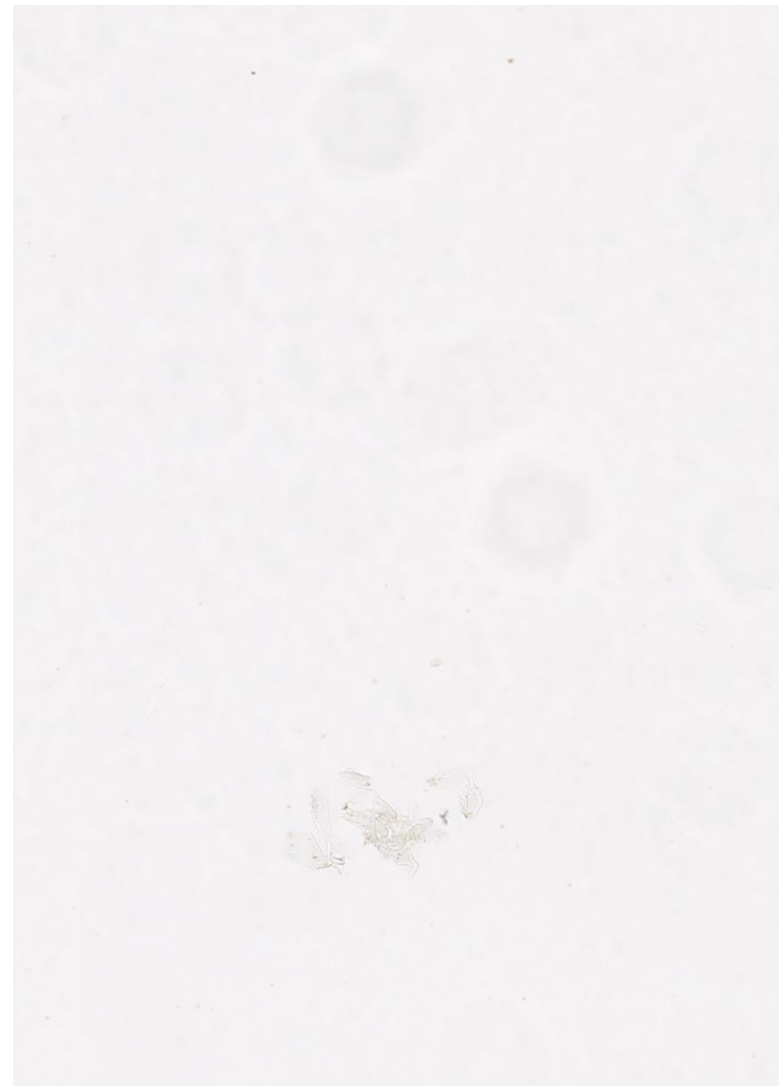
Unspecific / Background staining



Staining of other cells



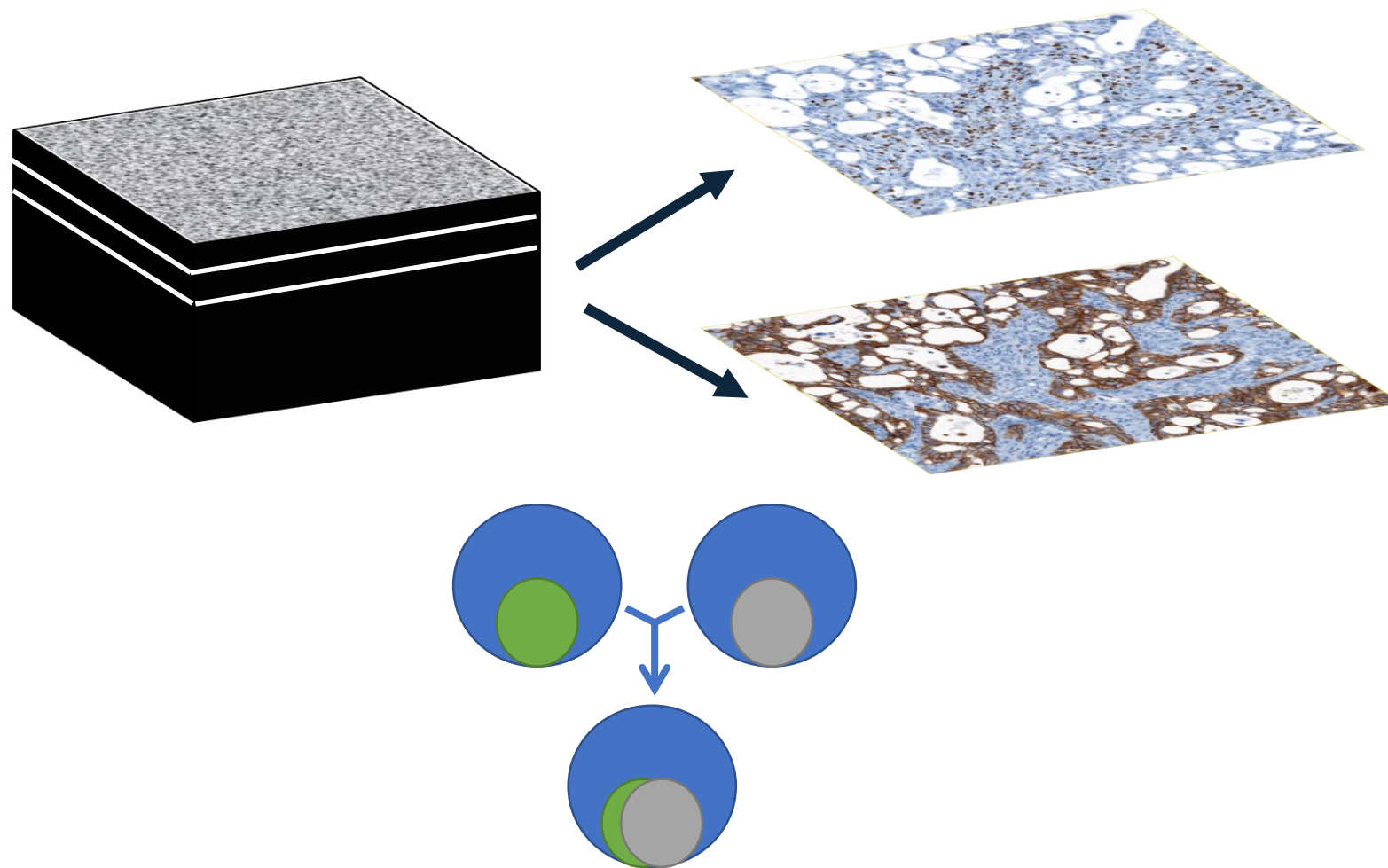
Scanning - background

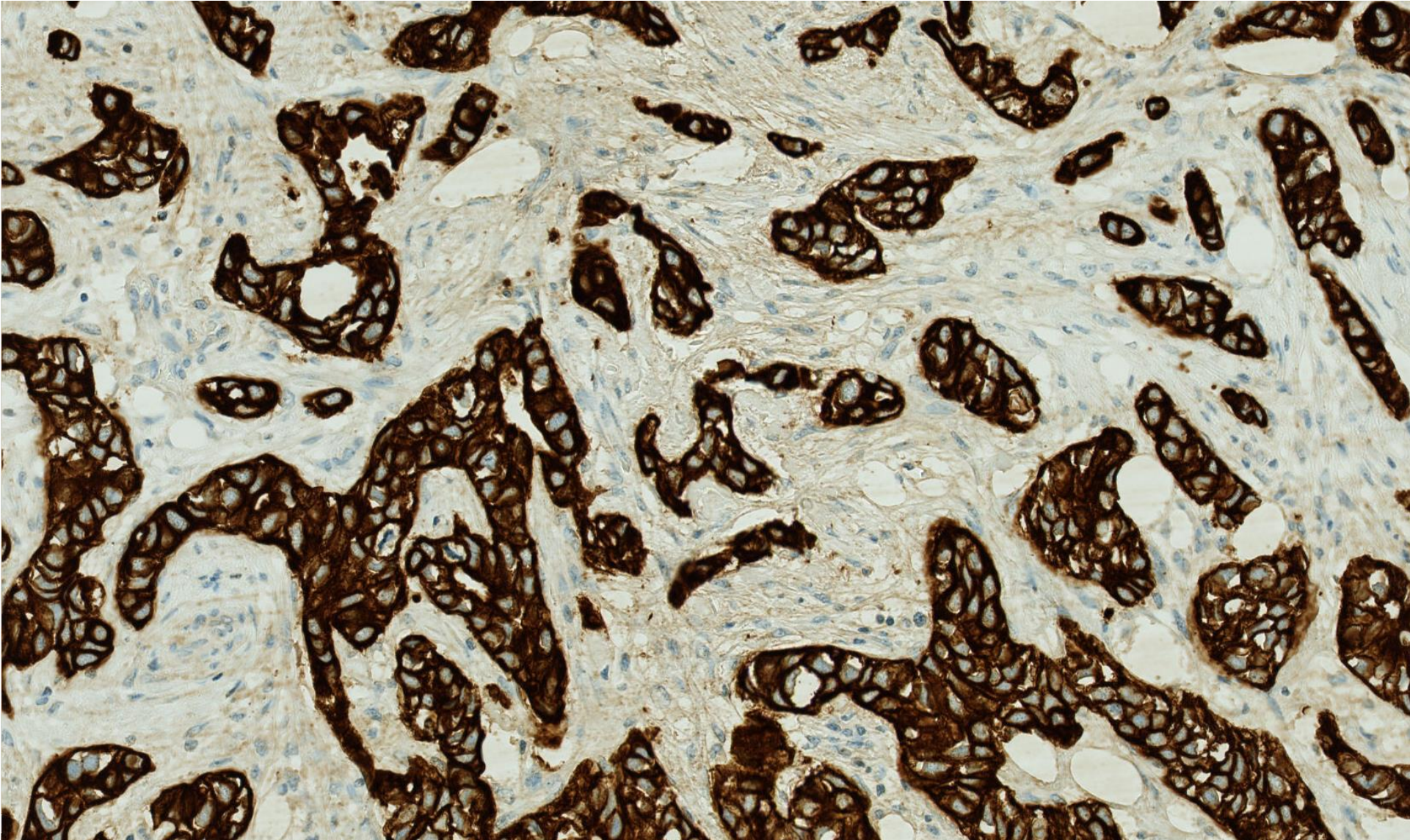


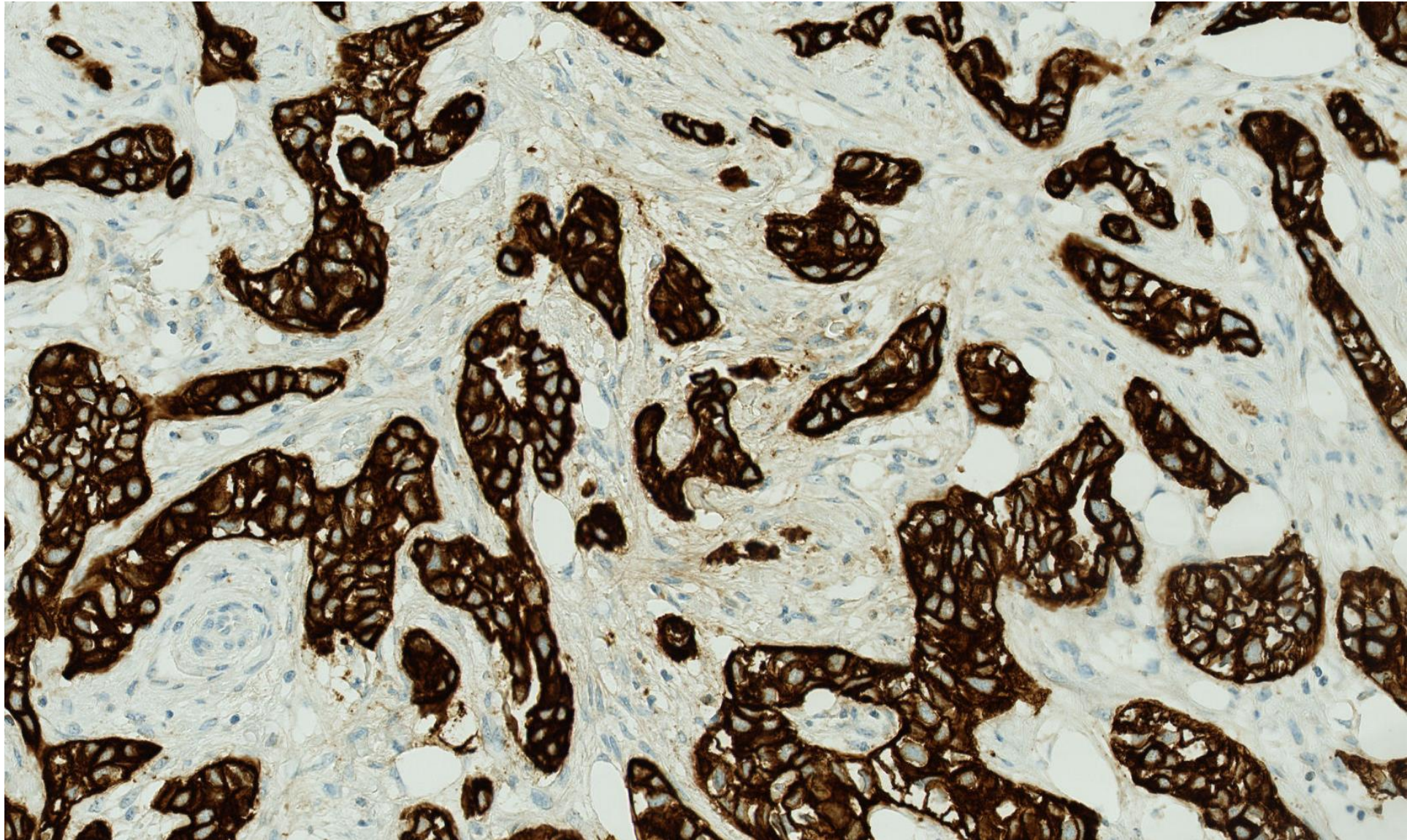
Validation of alignment

