

Machine Learning for Cancer Diagnostics, Prognostics and Prediction

Fayyaz Minhas

Associate Professor, Department of Computer Science

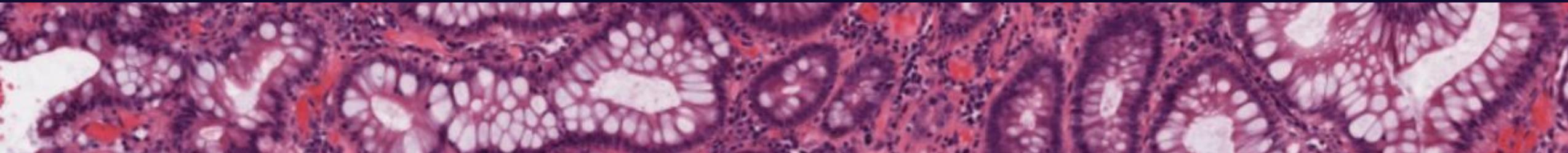
Deputy Director, Tissue Image Analytics Centre

Warwick Cancer Research Centre

University of Warwick



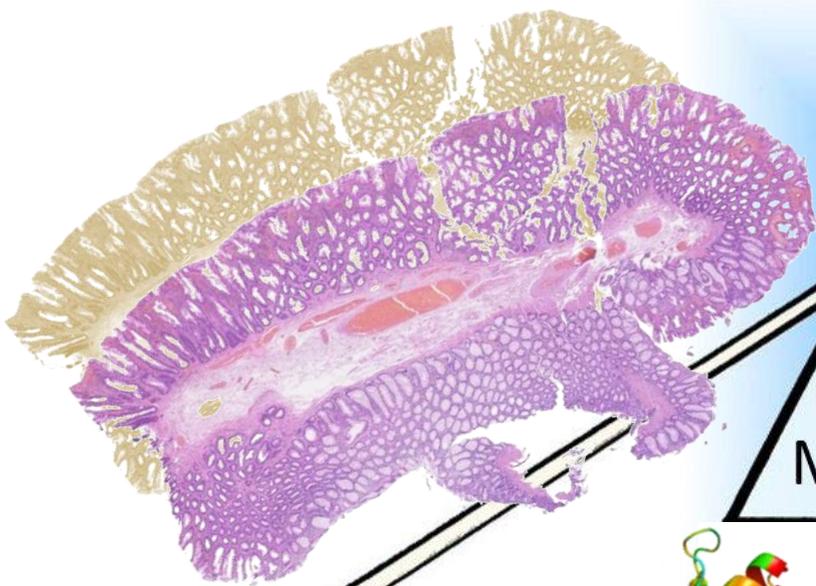
@fayyazhere



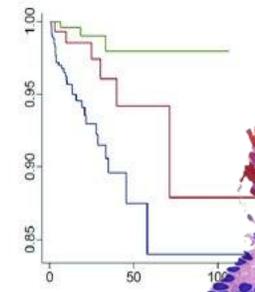
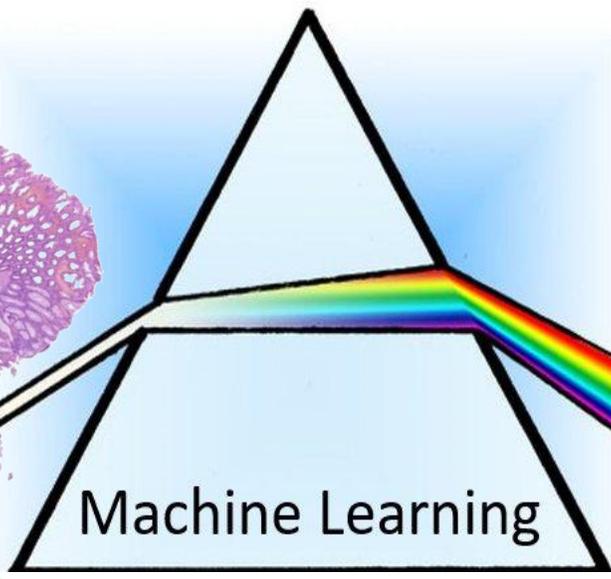
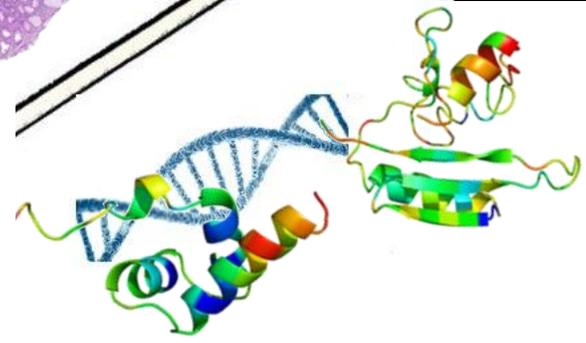
Case Number: XXXXX
 Diagnosis:
 A: Colon, transverse, partial colectomy
 Tumour Histologic Type: Invasive adenocarcinoma
 Histologic Grade: moderately differentiated (2 of 4)
 Tumour Location: transverse colon

Depth of Invasion:
 -Through muscularis propria and into and pericolic soft tissue MDO-3
 Lymphovascular Invasion: not identified
 Perineural Invasion: not identified

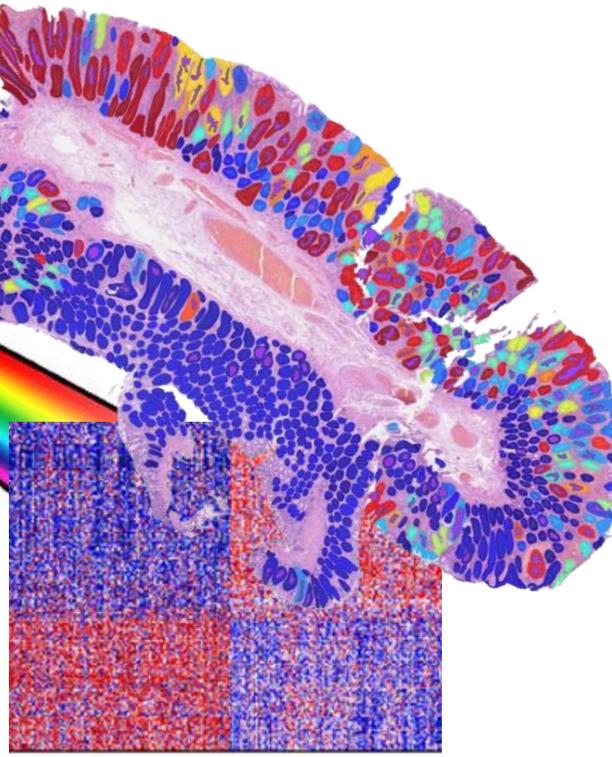
Margins:
 Proximal margin: negative
 Distal margin: negative
 Mesenteric margin: negative
 Distance of carcinoma from closest margin (specify): 6.6 cm to the closest distal margin of resection



Biomedical Data



Predictive & Prognostic Insights

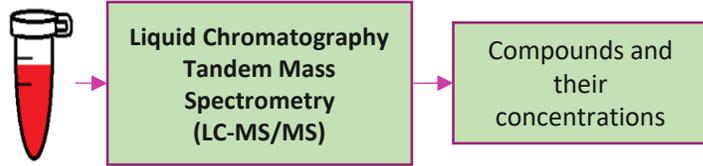


AI and ML to help discover and understand biology and pathology

 [@fayyazhere](https://twitter.com/fayyazhere) 



