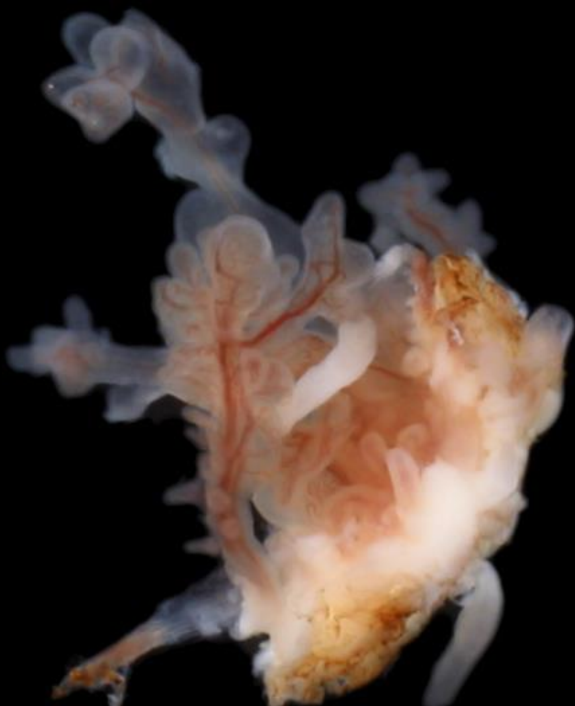
Thomas Fuchs (JPL, Caltech)

Application: Cancer Research

Sponsored by the Jet Propulsion Laboratory and by the Keck Institute for Space Studies, California Institute of Technology



Bladder

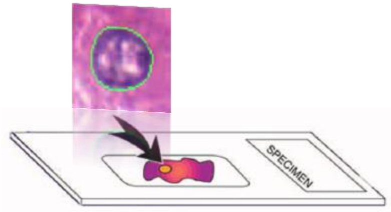
Kidney

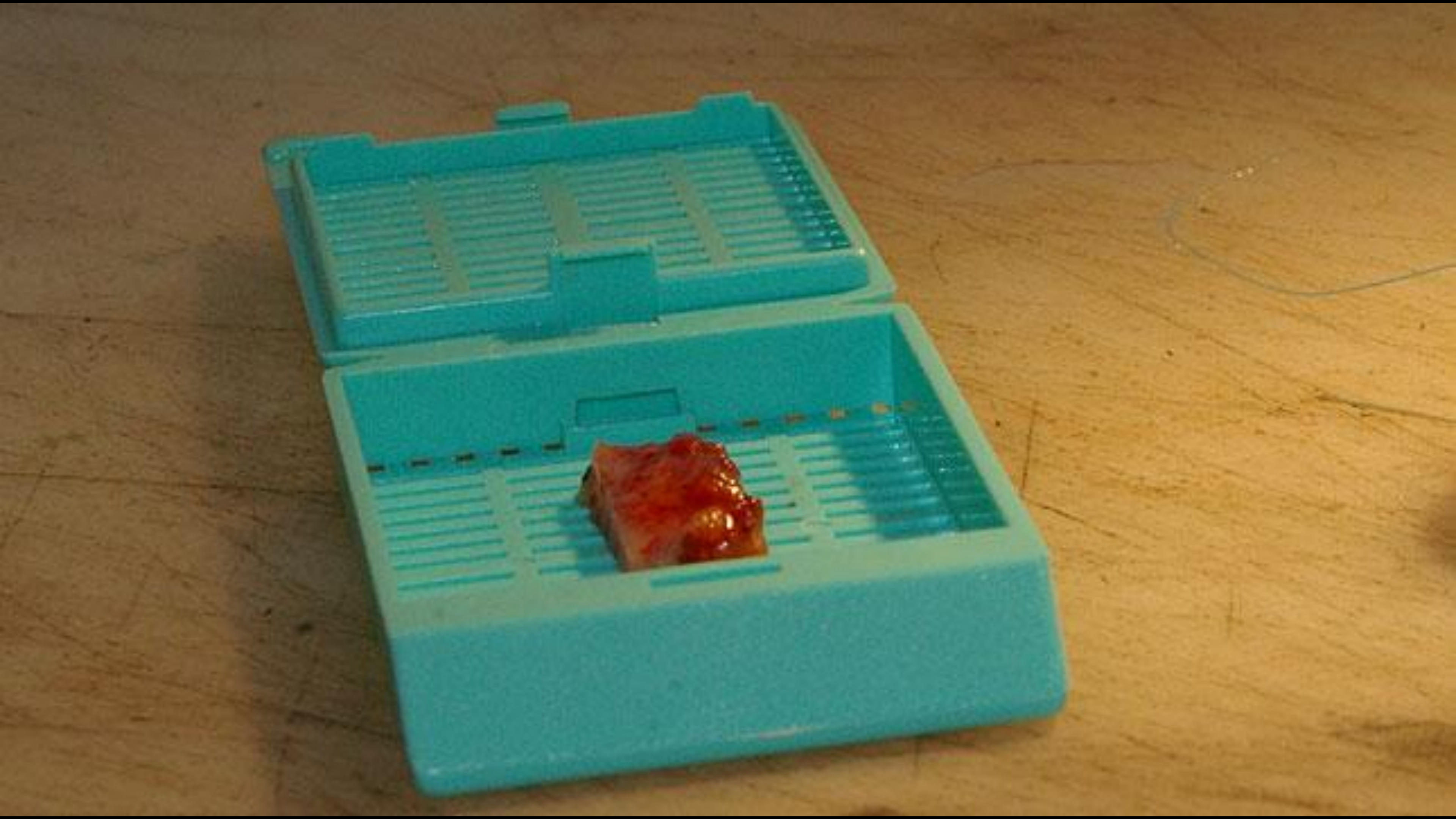
Lymphnodes

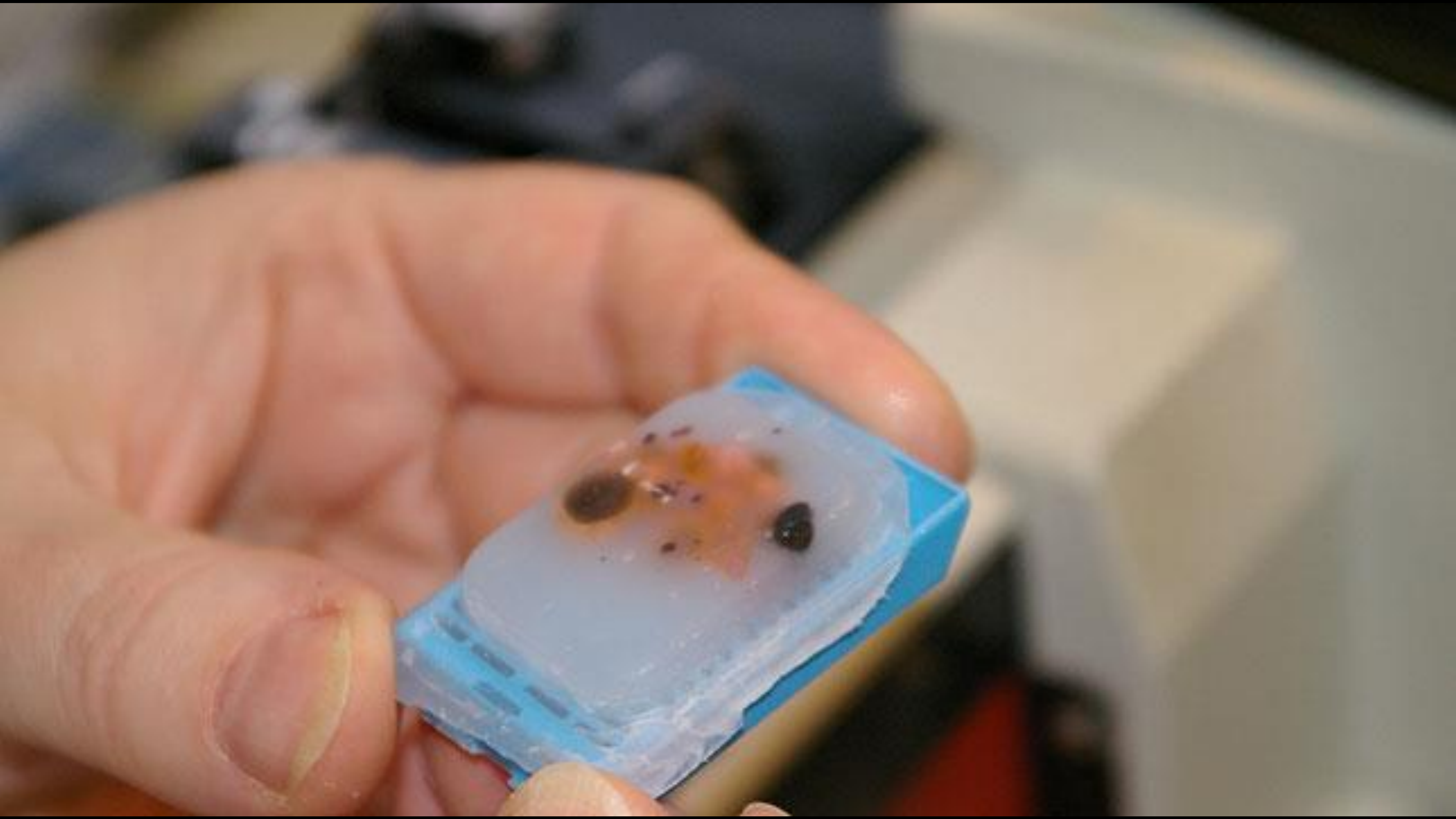
Definition

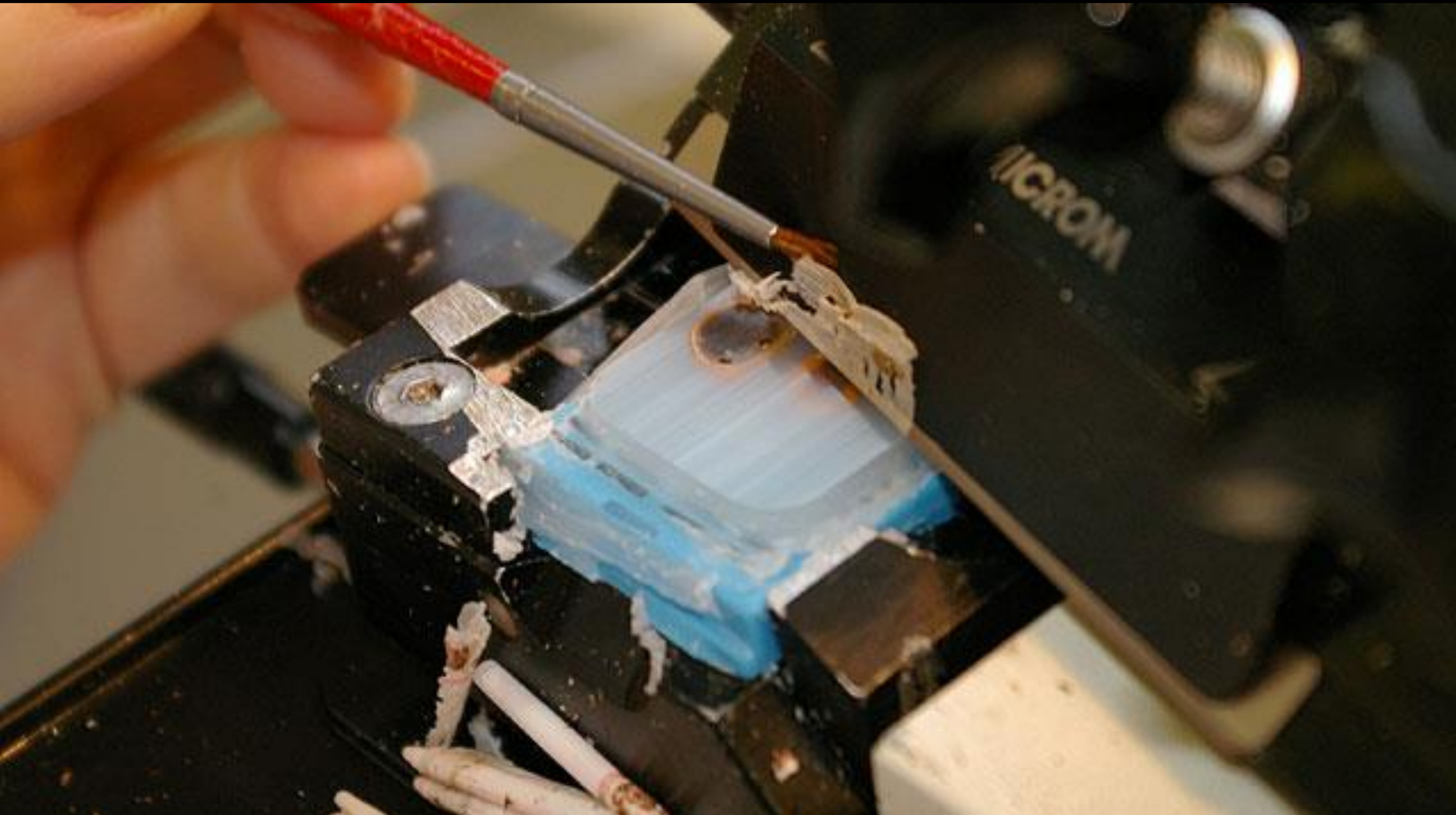
Computational Pathology investigates a **complete probabilistic treatment** of scientific and clinical workflows in general pathology, i.e. it combines experimental design, statistical pattern recognition and survival analysis within an **unified framework** to answer scientific and clinical questions in pathology.