Digital Image Analysis – Ki67

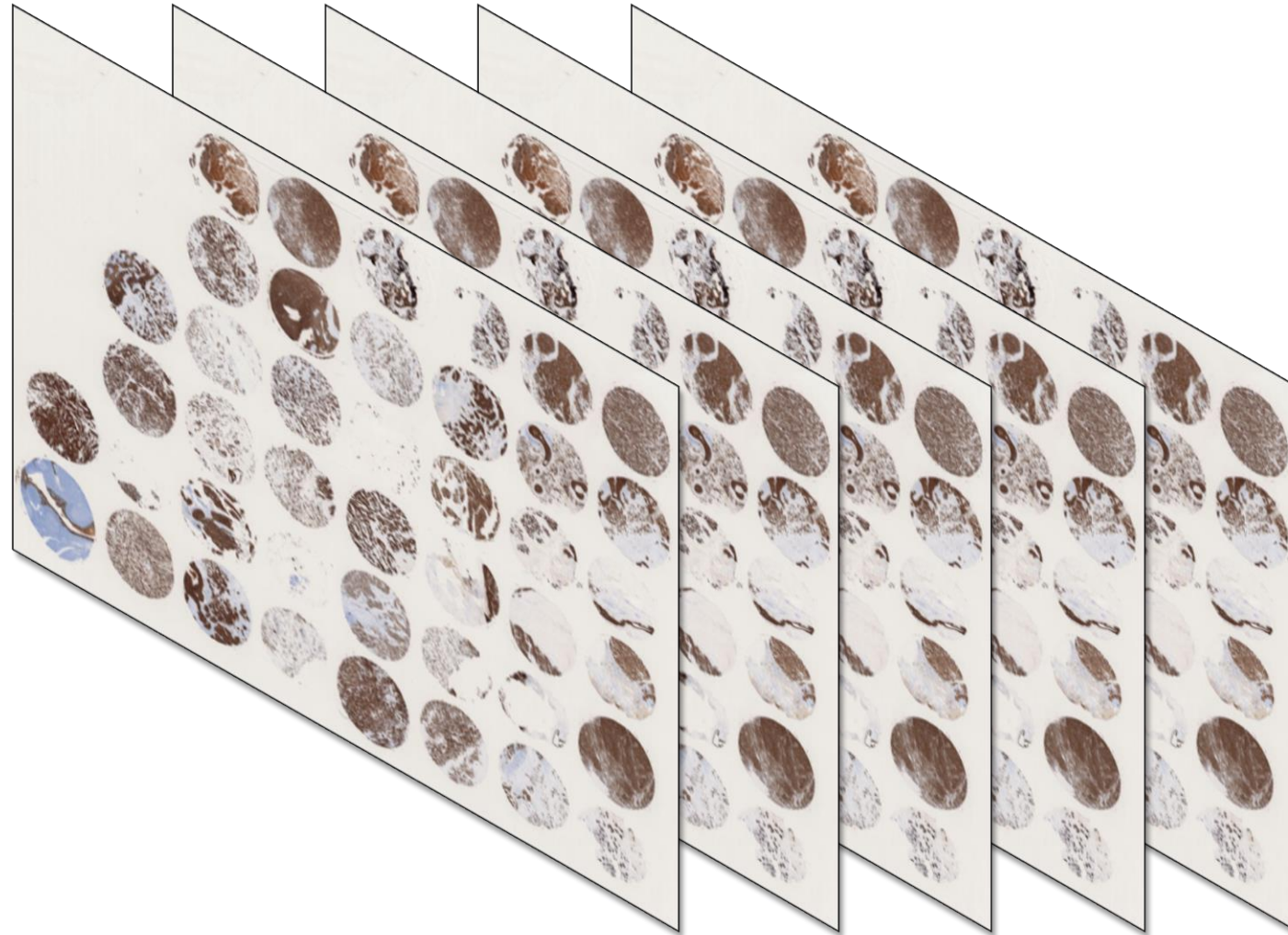
Validation of alignment







Five parallel slides of PCK

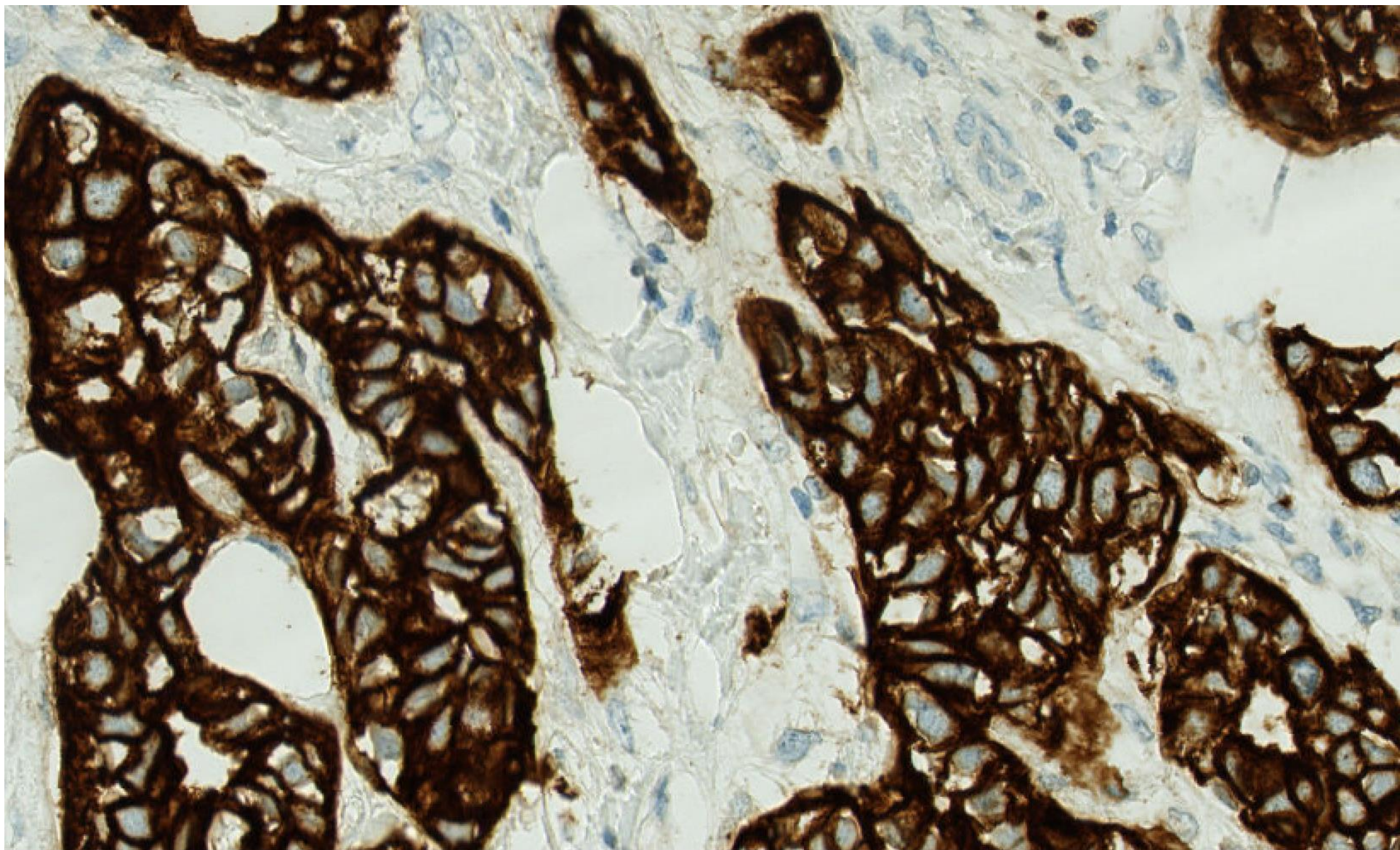


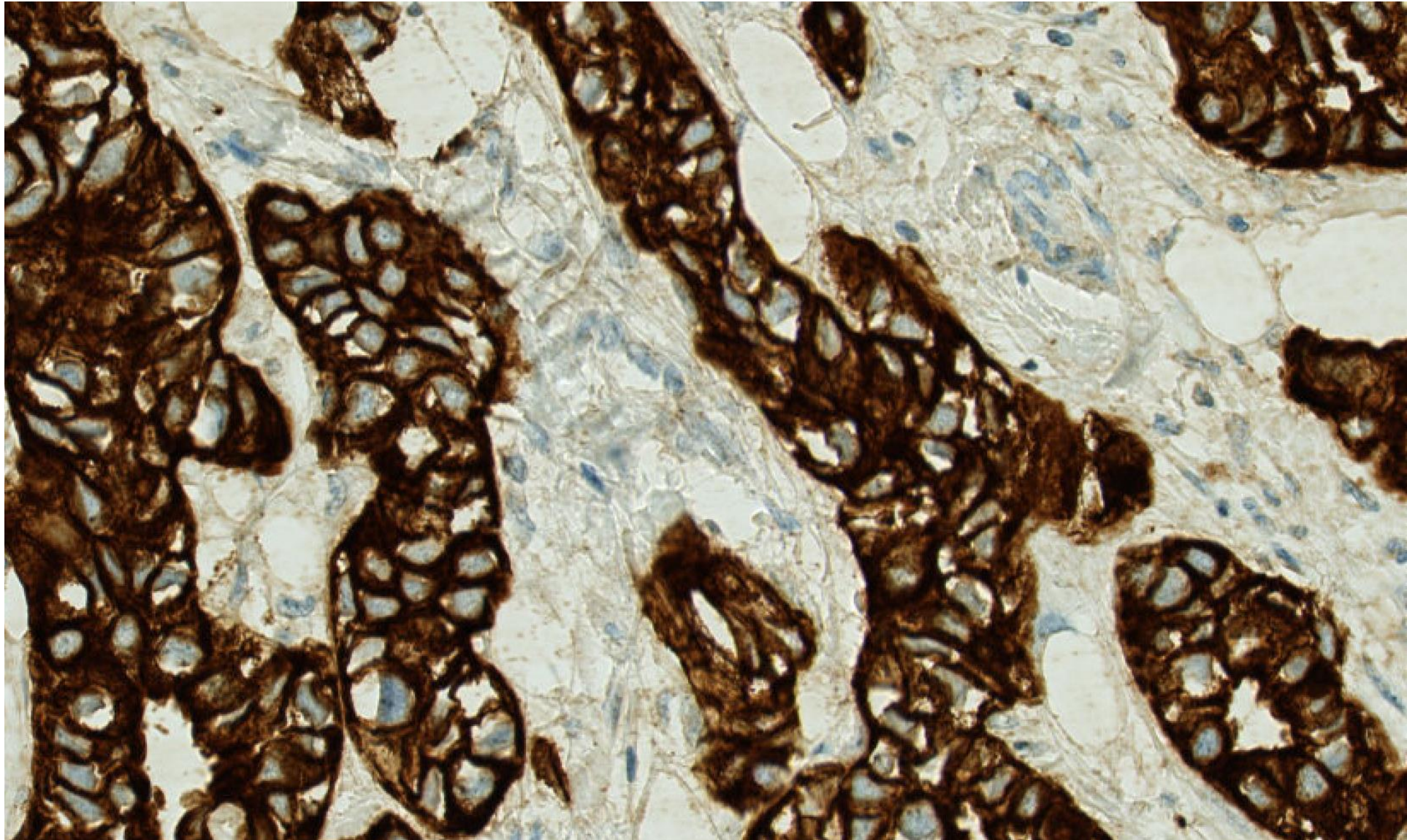
PCK-Alignment

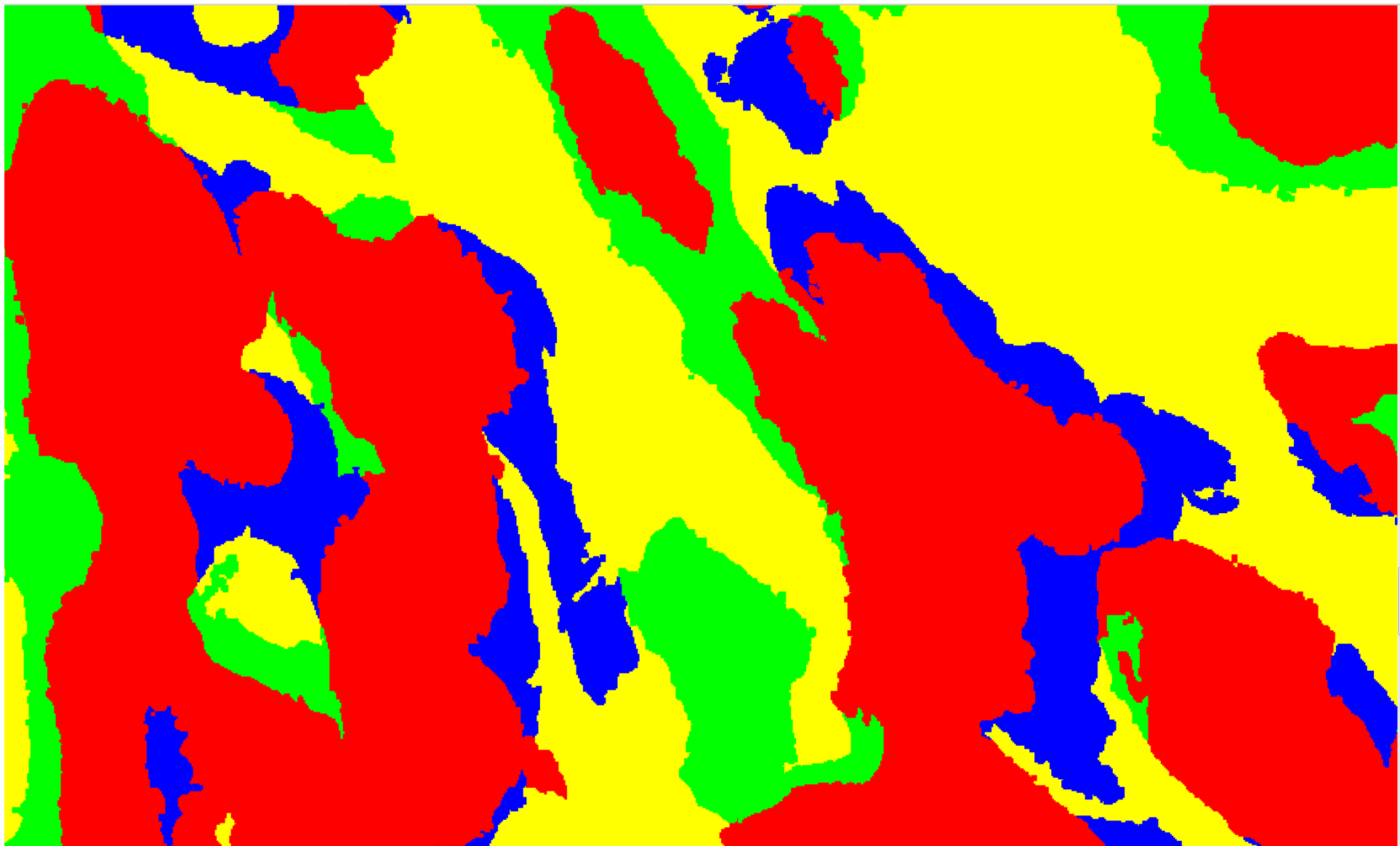