• AI for Biology



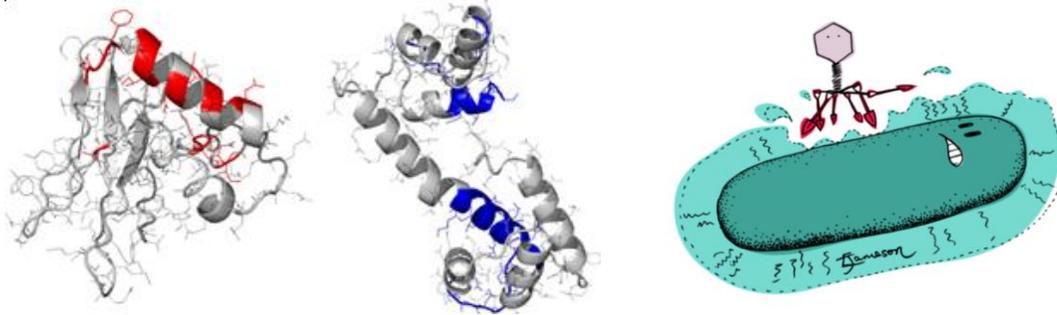
RAMClust: Spectral-matching-based annotation for metabolomics data

Nucleic Acids Research, 2020 1
doi: 10.1093/nar/gkaa219



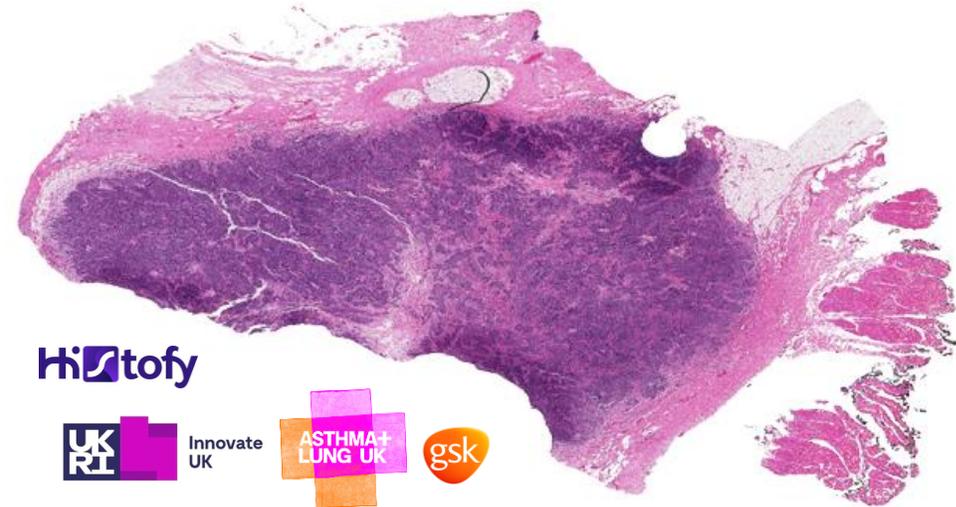
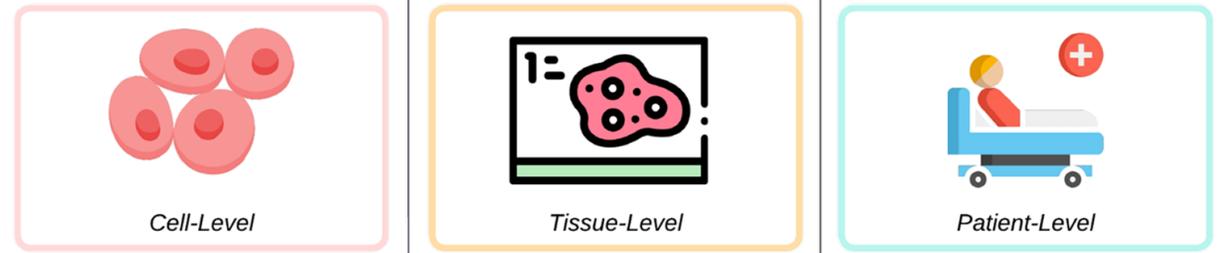
Machine learning predicts new anti-CRISPR proteins

Simon Eitzinger^{1,†}, Amina Asif^{2,3,†}, Kyle E. Watters^{1,†}, Anthony T. Iavarone⁴, Gavin J. Knott¹, Jennifer A. Doudna^{1,5,6,7,8,9,*} and Fayyaz ul Amir Afsar Minhas^{2,10,*}

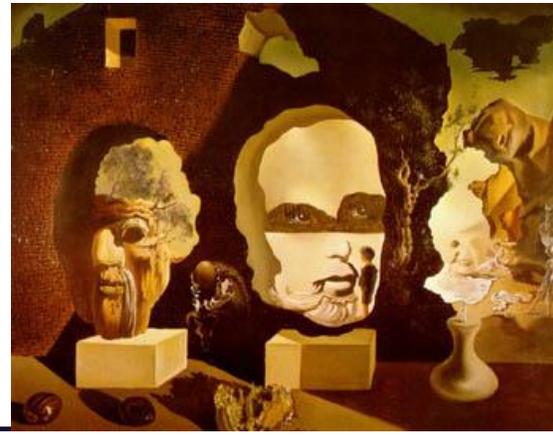
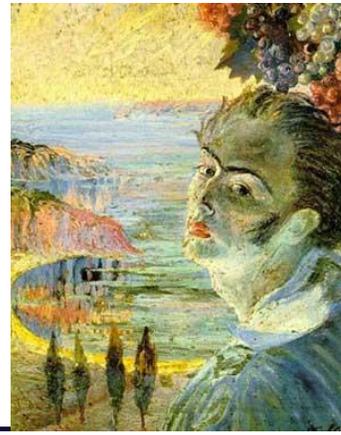
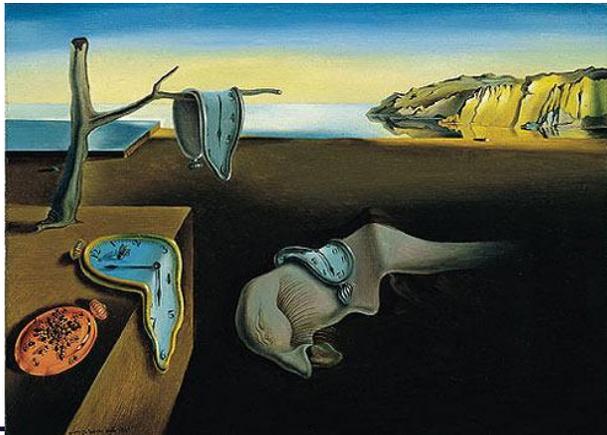


Prediction of Protein Interfaces, Drug-Protein Interactions and Protein Function (Anti-CRISPR proteins, Anti-Microbial & Hemolytic peptides, Amyloid & Prion activity, Phage antibacterial proteins)

• AI for Pathology & Diagnostics



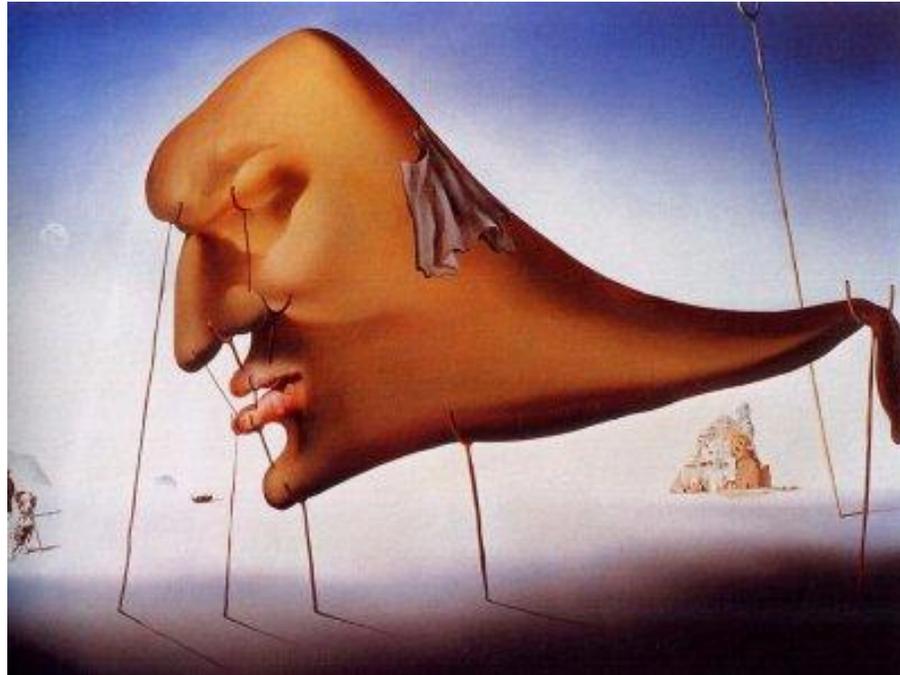
Paintings by two different painters



Who's painting is this?