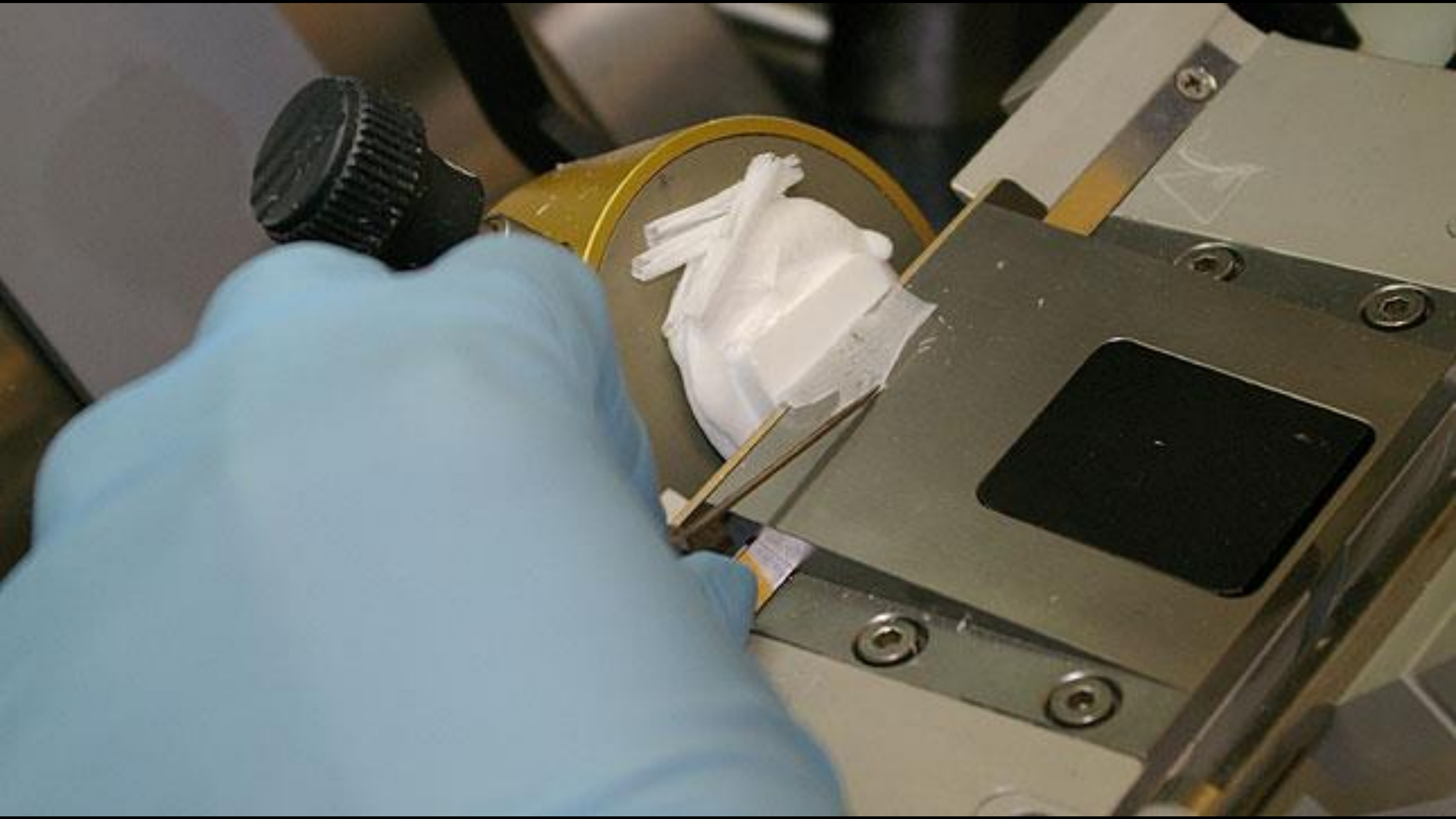
[CMIG 2011]