- 5 parallel slides from TMA containing 40 breast cancers
- All stained for PCK TMA
- Only 26 (of 40) cores were usable
- Exclusion were due to
 - Missing cores in one or more slides
 - Damaged cores

PCK-Alignment

- Algorithm was developed that segmented 2 slides based on PCK expression
- Four categories based on PCK status in slide 1 and slide 2:
 - + / + : PCK positive in both slides
 - / - : PCK negative in both slides
 - + / - or - / +: PCK positive in only one slide





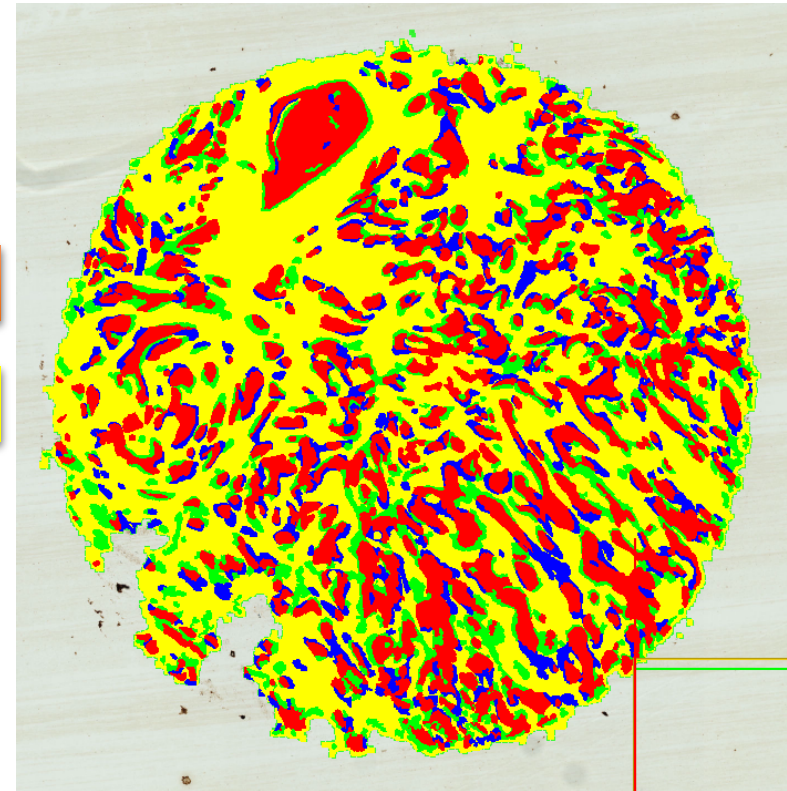


Overlap/agreement (%)