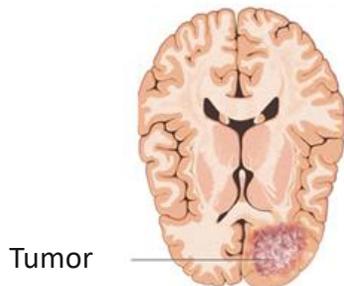


And this?

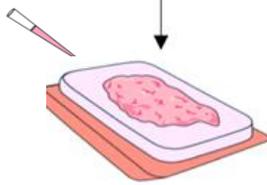


learning from data for generalization to unseen cases
inductive inference

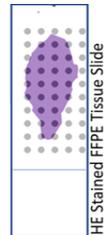
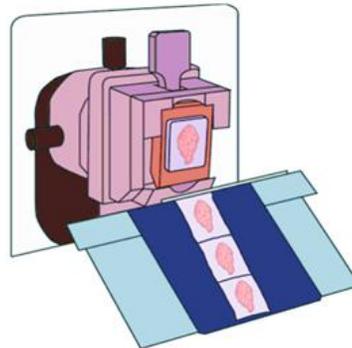
Workflow: Conventional Histopathology



Tumor



Block preparation
Addition of dyes



Small bowel biopsy worksheet

Case#: _____ PT. Name: _____ Unit#: _____ Date: _____ Pathologist's name: _____

<p>I. Clinical information None provided Rule out coeliac disease Previous diagnosis of coeliac disease Other (specify): _____</p> <p>II. Type of mucosa Duodenal Jejunal</p> <p>III. Number of biopsy pieces</p> <p>IV. Adequacy of specimen orientation (At least 4 consecutive well oriented crypt-villus units) Yes No</p> <p>V. Villous length Normal Abnormal Cannot be determined</p> <p>VI. Crypt length Normal Elongated Shortened</p> <p>VII. Crypt-to-villus ratio Normal (1.3 - 1.5) Abnormal (specify): _____</p> <p>VIII. Villous atrophy None Total Subtotal Partial</p>	<p>IX. Intraepithelial lymphocytes (IELs) Normal (up to 30 per 100 epithelial nuclei) Increased, specify: mild, moderate, severe, focal, diffuse</p> <p>X. Lamina propria inflammatory cells Normal Increased, specify: plasma cells, eosinophils, lymphocytes, neutrophils</p> <p>XI. Gastric metaplasia Present Absent</p> <p>XII. Subepithelial collagen Normal Increased</p> <p>XIII. Diagnosis Small intestinal mucosa, histologically unremarkable Focal chronic inflammation with or without focal intraepithelial lymphocytes, nonspecific Diffuse intraepithelial lymphocytosis, normal villi, consistent with coeliac disease (MARSH-1) Diffuse intraepithelial lymphocytosis and crypt hyperplasia, consistent with coeliac disease (MARSH-2) Partial villous atrophy, crypt hyperplasia and intraepithelial lymphocytosis, consistent with coeliac disease (MARSH-3a) Subtotal villous atrophy, crypt hyperplasia and intraepithelial lymphocytosis, consistent with coeliac disease (MARSH-3b) Total villous atrophy, crypt hyperplasia and intraepithelial lymphocytosis, consistent with coeliac disease (MARSH-3c) Villous atrophy, crypt hyperplasia and intraepithelial lymphocytosis, consistent with coeliac disease (MARSH-4) Other (specify): _____</p>
---	--

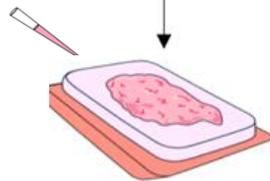
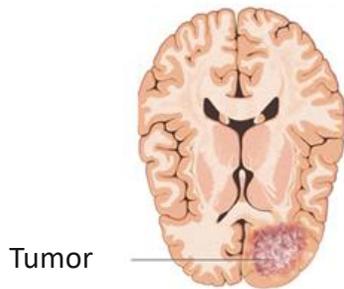
Workflow: Digital Pathology

Tissue
Acquisition

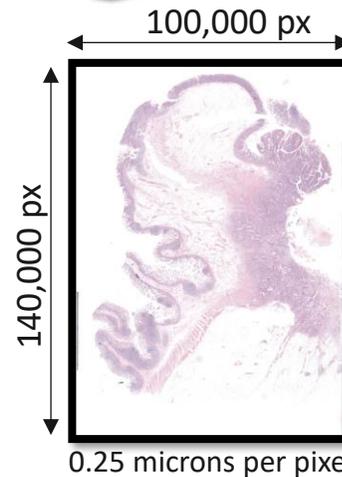
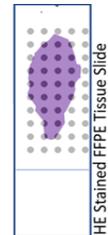
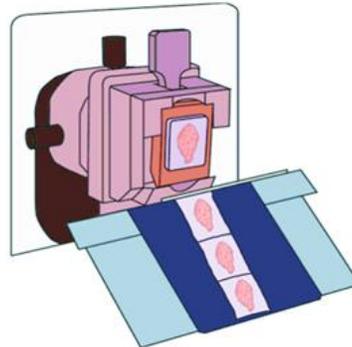
Slide
Preparation

Whole Slide
Scanning

Clinical
Decision
Making



Block preparation
Addition of dyes

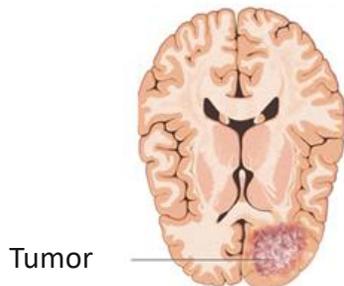
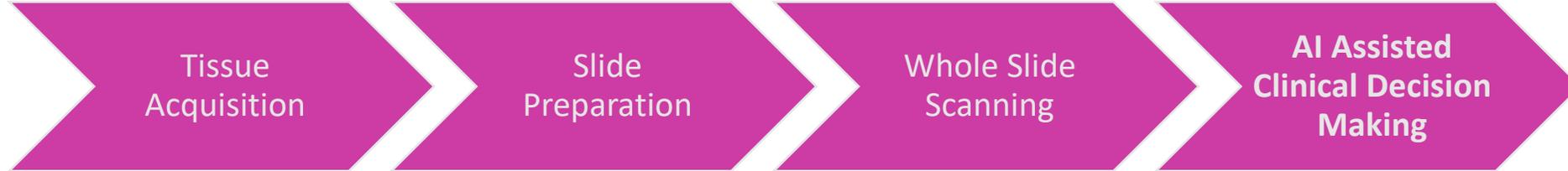


- Flexible working
- Good concordance with slide based clinical decision making based on equivalence studies
- However: *digitization of glass slides alone does not resolve the pressures of an increasing workload on a diminishing workforce of pathologists*

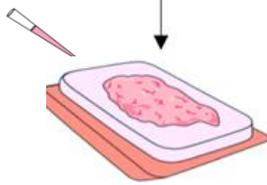
Example equivalence studies: Snead, David R. J., Yee-Wah Tsang, Aisha Meskiri, Peter K. Kimani, Richard Crossman, Nasir M. Rajpoot, Elaine Blessing, et al. "Validation of Digital Pathology Imaging for Primary Histopathological Diagnosis." *Histopathology* 68, no. 7 (June 2016): 1063–72. <https://doi.org/10.1111/his.12879>.