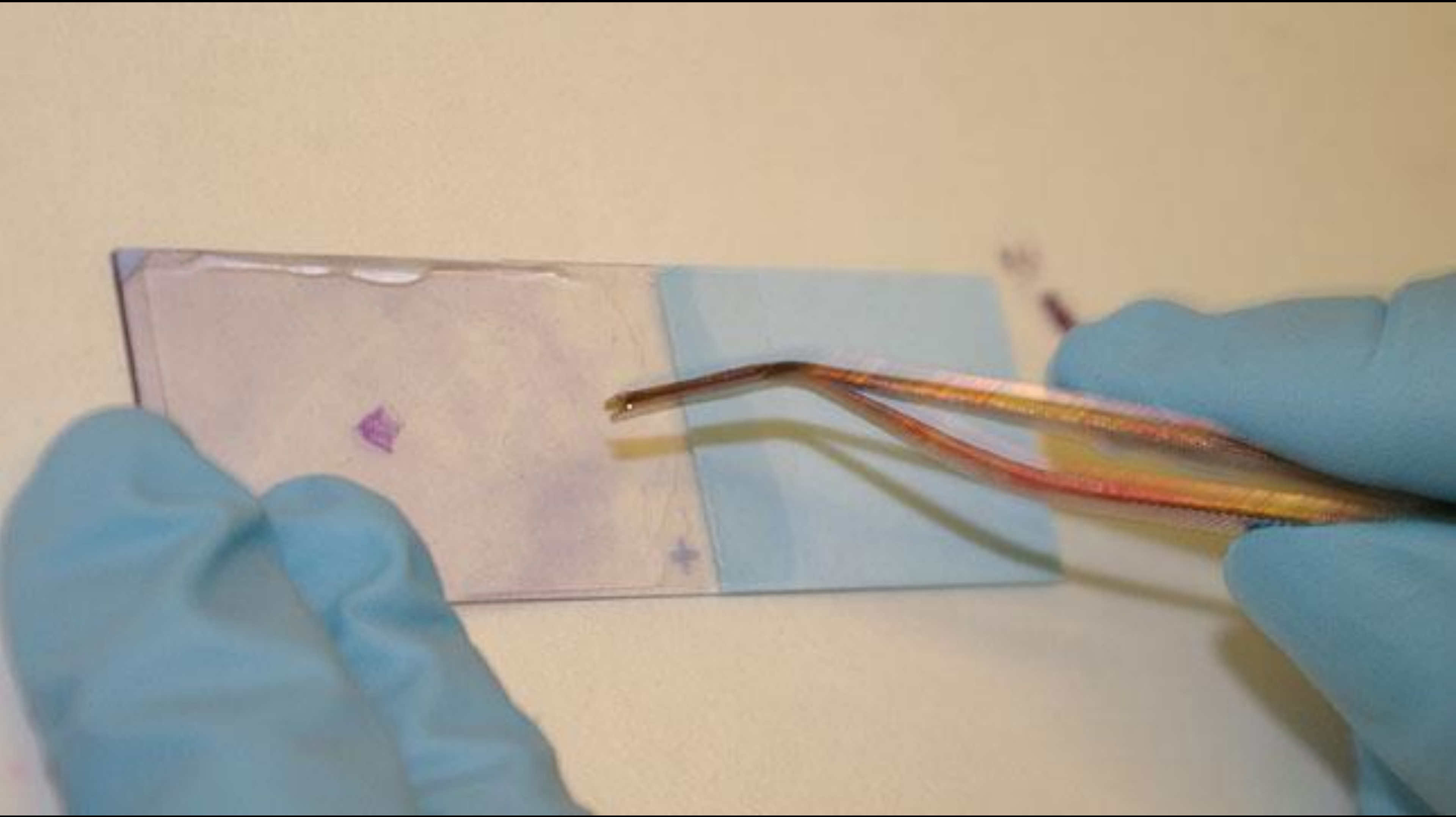


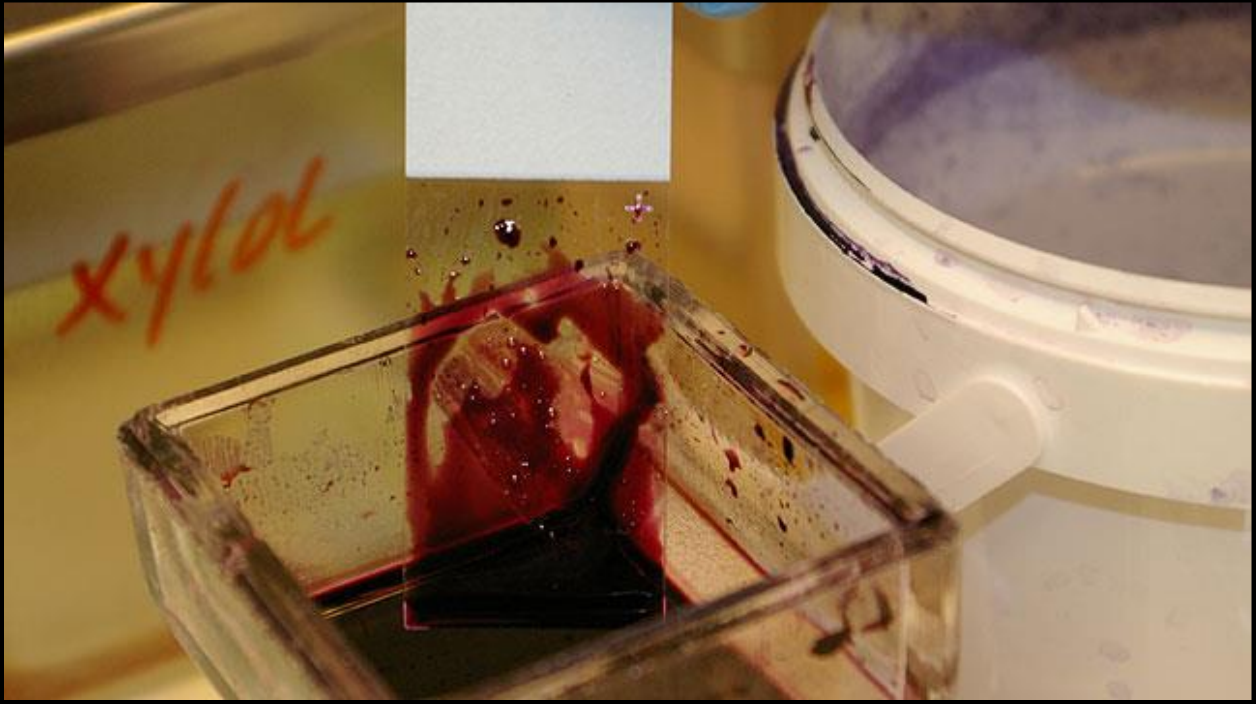


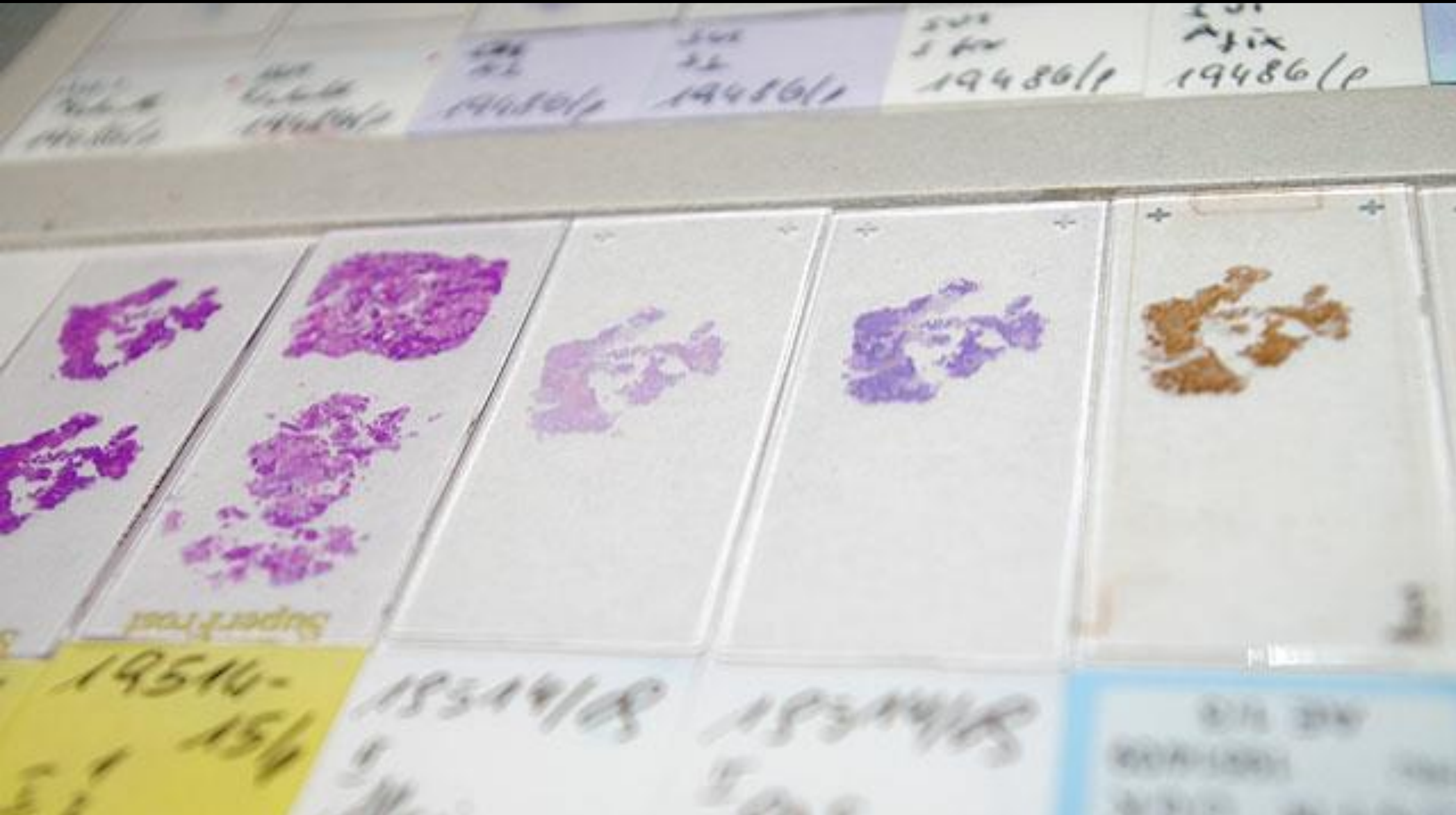









Xylol

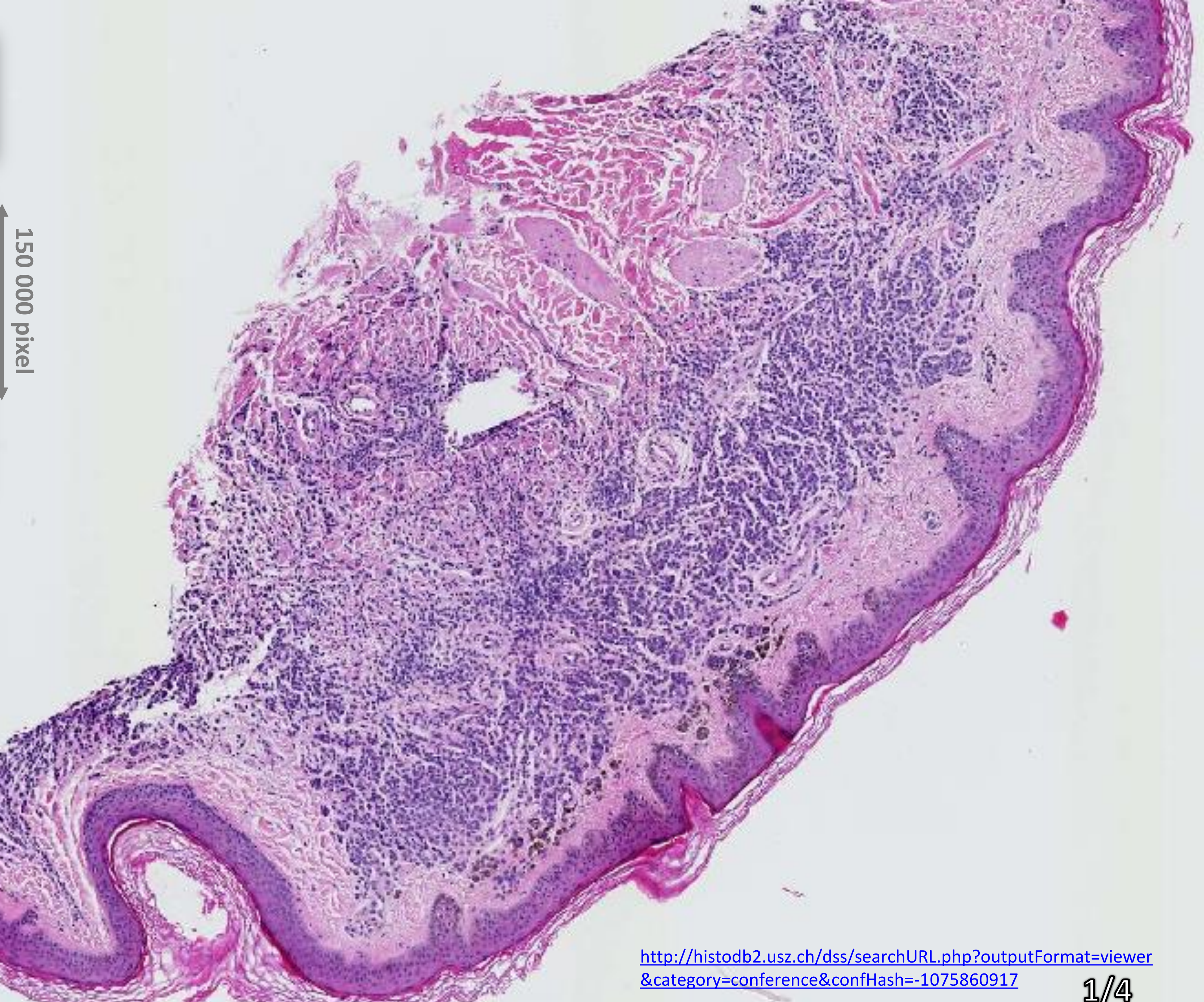


UniversitätsSpital Zürich
Departement Pathologie

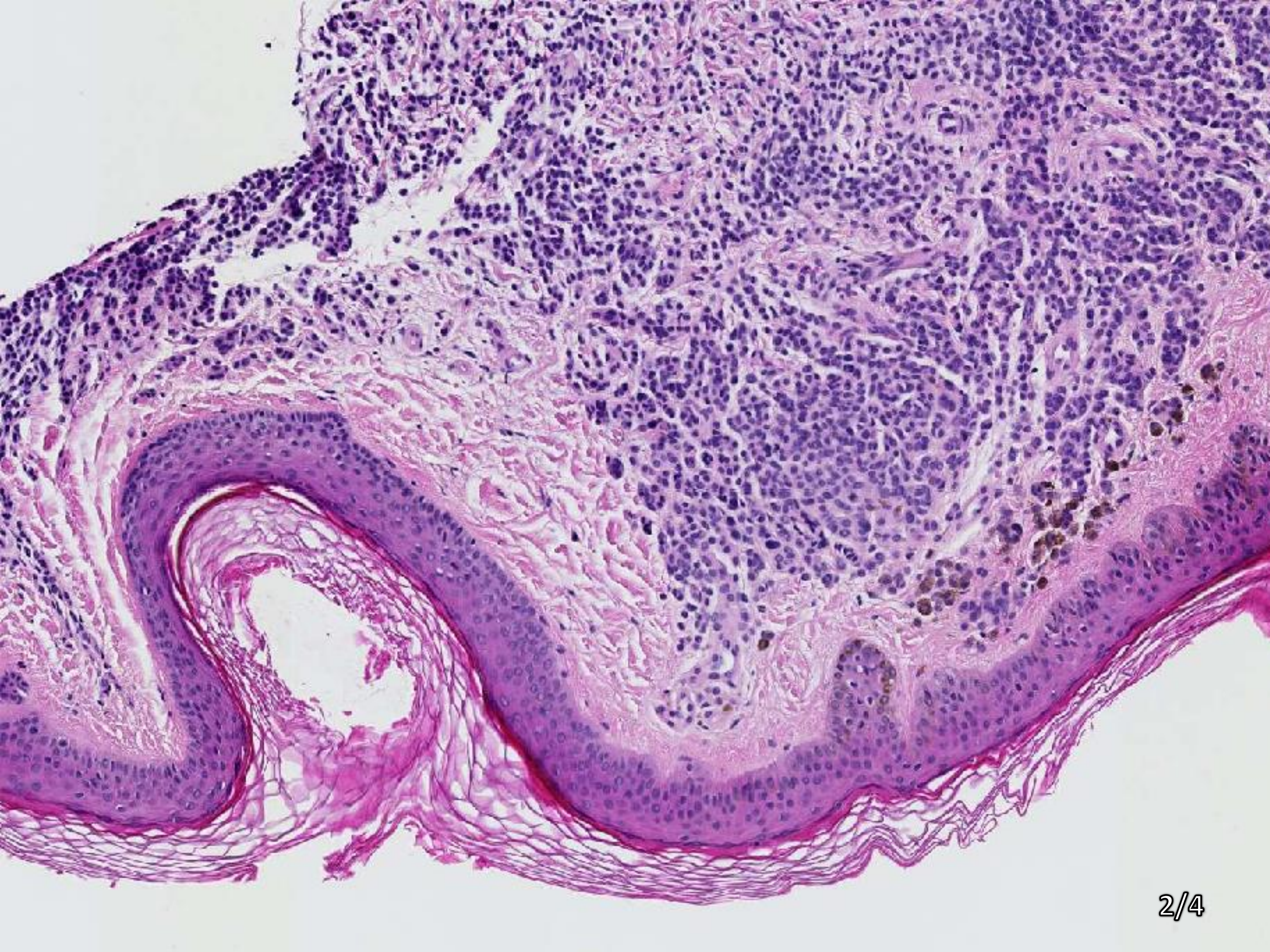
~67,000 cases per year

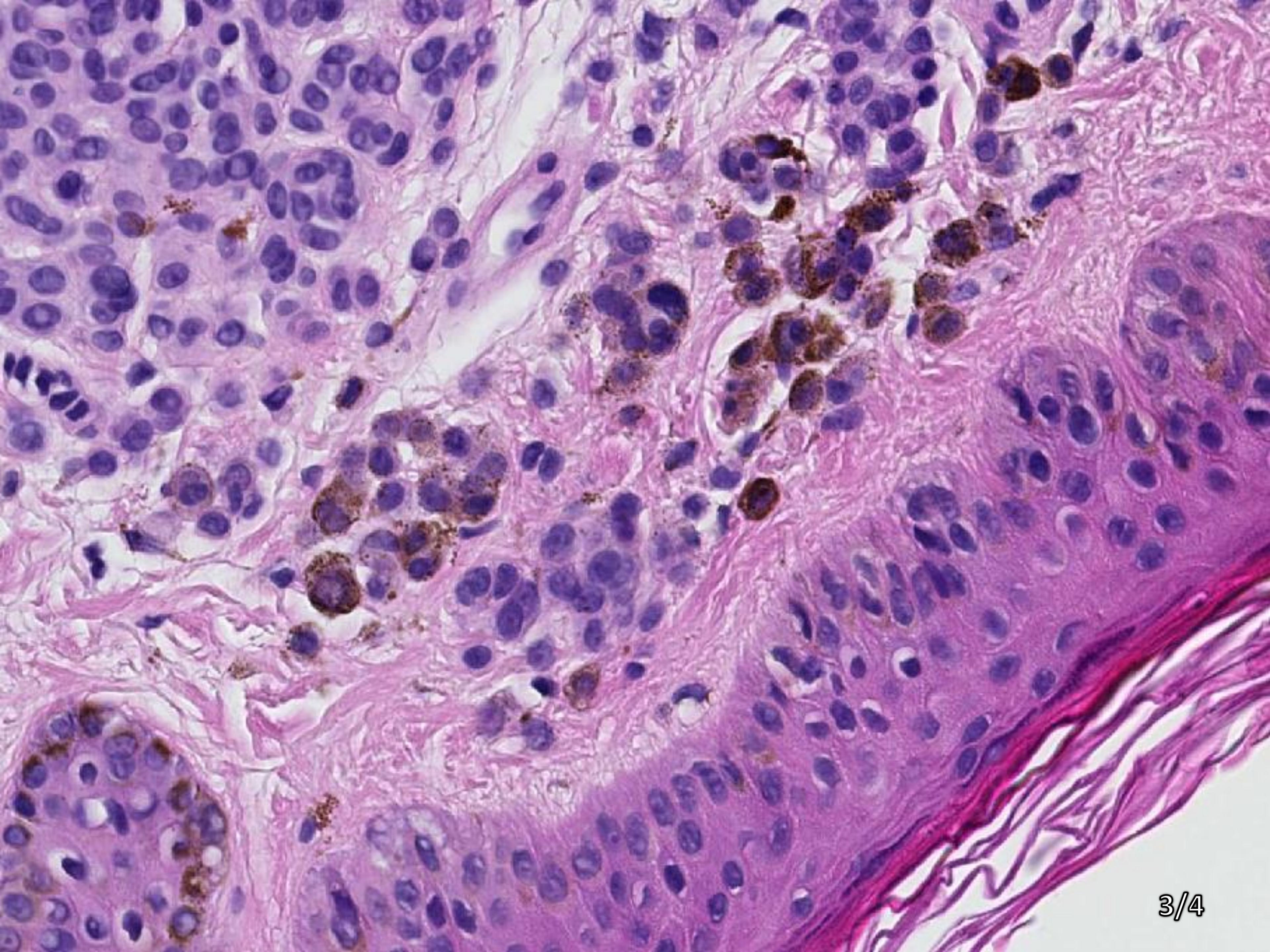

NDP

150 000 pixel

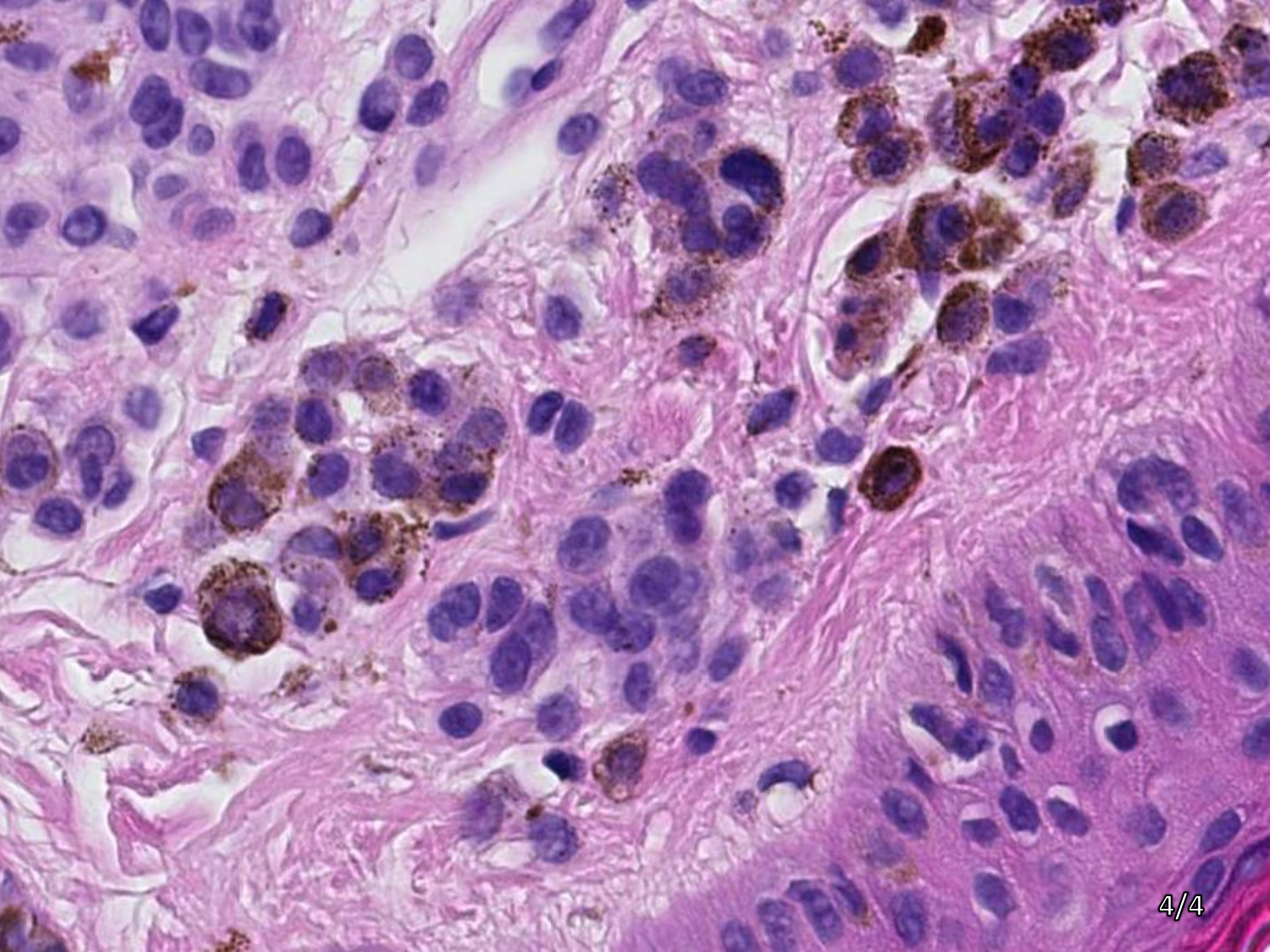