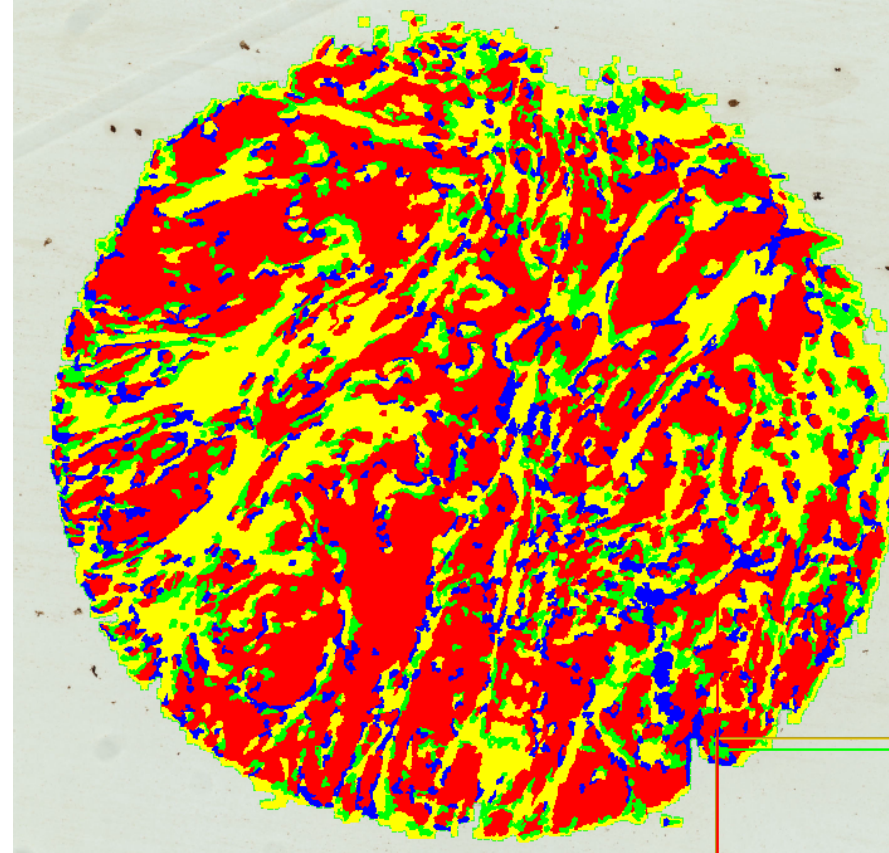
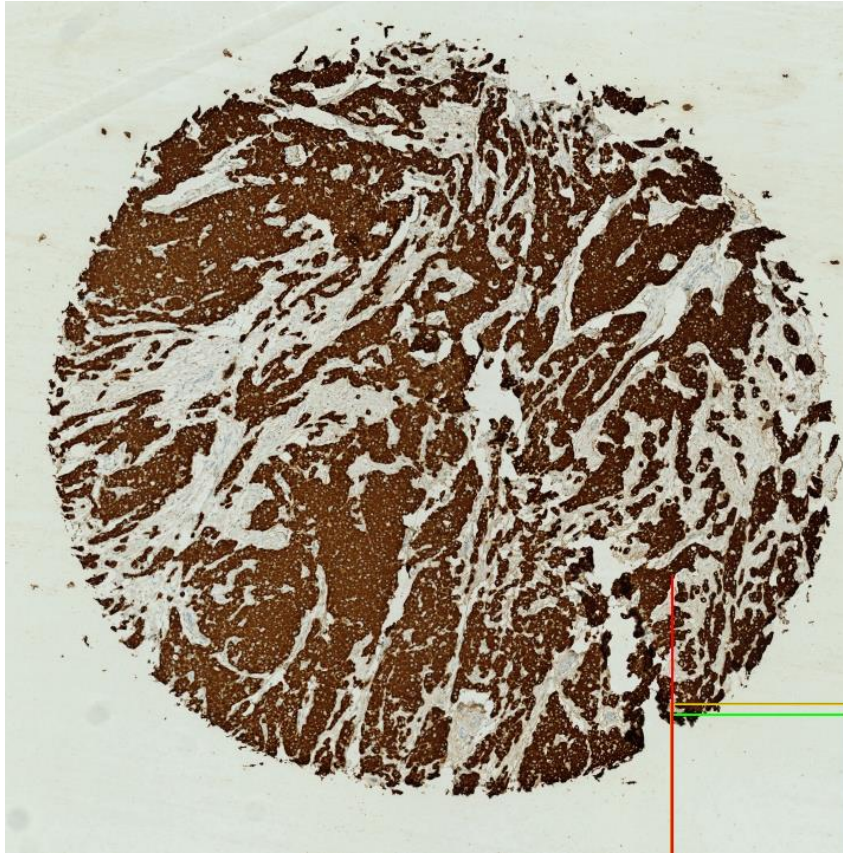
- Calculated as:

PCK positive area in both slides +
PCK negative area in both slides

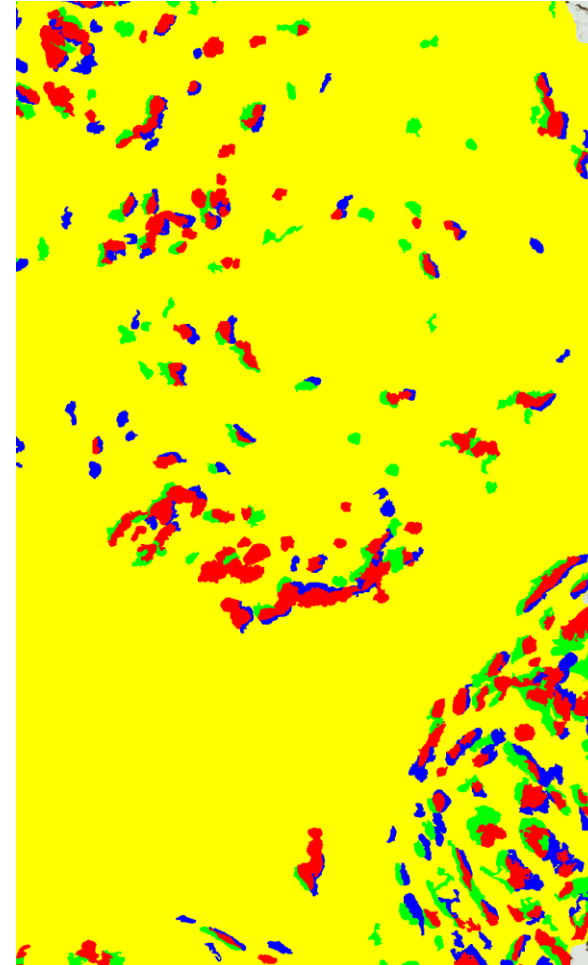
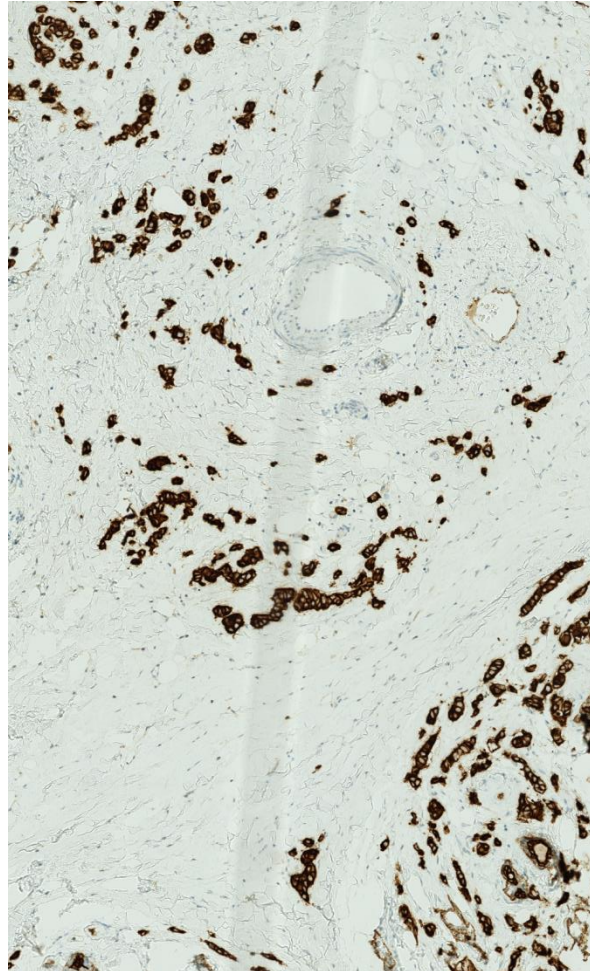
Divided by total area