Hanna, Matthew G., Victor E. Reuter, Meera R. Hameed, Lee K. Tan, Sarah Chiang, Carlie Sigel, Travis Hollmann, et al. "Whole Slide Imaging Equivalency and Efficiency Study: Experience at a Large Academic Center." *Modern Pathology* 32, no. 7 (July 2019): 916–28. <https://doi.org/10.1038/s41379-019-0205-0>.

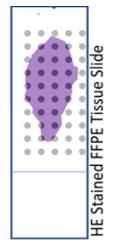
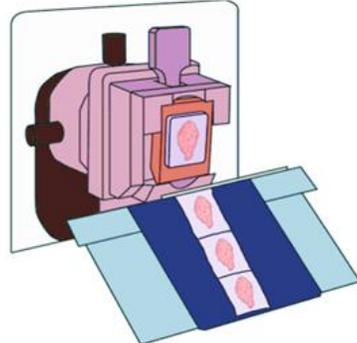
Workflow: Computational Pathology



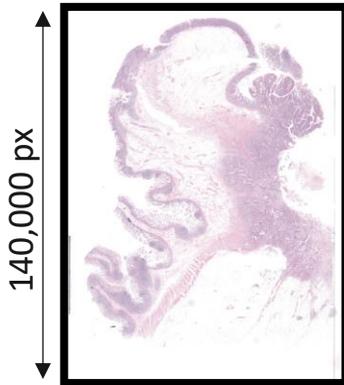
Tumor



Block preparation
Addition of dyes

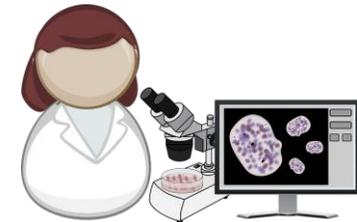


100,000 px

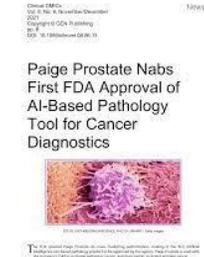


140,000 px

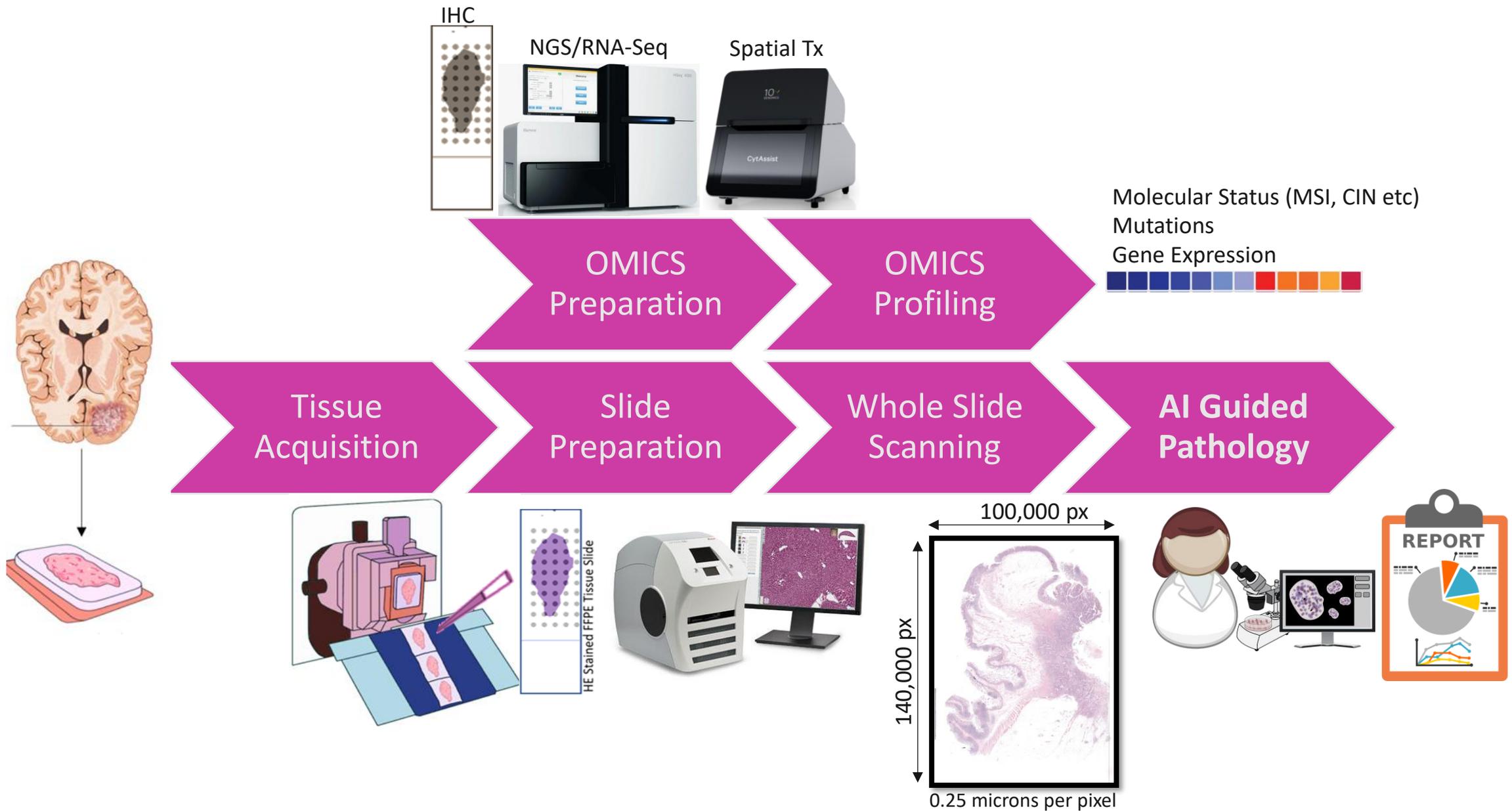
0.25 microns per pixel



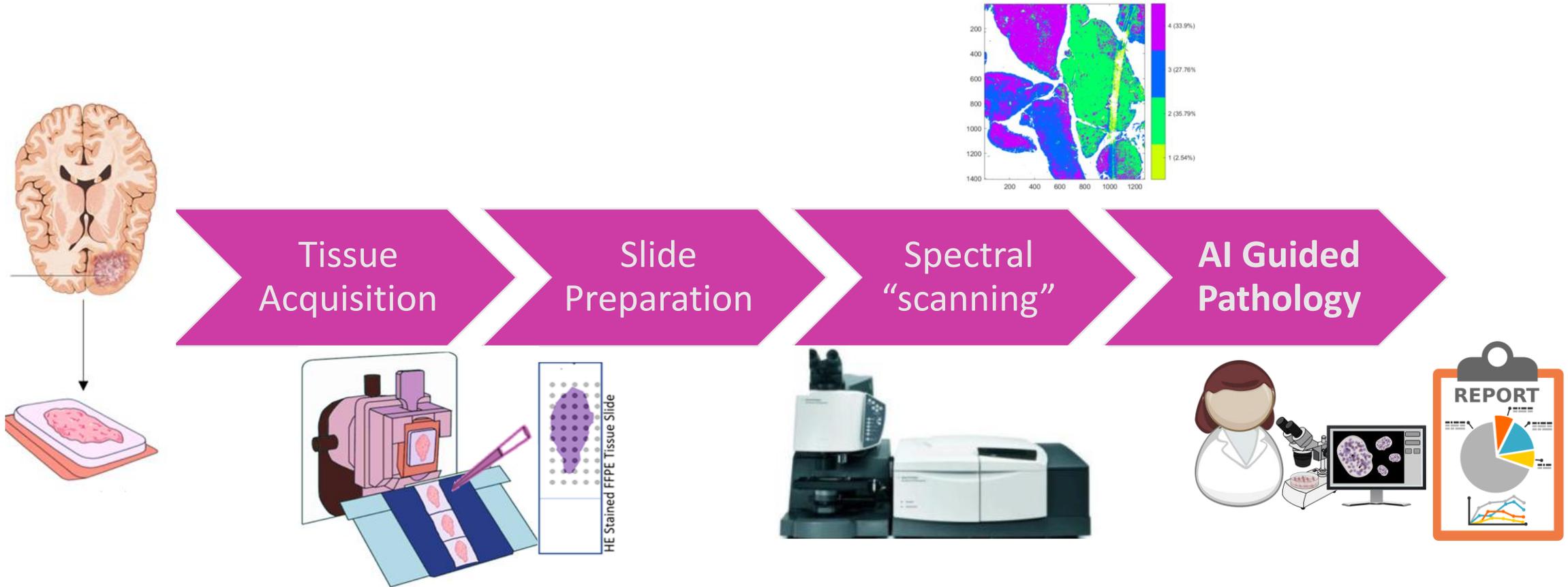
<https://info.paige.ai/prostate>



Example Independent validation study of PAIGE Prostate: Kanan, Christopher, Jillian Sue, Leo Grady, Thomas J. Fuchs, Sarat Chandarlapaty, Jorge S. Reis-Filho, Paulo G O Salles, Leonard