http://histodb2.usz.ch/dss/searchURL.php?outputFormat=viewer&category=conference&confHash=-1075860917

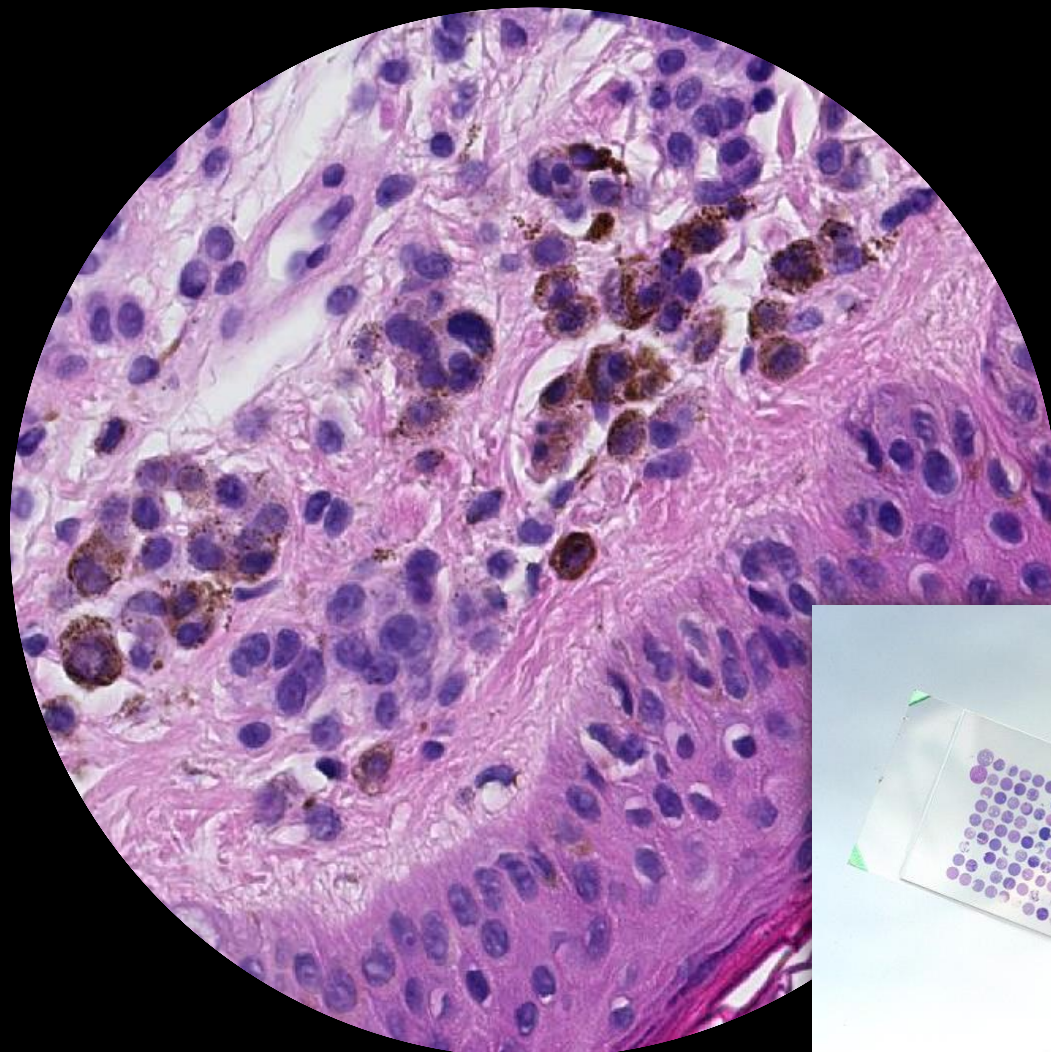


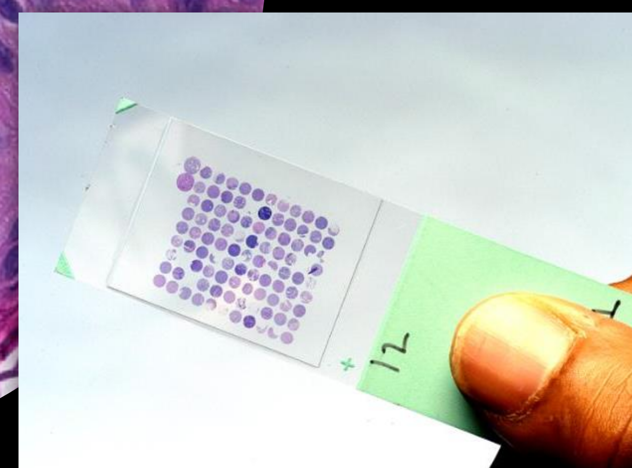


3/4



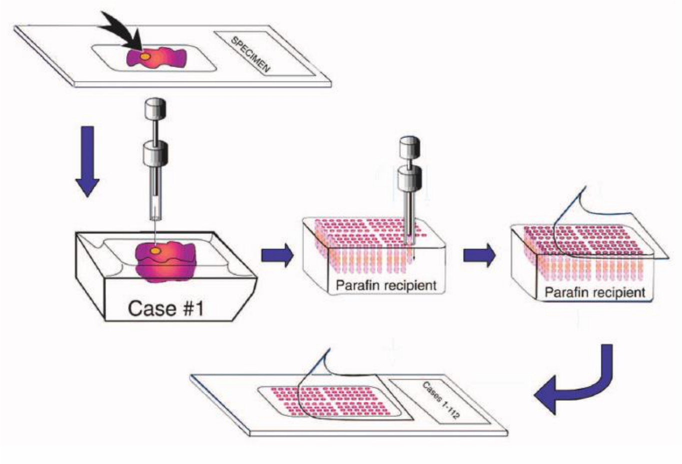
From Subjects to Cohorts

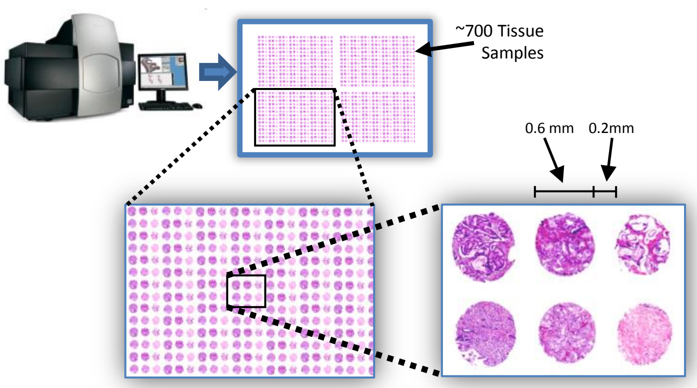


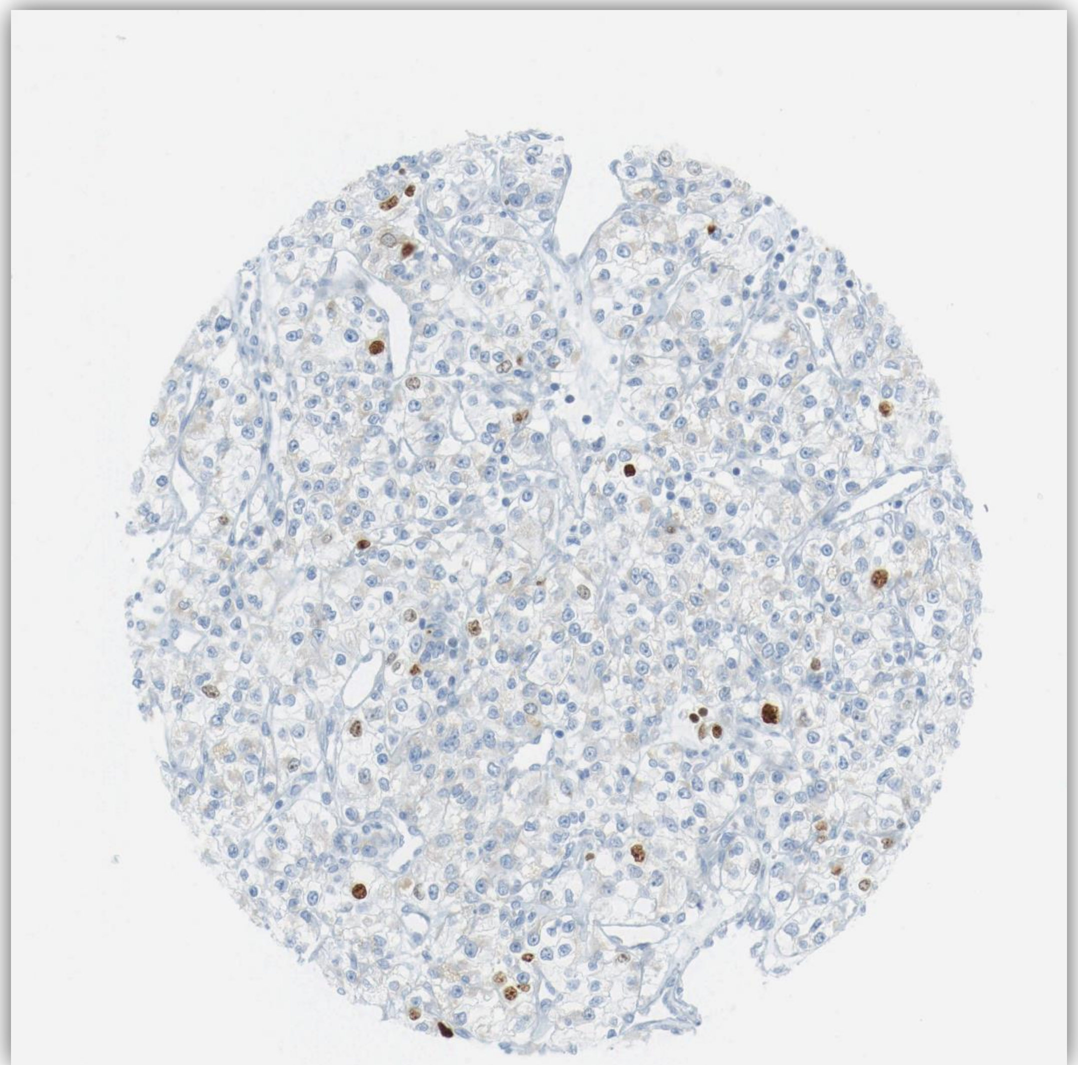


Tissue Microarray (TMA) Spot

Tissue Micro Array (TMA)