Good agreement (>90 %)



Less good agreement



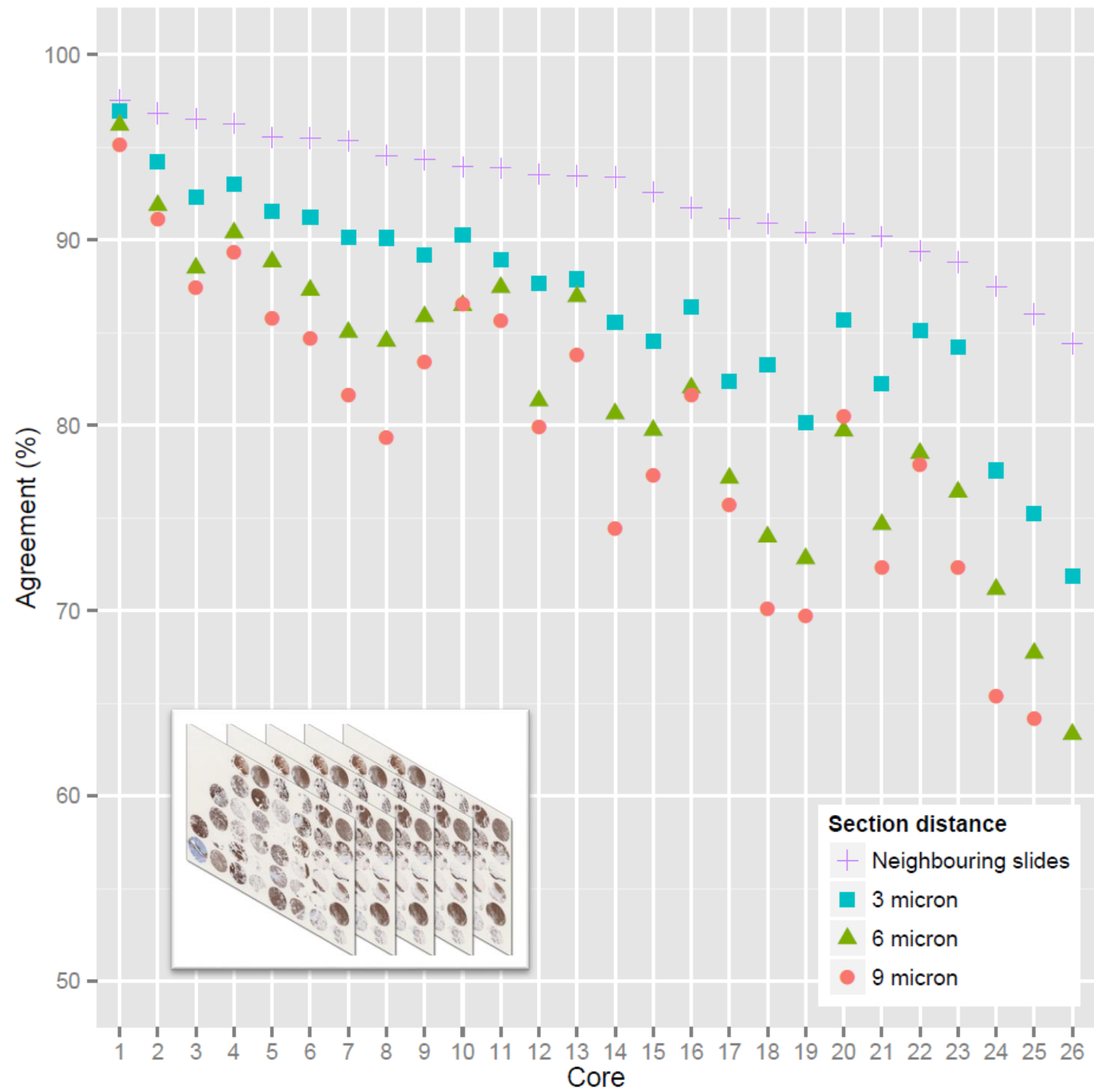


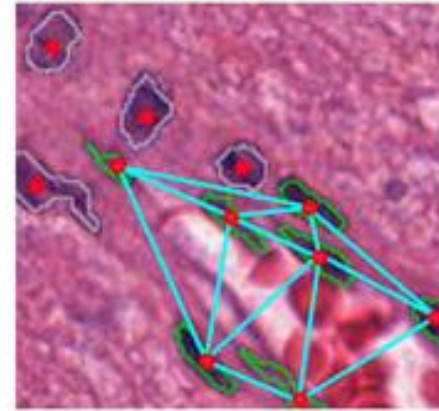
Image analysis – advanced algorithms

Image analysis in IHC - overview, considerations
and applications

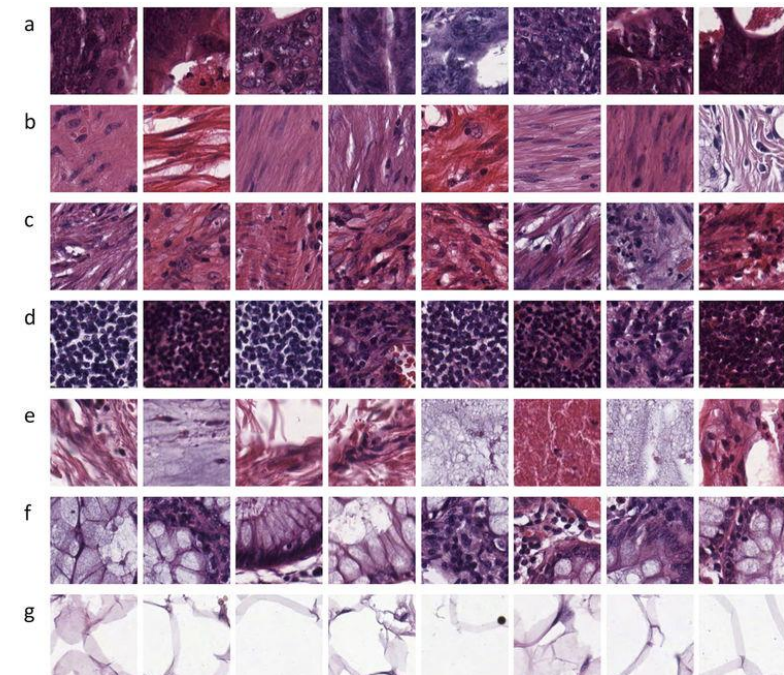


Advanced algorithms

- More complex algorithms
- Successive application of several algorithms
- Not only thresholds
- Texture-based
- Architecture-based
- Feature-based training
 - Feature may be selected statistically and unsupervised