Medeiros da Silva, Carlos Gil Ferreira, and Emilio Marcelo Pereira. "Independent Validation of Paige Prostate: Assessing Clinical Benefit of an Artificial Intelligence Tool within a Digital Diagnostic Pathology Laboratory Workflow." *Journal of Clinical Oncology* 38, no. 15_suppl (May 20, 2020): e14076–e14076. https://doi.org/10.1200/JCO.2020.38.15_suppl.e14076.

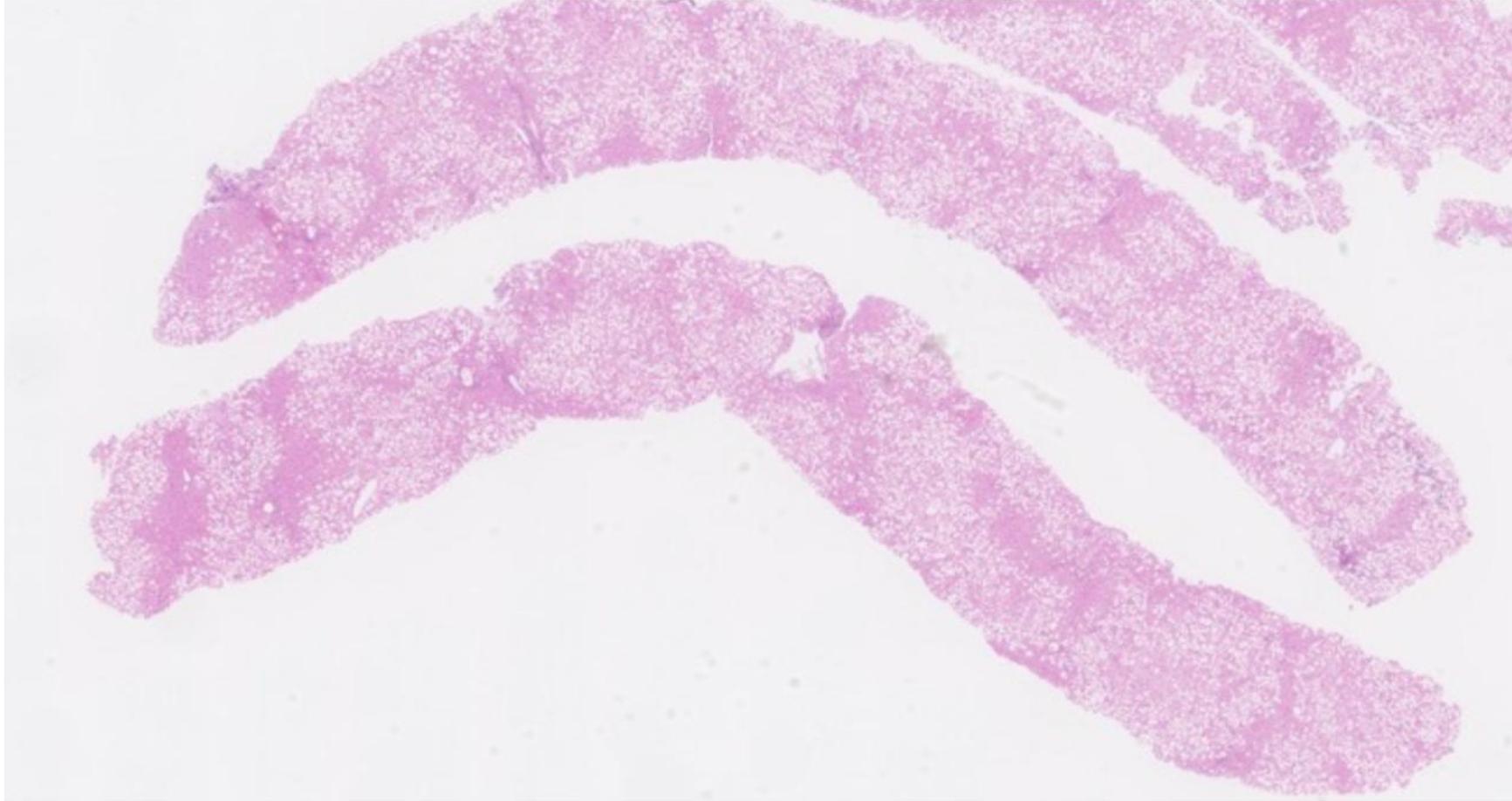


Example Independent validation study of PAIGE Prostate: Kanan, Christopher, Jillian Sue, Leo Grady, Thomas J. Fuchs, Sarat Chandarlapaty, Jorge S. Reis-Filho, Paulo G O Salles, Leonard Medeiros da Silva, Carlos Gil Ferreira, and Emilio Marcelo Pereira. "Independent Validation of Paige Prostate: Assessing Clinical Benefit of an Artificial Intelligence Tool within a Digital Diagnostic Pathology Laboratory Workflow." *Journal of Clinical Oncology* 38, no. 15_suppl (May 20, 2020): e14076–e14076. https://doi.org/10.1200/JCO.2020.38.15_suppl.e14076.