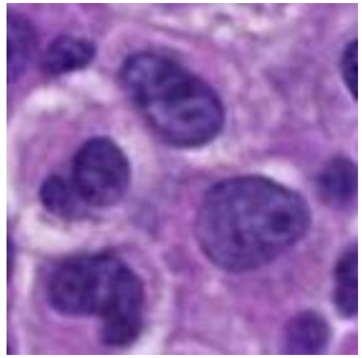


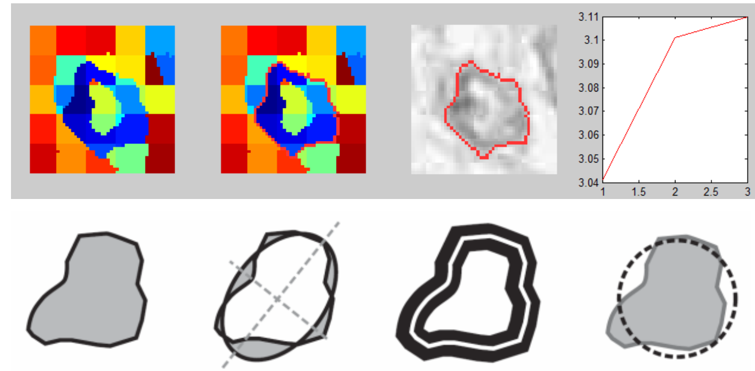




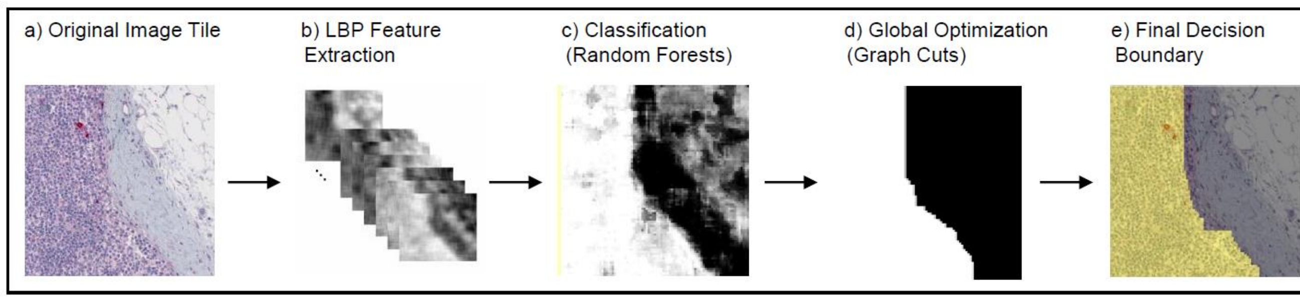
Computer Vision Tasks in Pathology

Nuclei Detection and Classification
Sub-cellular level

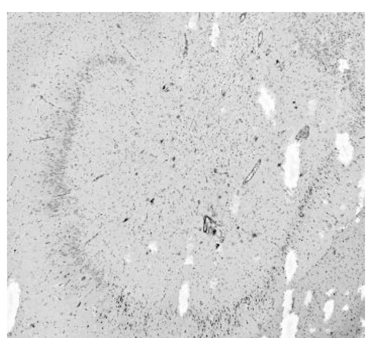


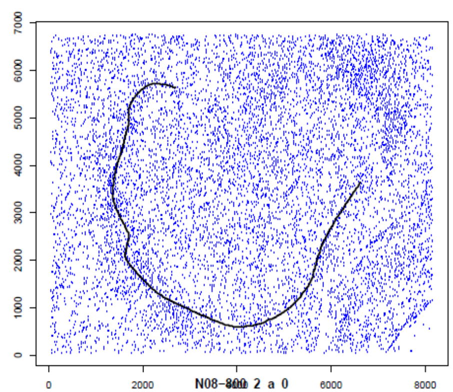


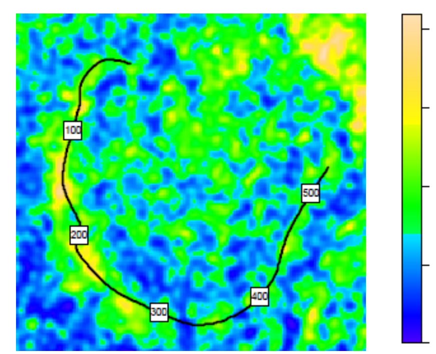
Segmentation



Structure Estimation Morphology

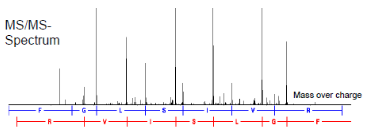




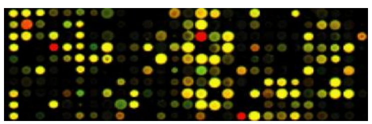


Biomarker Detection & Validation

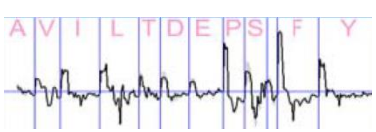
Proteomics



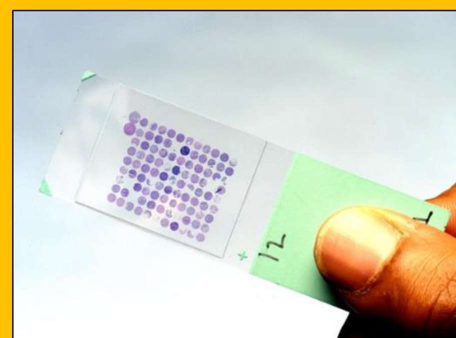
Transcriptomics



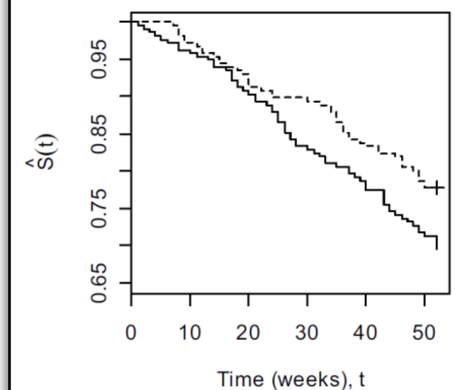
Metabolomics



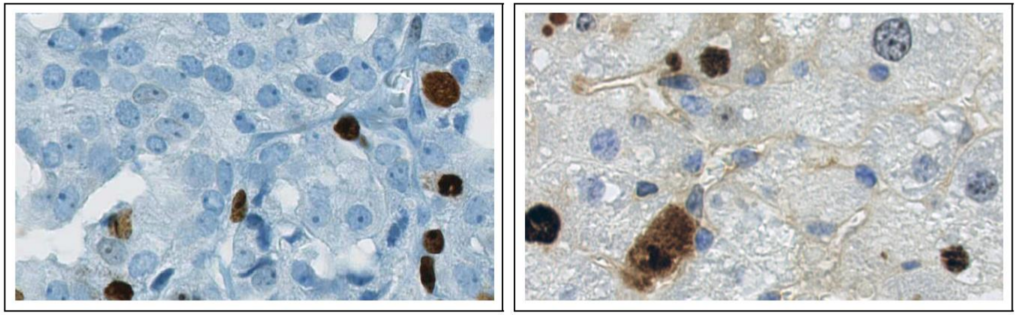
Human Tissue
TMA



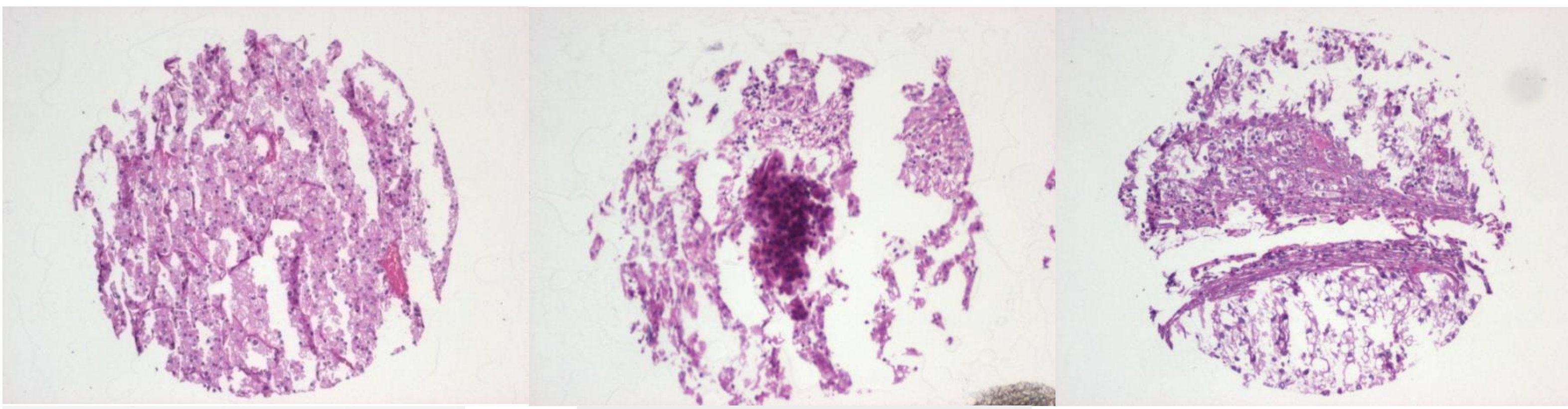
Clinical Trial



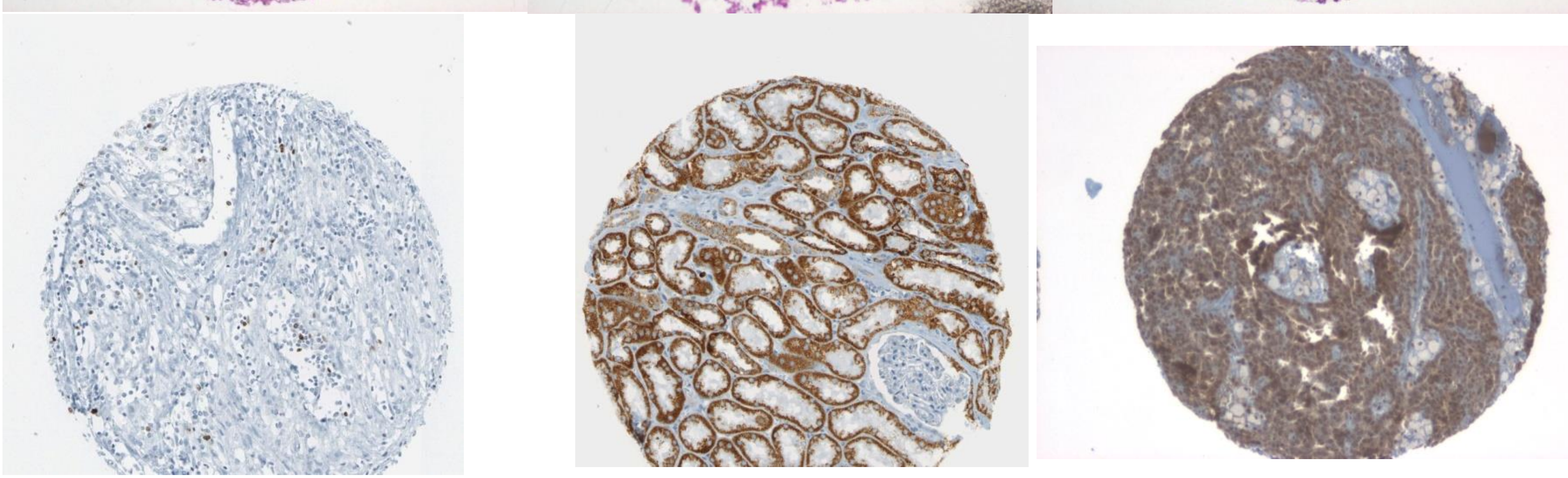
Variability



H & E

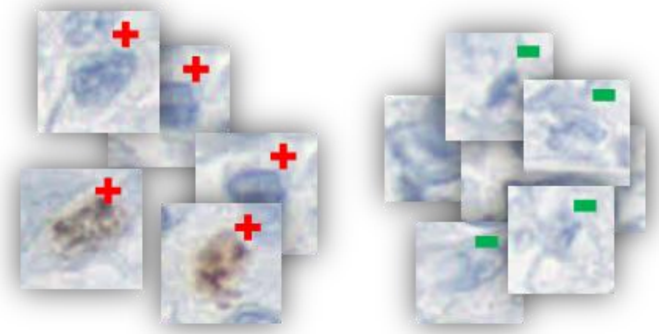


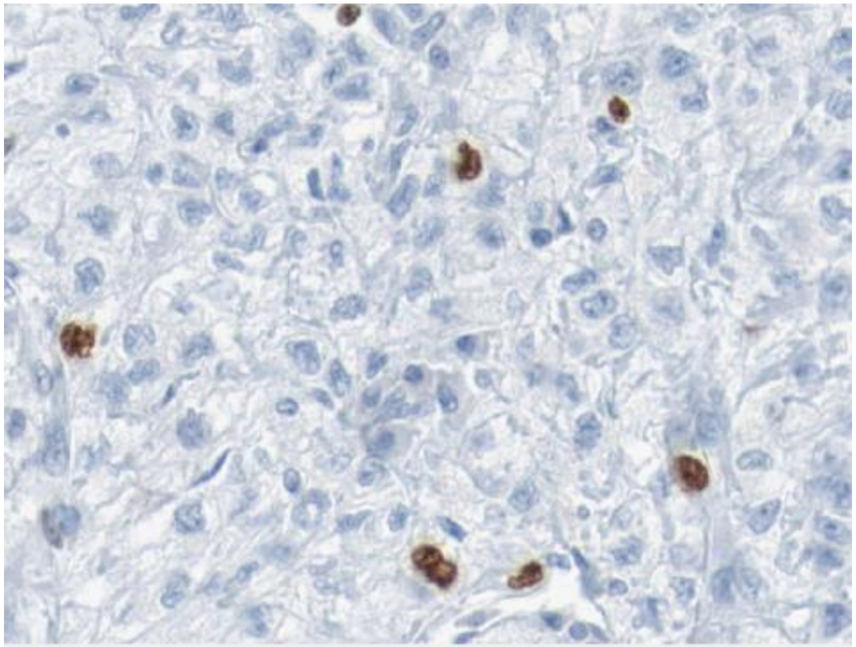
MIB-1



Ground Truth for Statistical Learning

Labeled samples are needed for training and validation.