Quantitative phenotyping



Advanced algorithms – architectural and texture

AUTOMATED GRADING OF PROSTATE CANCER USING ARCHITECTURAL AND TEXTURAL IMAGE FEATURES

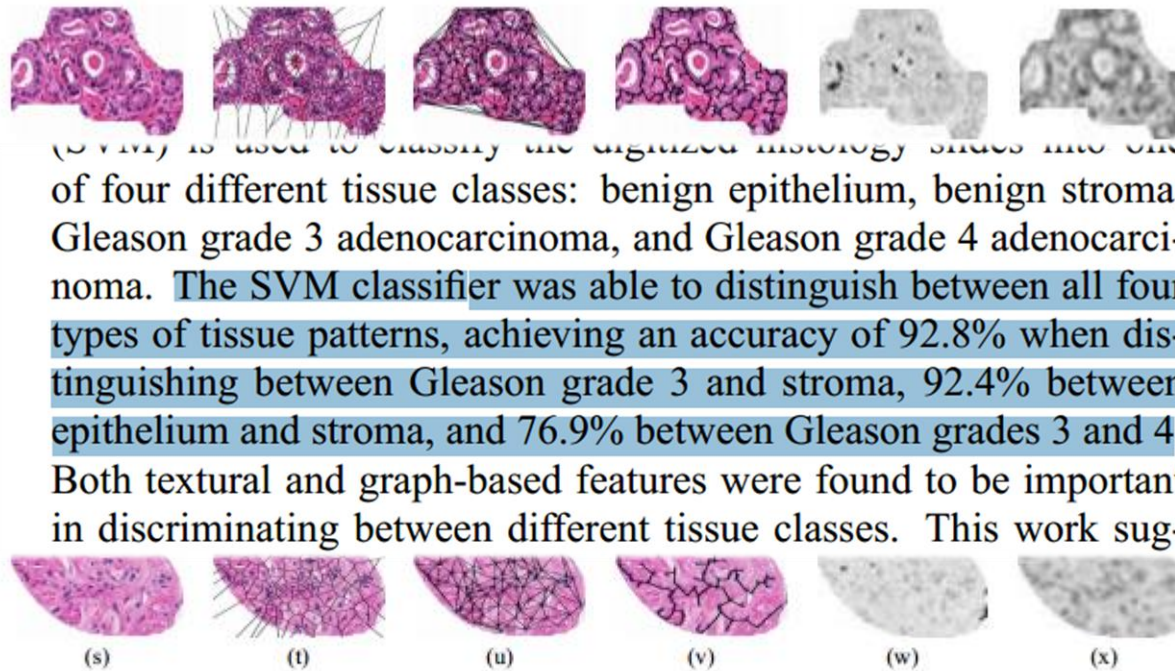
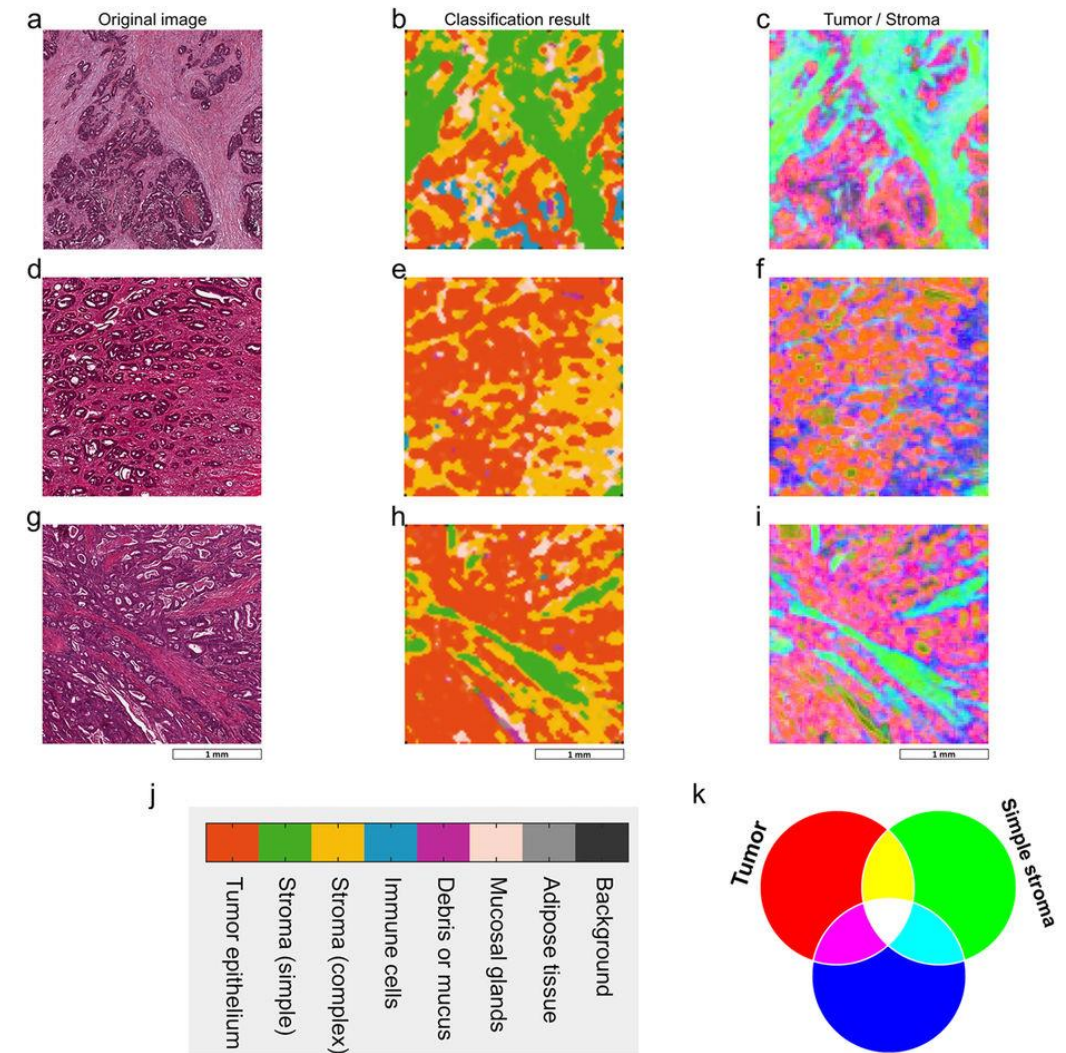
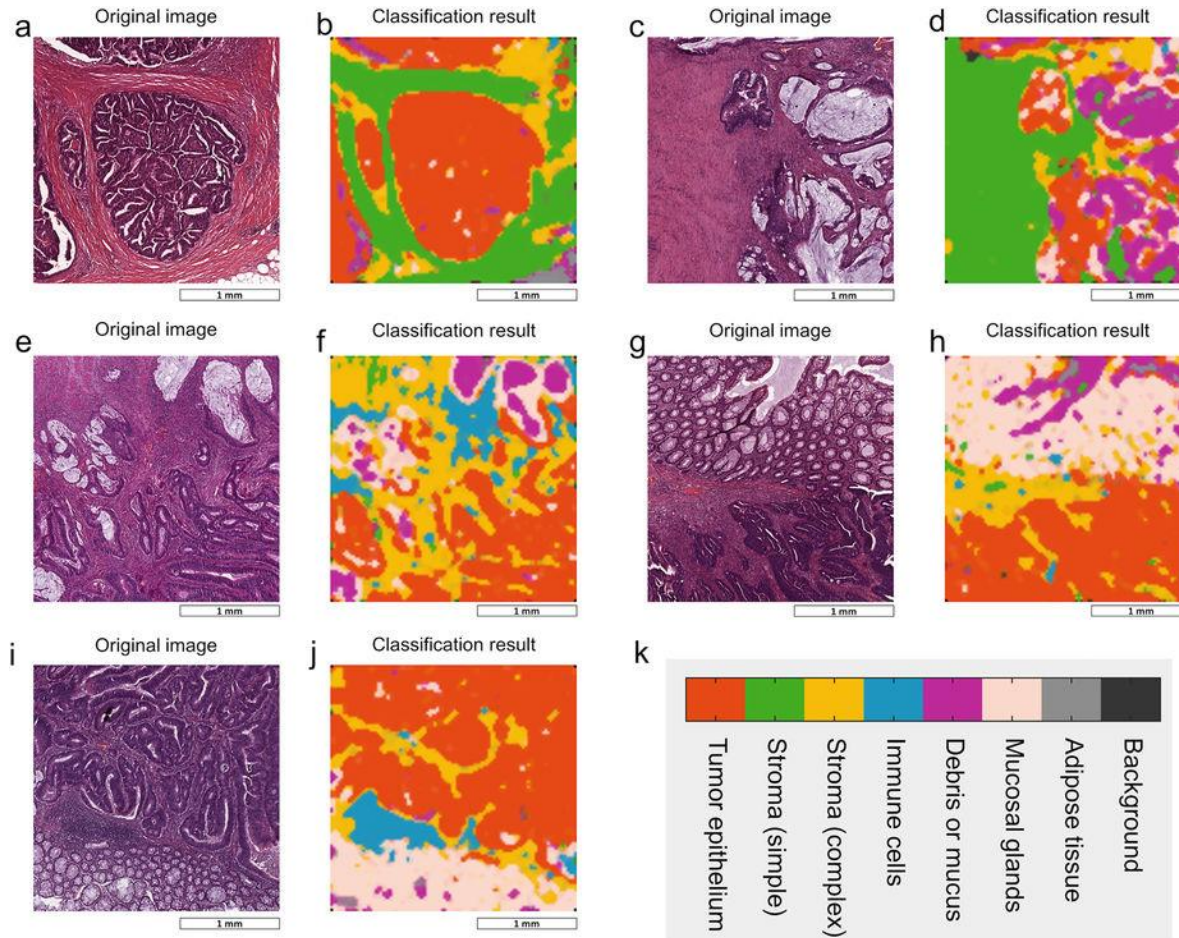


Fig. 3. Comparison of ((a)-(f)) Gleason grade 3 tissue, ((g)-(l)) grade 4 tissue, ((m)-(r)) benign epithelium, and ((s)-(x)) benign stroma. Superimposed on ((a), (g), (m), (s)) the original images are ((b), (h), (n), (t)) the Voronoi diagram, ((c), (i), (o), (u)) the Delaunay triangulation, ((d), (j), (p), (v)) the minimum spanning trees, ((e), (k), (q), (w)) pixel entropy texture feature, and ((f), (l), (r), (x)) Gabor filter ($s = 3$, $\theta = \frac{5\pi}{8}$).