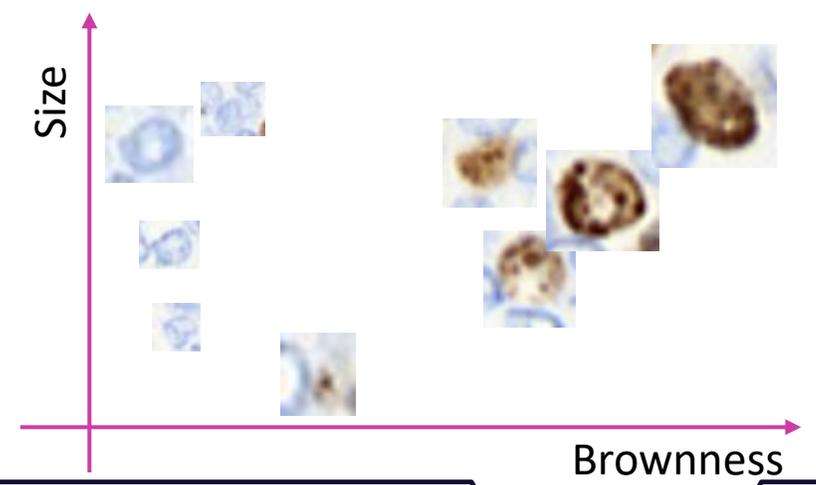
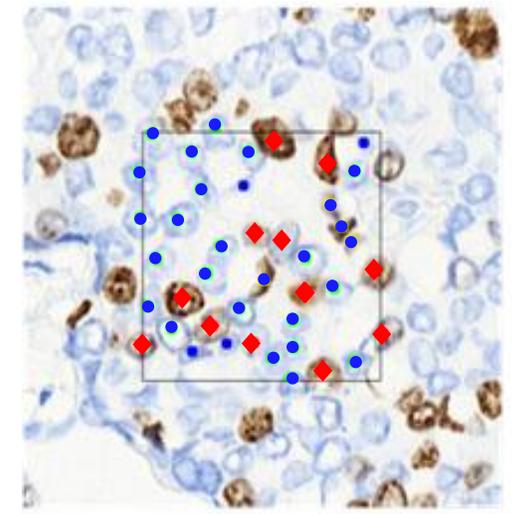
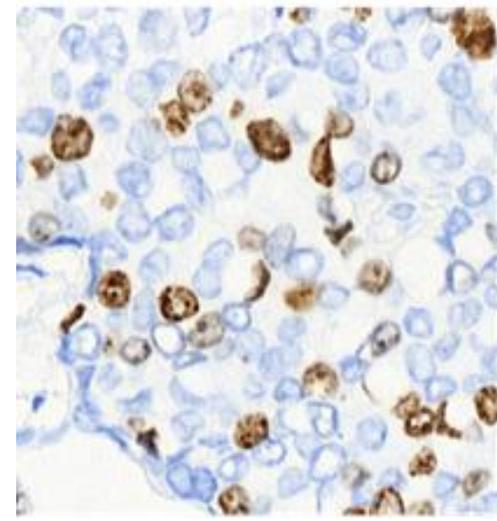
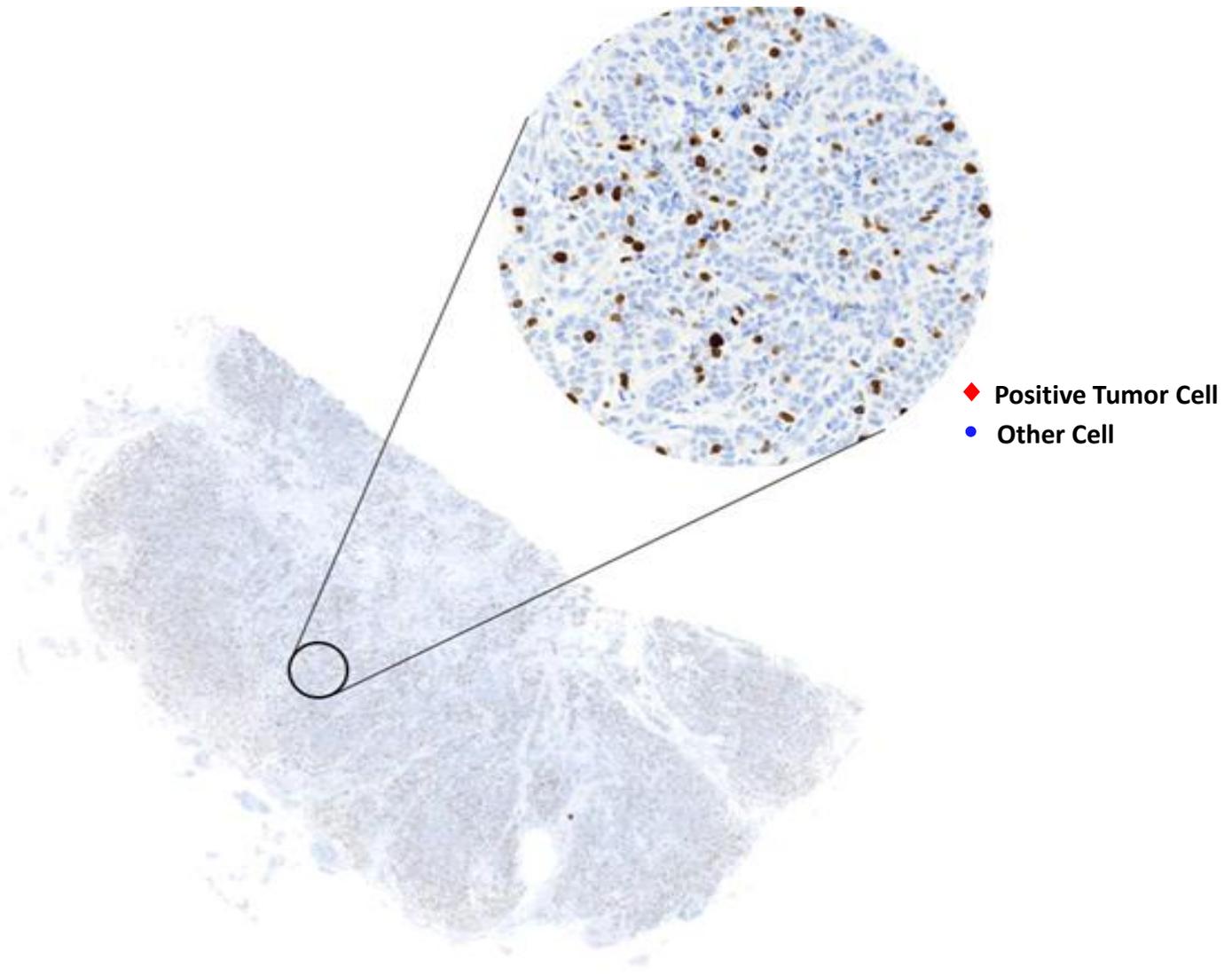
Example Independent validation study of PAIGE Prostate: Kanan, Christopher, Jillian Sue, Leo Grady, Thomas J. Fuchs, Sarat Chandarlapaty, Jorge S. Reis-Filho, Paulo G O Salles, Leonard Medeiros da Silva, Carlos Gil Ferreira, and Emilio Marcelo Pereira. "Independent Validation of Paige Prostate: Assessing Clinical Benefit of an Artificial Intelligence Tool within a Digital Diagnostic Pathology Laboratory Workflow." *Journal of Clinical Oncology* 38, no. 15_suppl (May 20, 2020): e14076–e14076. https://doi.org/10.1200/JCO.2020.38.15_suppl.e14076.

Why is applying AI to pathology a good idea?