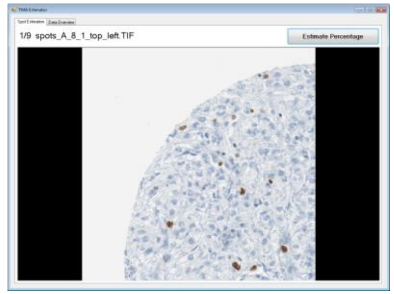


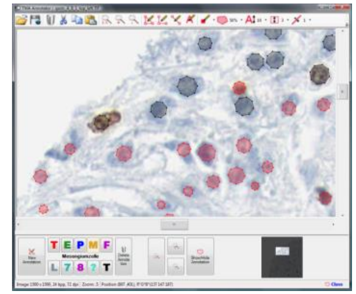


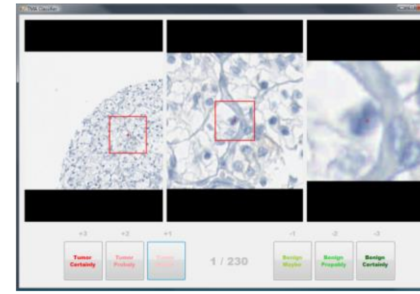
What is the „Ground Truth"?

Expert & Crowd Sourcing

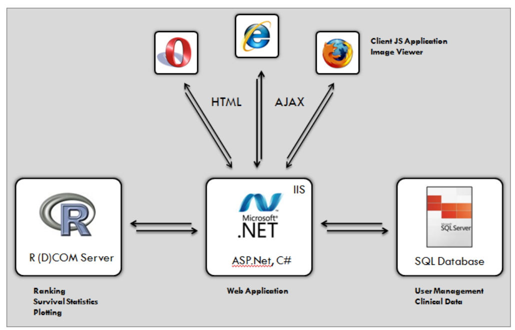
Past

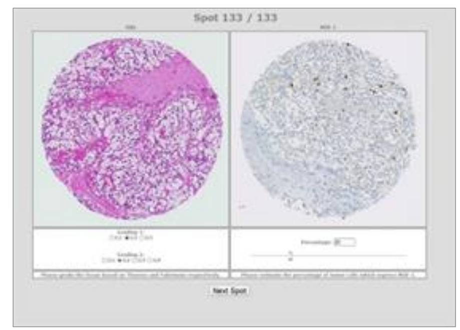




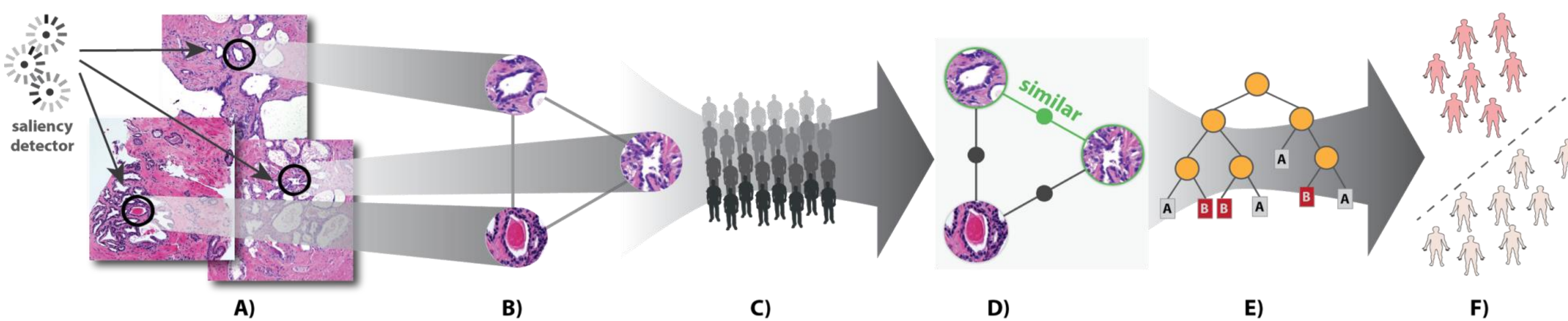


Present





Future



[BD2K Proposal 2014]

Staining Estimation Pipeline

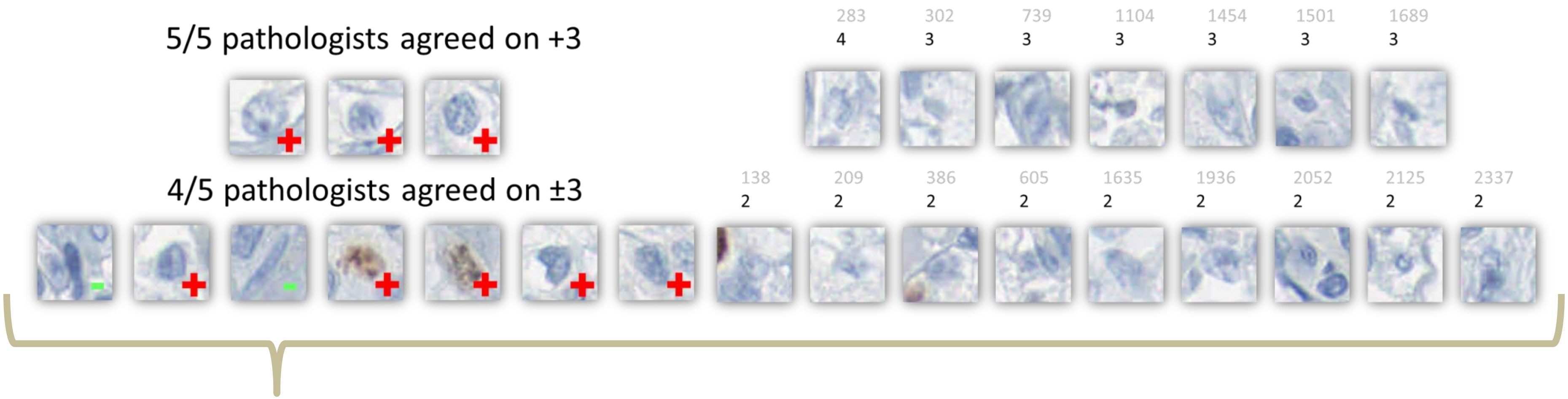
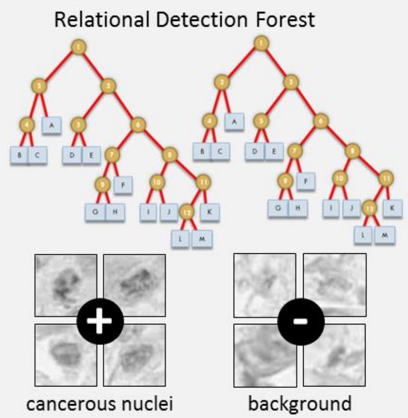


Image Analysis

Learning

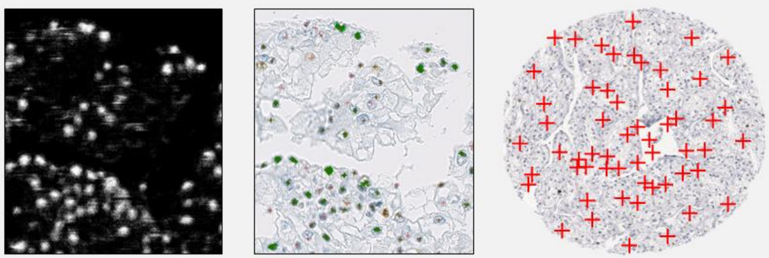


Prediction

Classifying every pixel of a spot results in a probability map

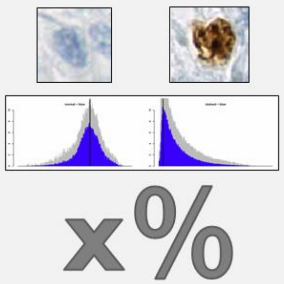
Nuclei centers are found by applying mean shift clustering

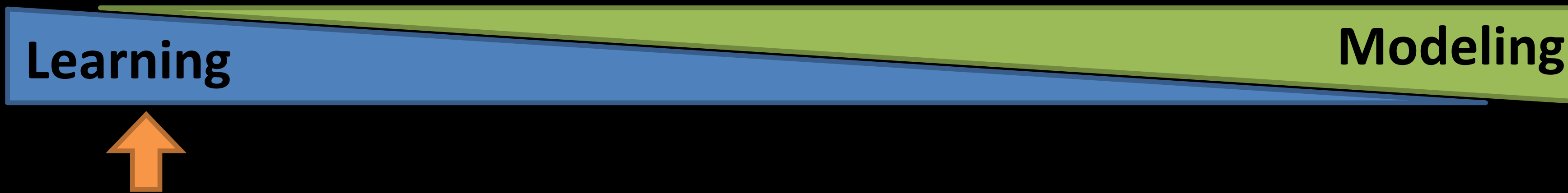
Several hundred nuclei are detected on each Image of a TMA spot



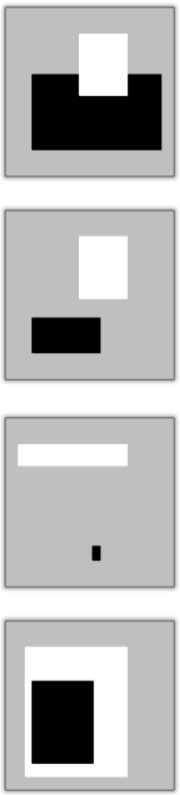
Estimation

Overall staining per patient is Calculated by assessing the staining of all detected nuclei





Relational Detection Forest



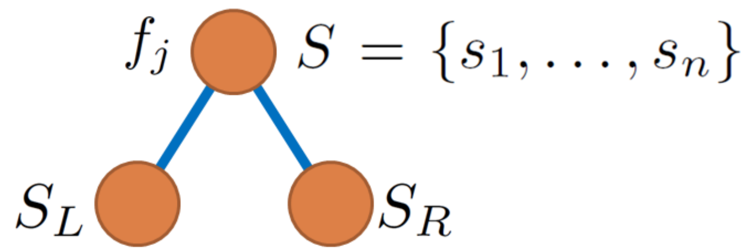
Procedure `LearnTree`

Input: set of samples $S = \{s_1, s_2, \dots, s_n\}$
Input: depth d
Input: max depth d_{max}
Input: features to sample $mTry$

Init: $\widehat{label} = null$; $g = -\inf$
Init: $N_{left} = null$; $N_{right} = null$
if $(d = d_{max})$ *OR* $(\mathtt{isPure}(S))$ **then**

$$\widehat{label} = \arg \max_{l \in \{true, false\}} \sum_{i | s_i = l} 1$$

else
  for $i = 0, i < mTry, i++$ **do**
    $f_i = \mathtt{SampleFeature}()$
    $S_L = \{s_j | f_i(s_j) = true\}$
    $S_R = \{s_j | f_i(s_j) = false\}$
    $g_i = \widehat{\Delta G}(S_L, S_R)$
    if $g_i > g$ **then**
      $f_{best} = f_i; \quad g = g_i$
    end
  end
  $N_{left} = \mathtt{LearnTree}(\{s_i | f_{best}(s_i) = true\})$
  $N_{right} = \mathtt{LearnTree}(\{s_i | f_{best}(s_i) = false\})$
end



Gini Index:

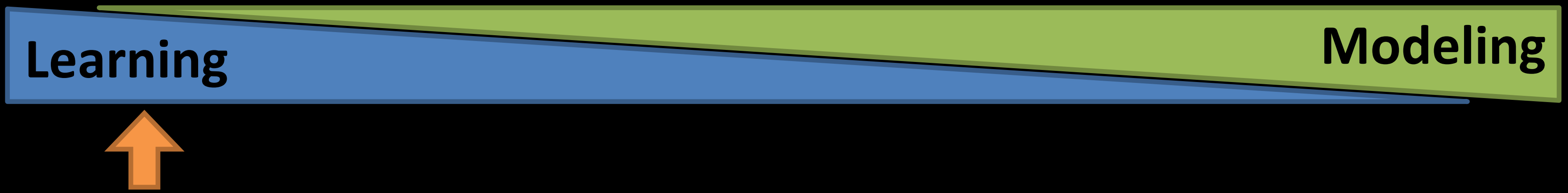
$$\widehat{G}(S) = 2\frac{N_{false}}{|S|}\left(1 - \frac{N_{false}}{|S|}\right)$$

$$N_{false} = \sum_{s_i} I\left(f_j(s_i) = false\right)$$

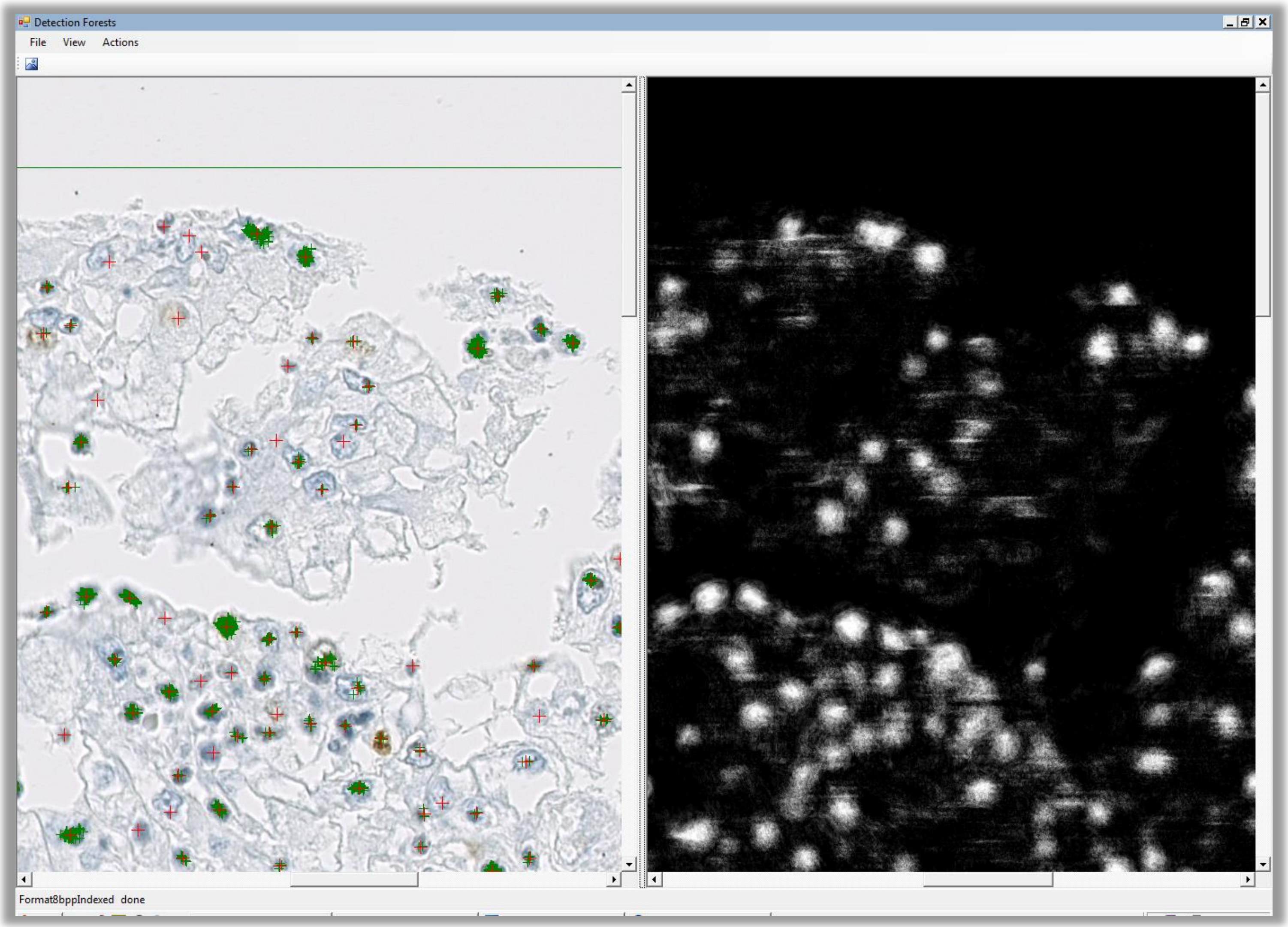
Gini Gain:

$$\widehat{\Delta G}(S_L, S_R) = \widehat{G}(S) - \left(\frac{|S_L|}{|S|}\widehat{G}(S_L) + \frac{|S_R|}{|S|}\widehat{G}(S_R)\right)$$

MICCAI 2006, ISVC 2009

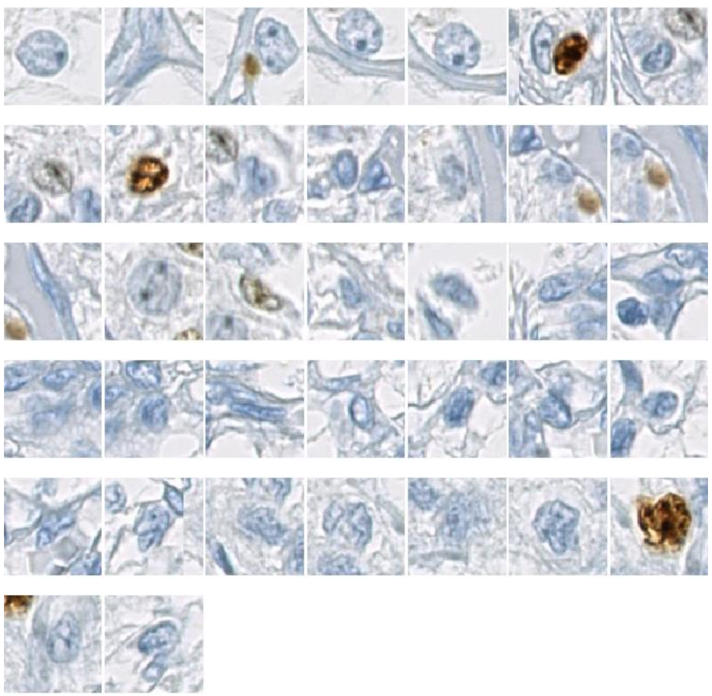


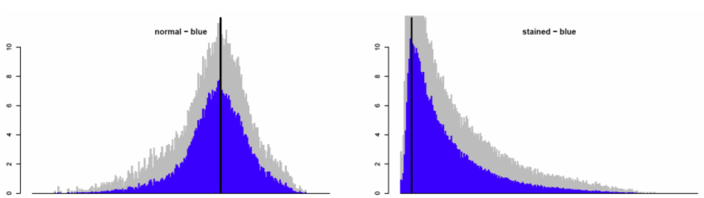
Relational Detection Forest

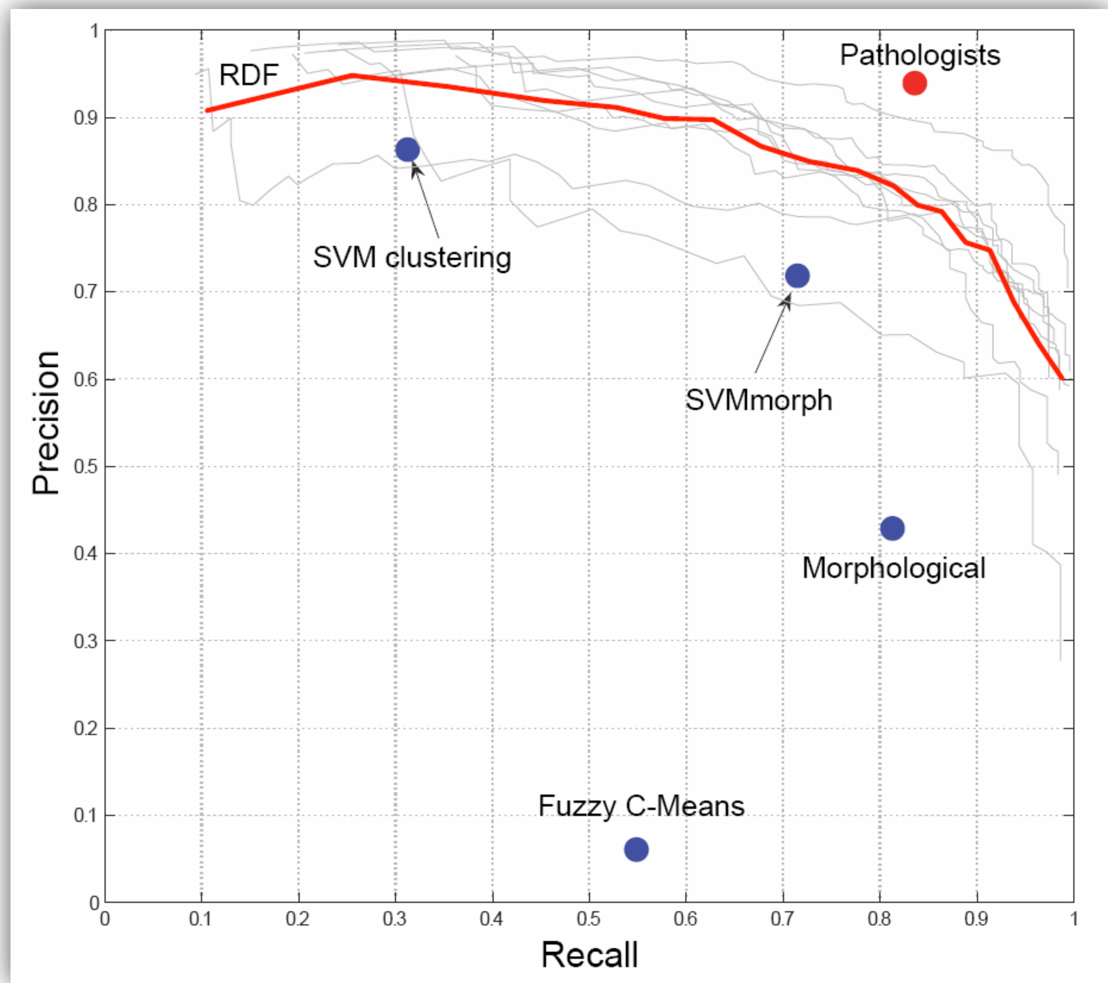


Learning — Modeling

Relational Detection Forest







ISVC 2009

Nucleus Based Analysis

$$\mathrm{NucleusIntensity}(n) = \frac{1}{|n|}\sum_{x\in n} x$$

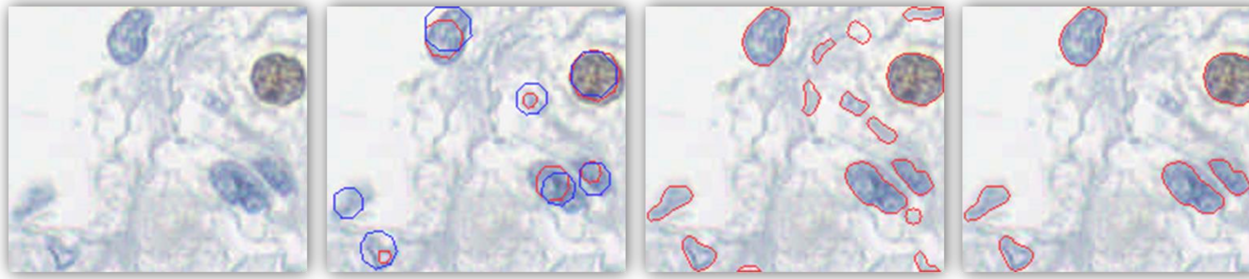
$$\mathrm{InnerIntensity}(n) = \frac{1}{|n|}\sum_{x\in[n\setminus\epsilon_B(n)]} x$$

$$\mathrm{OuterIntensity}(n) = \frac{1}{|n|}\sum_{x\in[\delta_B(n)\setminus n]} x$$

$$\mathrm{InnerHomogeneity}(n) = \mathrm{std}(x \in [n\setminus\epsilon_B(n)])$$

$$\mathrm{OuterHomogeneity}(n) = \mathrm{std}(x \in [\delta_B(n)\setminus n])$$

$$\mathrm{IntensityDifference}(n) = \frac{1}{|n|}\sum_{x\in(\delta_B(n)\setminus n)} x \cdot \left(\frac{1}{|n|}\sum_{x\in(n\setminus\epsilon_B(n))} x\right)^{-1}$$

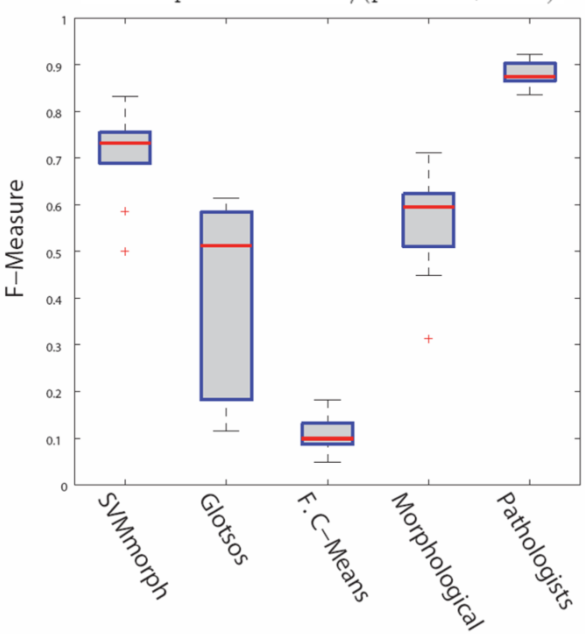


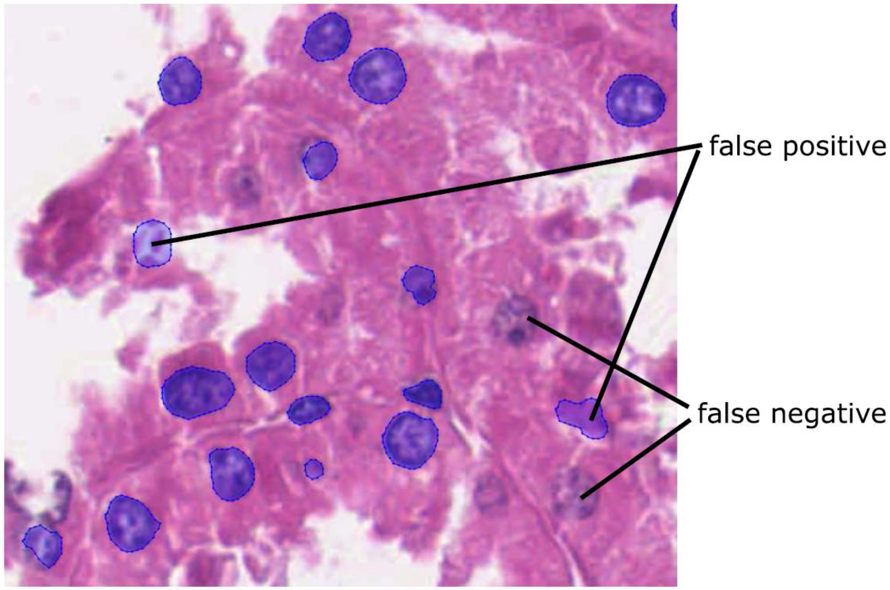
$$\mathrm{Size}(n) = \frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{(|n|-\mu)^2}{2\sigma^2}\right), \quad \mu = 600, \quad \sigma = 300$$