Advanced algorithms - texture



Kather, J. N., Weis, C. A., Bianconi, F., Melchers, S. M., Schad, L. R., Gaiser, T., ... & Zöllner, F. G. (2016). Multi-class texture analysis in colorectal cancer histology. *Scientific reports*, 6, 27988.

Advanced algorithms – cell nuclei texture

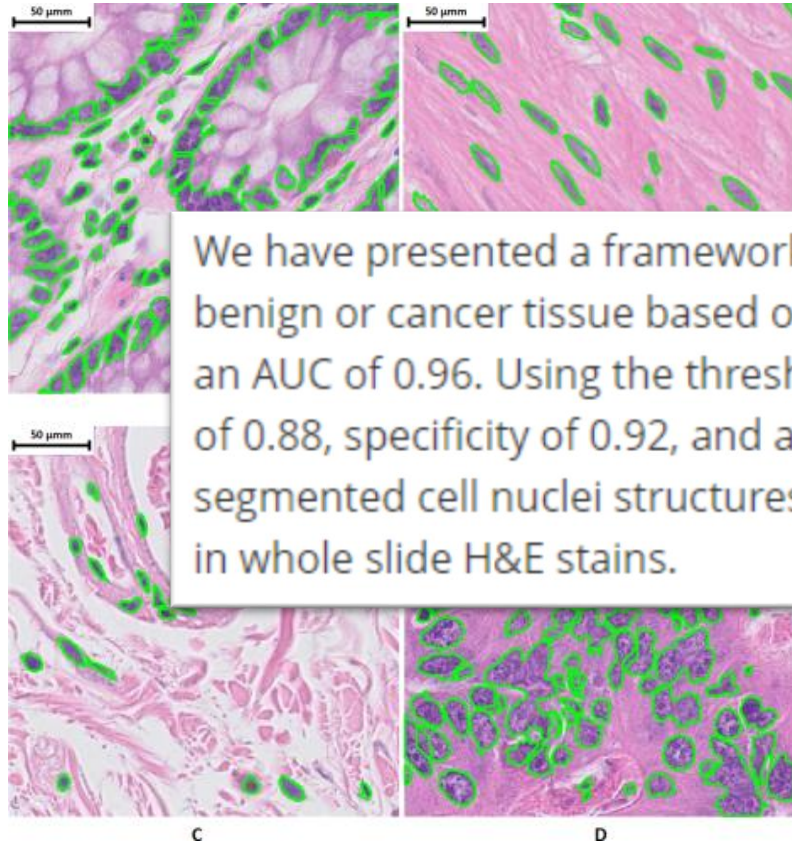
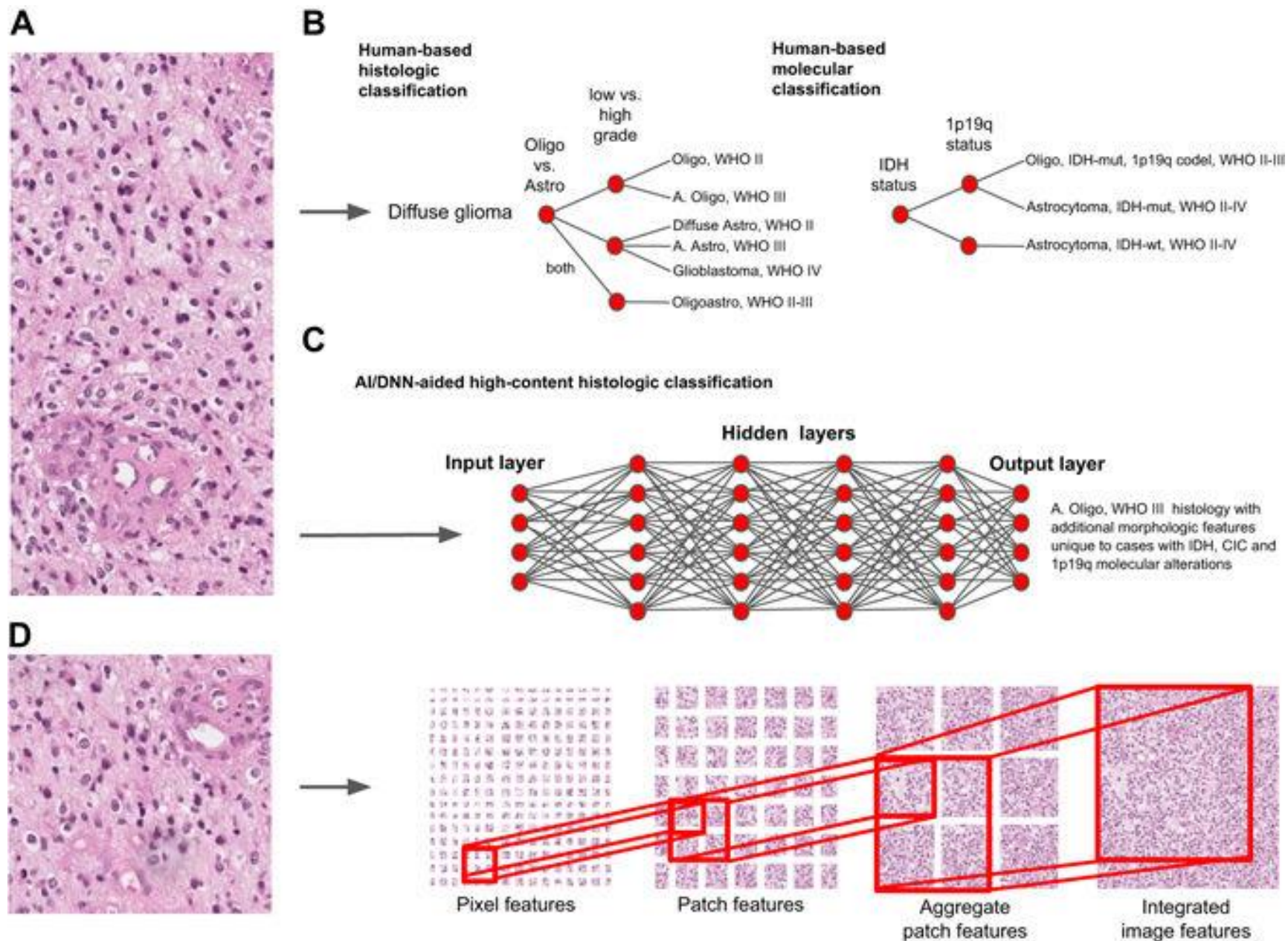


Table 1. The nine most important features selected for the ROI classifier

Feature number	Feature name	Feature image	Variable (Gini) importance
1	Maximum intensity	Saturation	16.27
7	Mean long run emphasis	Blue	6.85
8	Maximum intensity	Hue	6.25
9	Long run low gray level emphasis	Blue	5.82

We have presented a framework for classification of ROI's within whole slide H&E stains as containing benign or cancer tissue based on features extracted from segmented cell nuclei structures obtaining an AUC of 0.96. Using the threshold maximizing sensitivity and specificity simultaneously, a sensitivity of 0.88, specificity of 0.92, and an accuracy of 0.91 was obtained. The results indicate that features from segmented cell nuclei structures can be used to discriminate between benign or cancerous colon tissue in whole slide H&E stains.

AI



Djuric, Ugljesa, et al. "Precision histology: how deep learning is poised to revitalize histomorphology for personalized cancer care." *npj Precision Oncology* 1.1 (2017): 22.