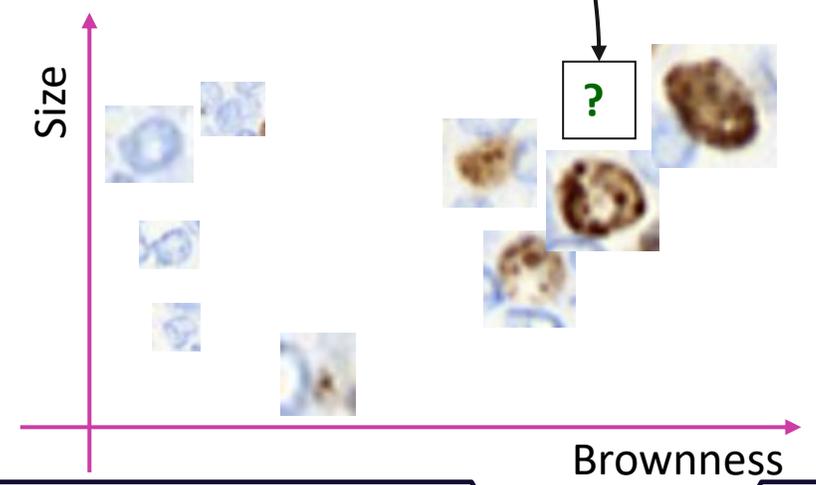
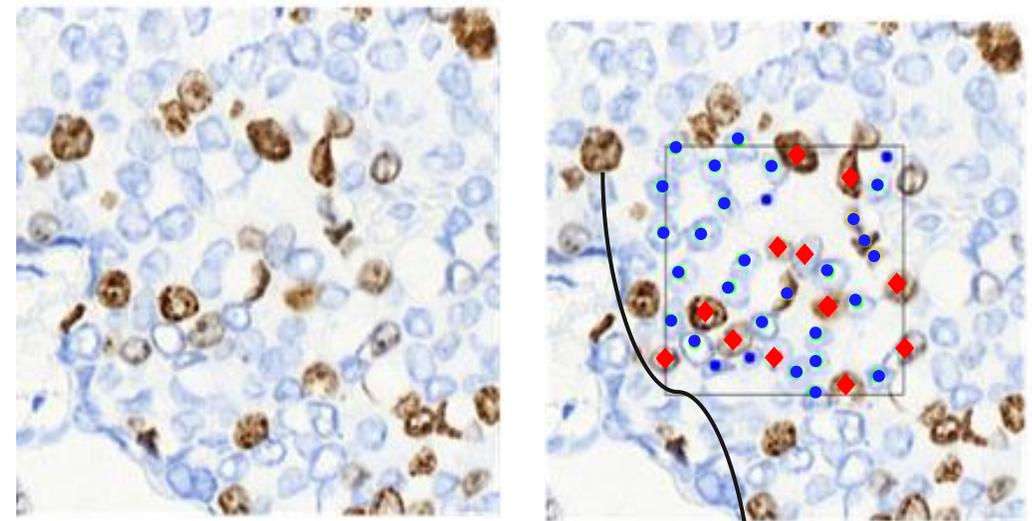
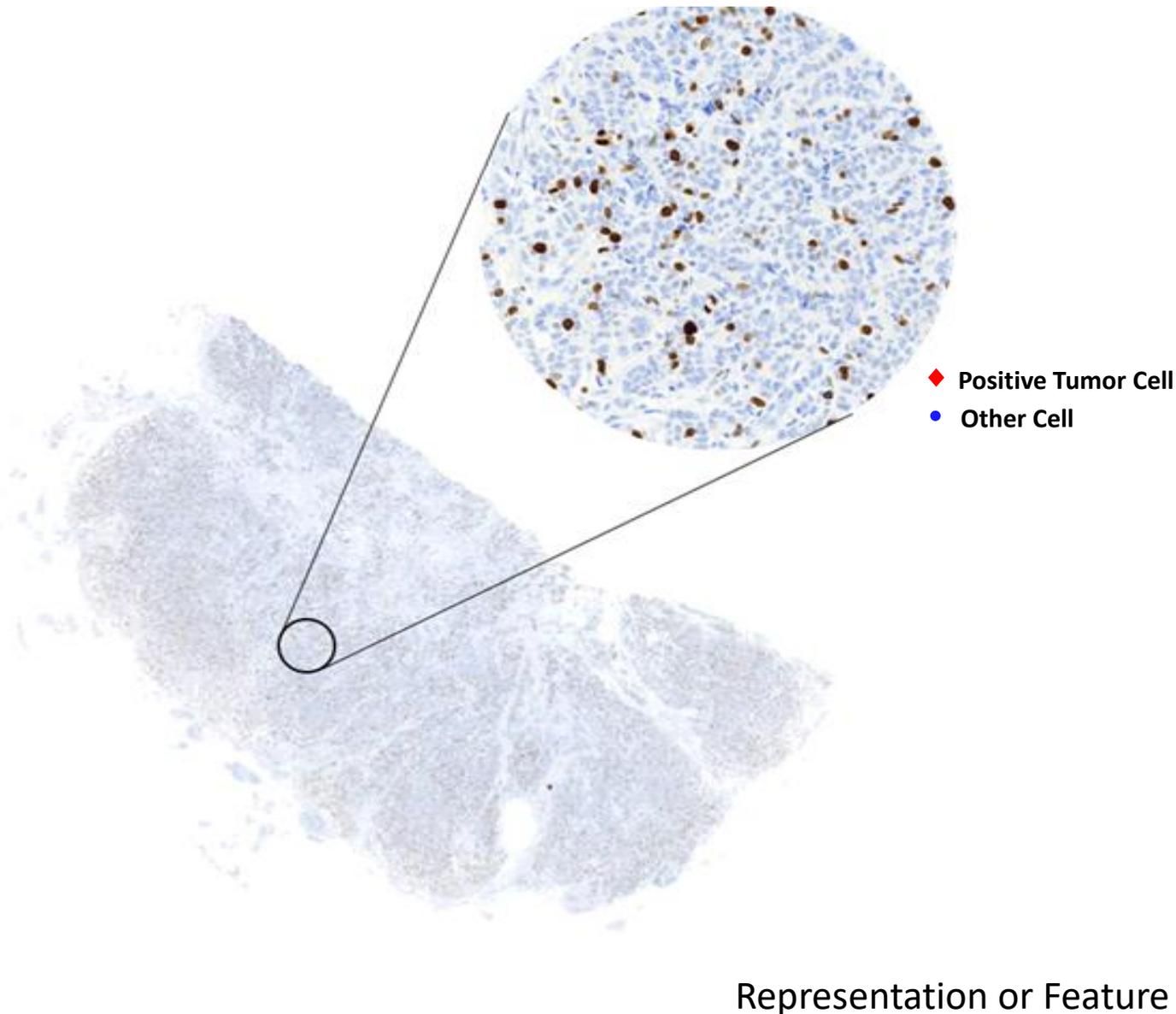
What percentage of cells in this sample are fat cells?

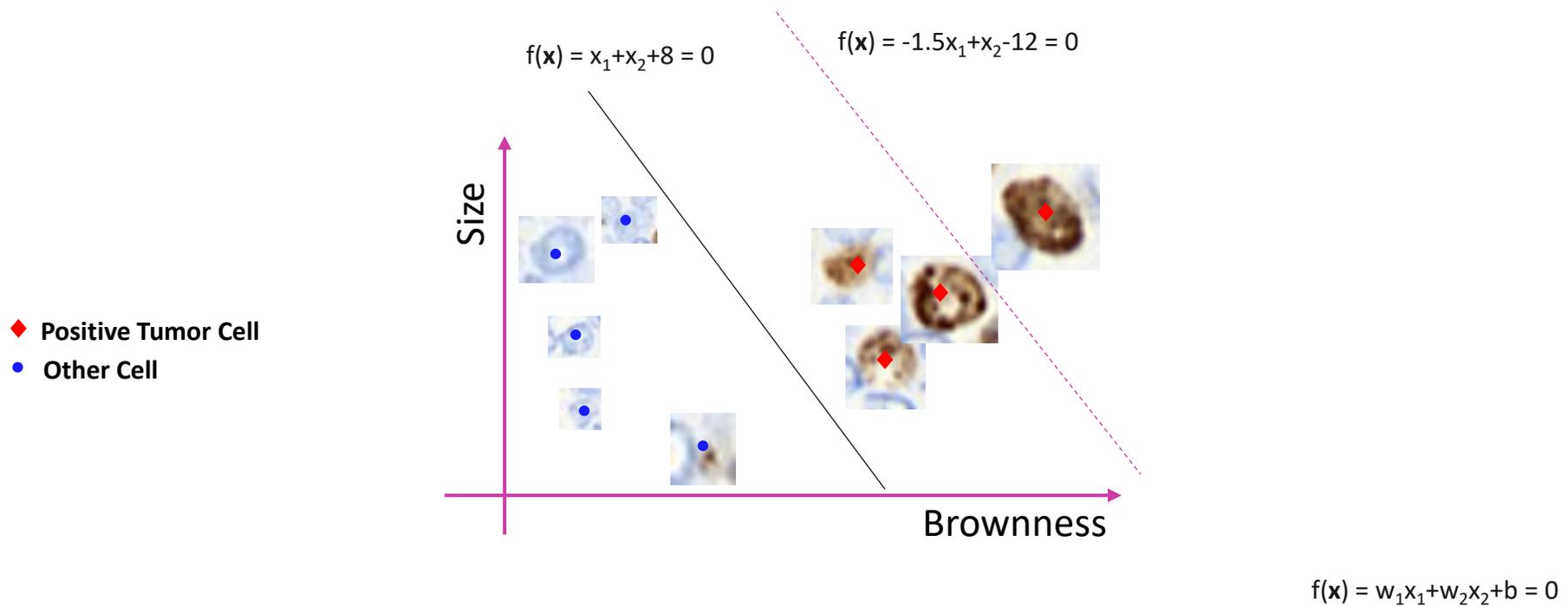
How does ML work?