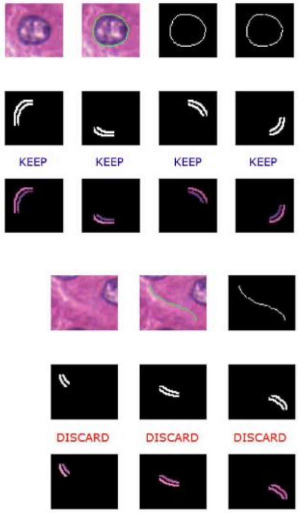
$$\mathrm{Ellipticity}(n) = \frac{|n_{Ellipse}|}{|n_{Ellipse}| + |(n_{Ellipse}\cup n)\setminus(n_{Ellipse}\cap n)|}$$

$$\mathrm{ShapeRegularity}(n) = \frac{2\pi\sqrt{\frac{n_{Area}}{\pi}}}{n_{Perim}}$$

$$F = 2 \times \mathrm{precision} \times \mathrm{recall}/(\mathrm{precision} + \mathrm{recall}).$$

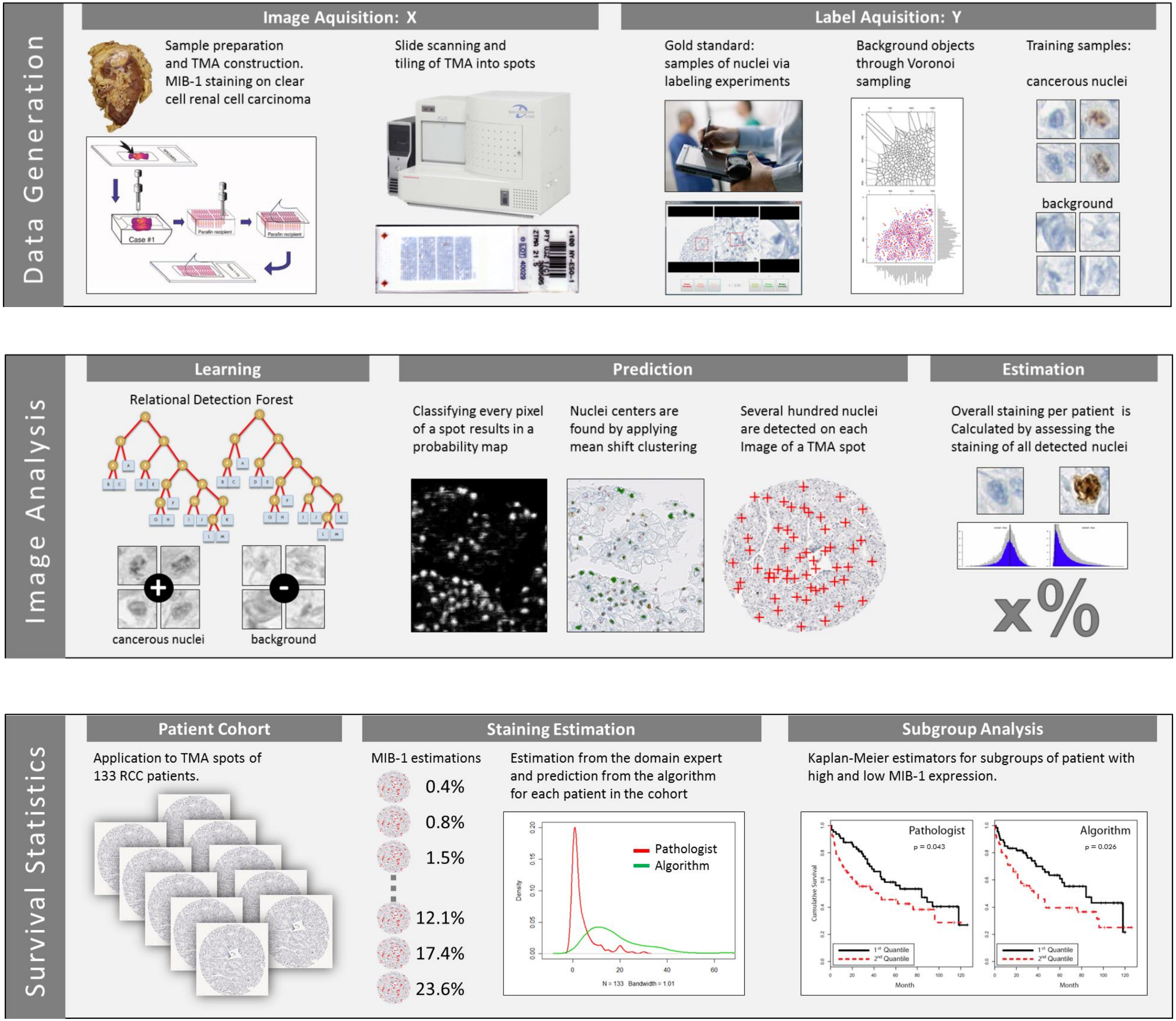




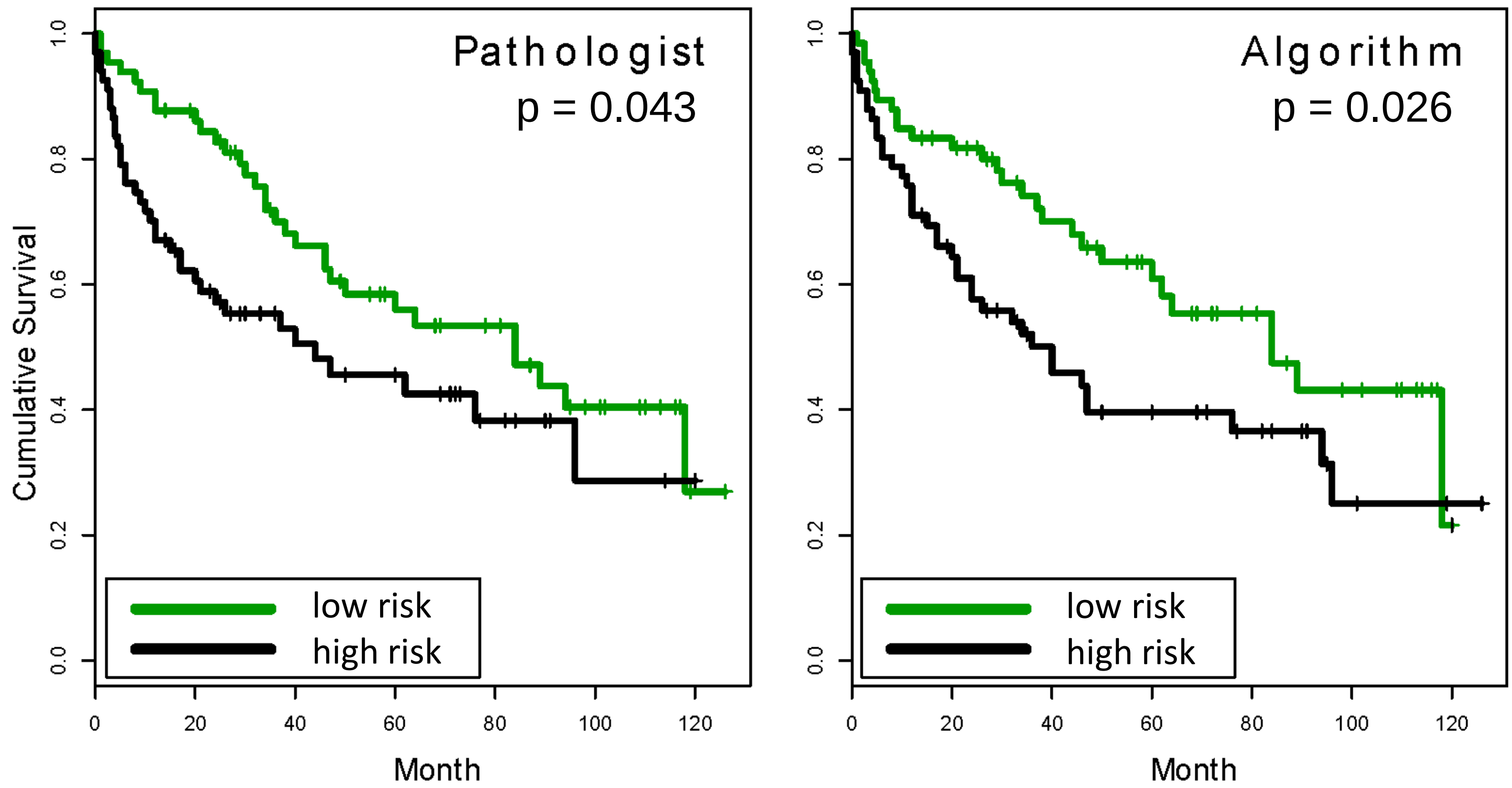


DAGM 2008

Computational Pathology

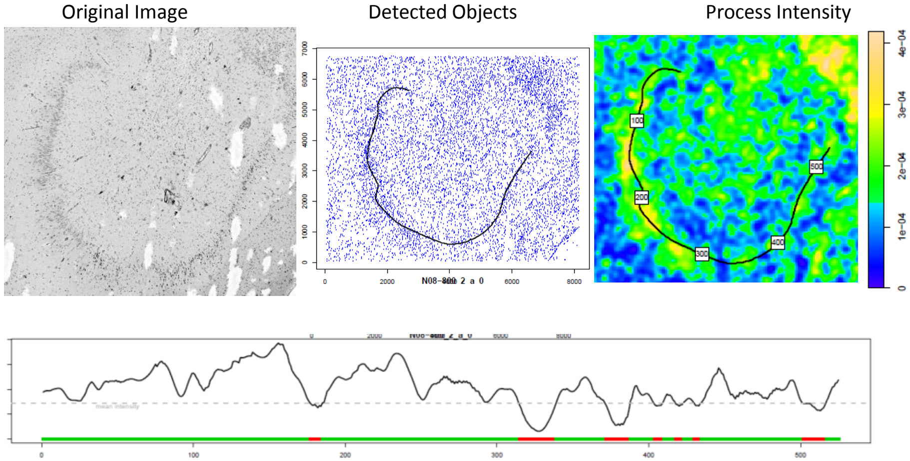


Survival Analysis

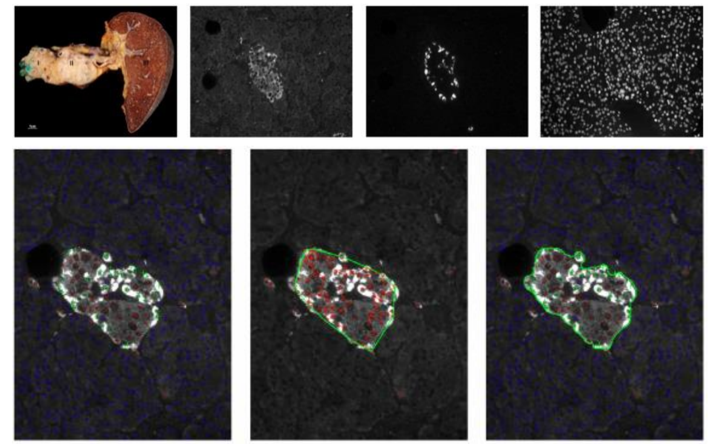


Applications of the Framework

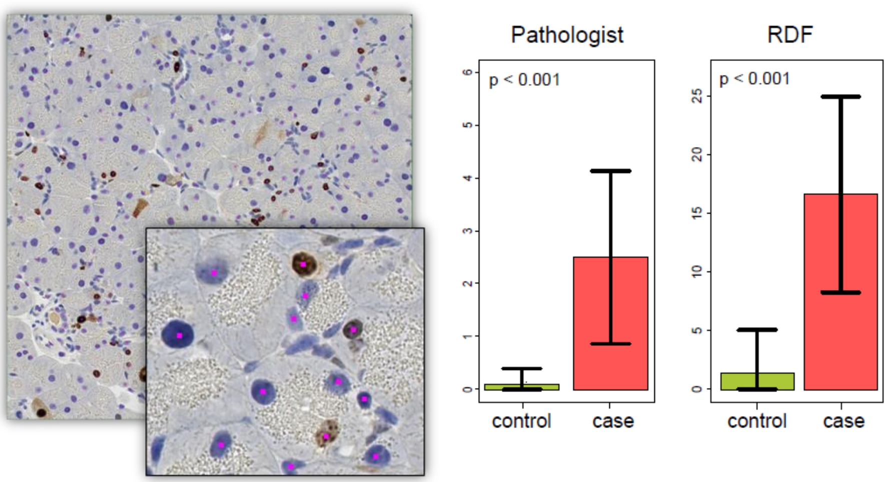
Spatial Processes for Hippocampal Sclerosis



Pancreatic Islet Segmentation for T2 Diabetes



Counting of Mouse Liver Hepatocytes



Detection in IHC Stained Cell Cultures