Representation or Feature Space

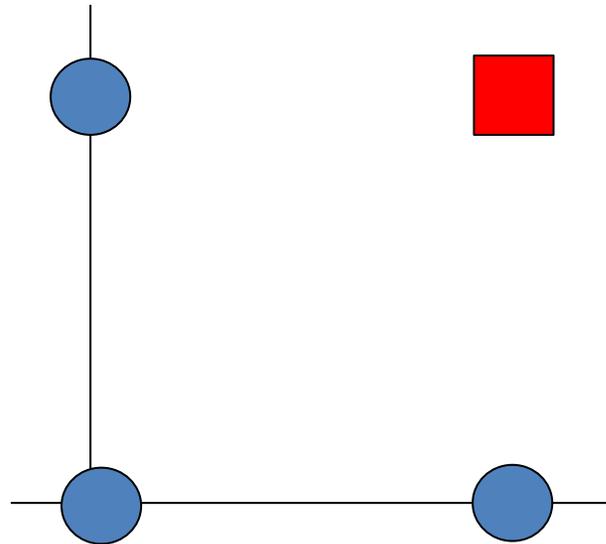
How does ML work?



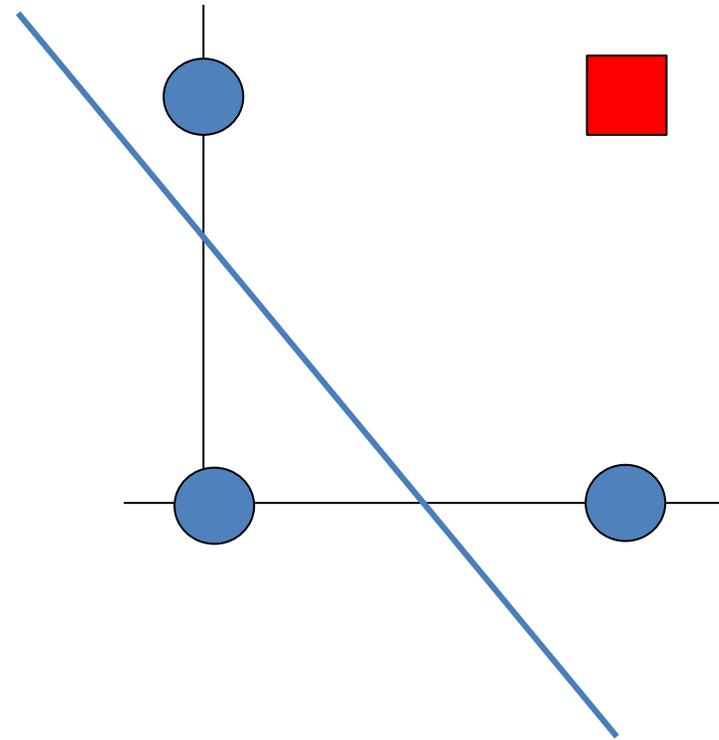


Goal is to be able to generalize to unseen data
 With Deep Learning – we don't even need to define features!

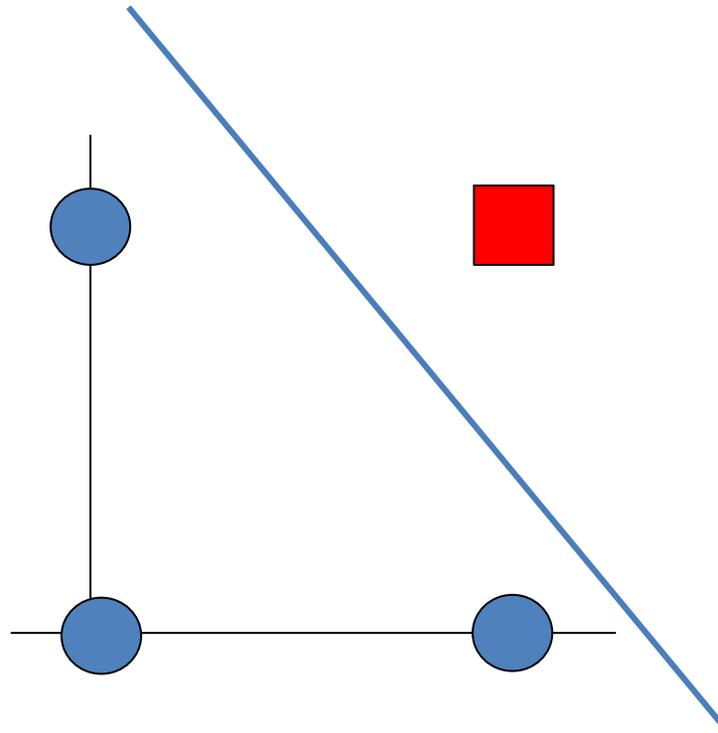
Exercise



Exercise



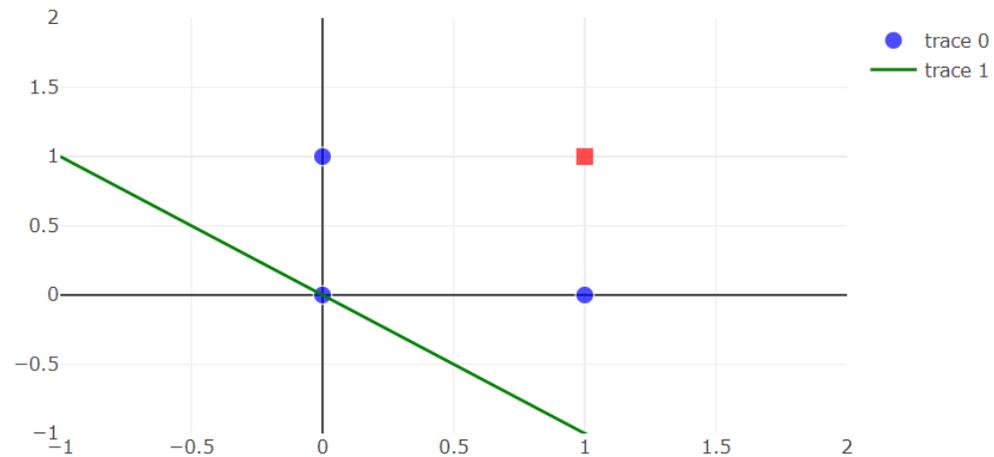
Exercise



Doing it interactively

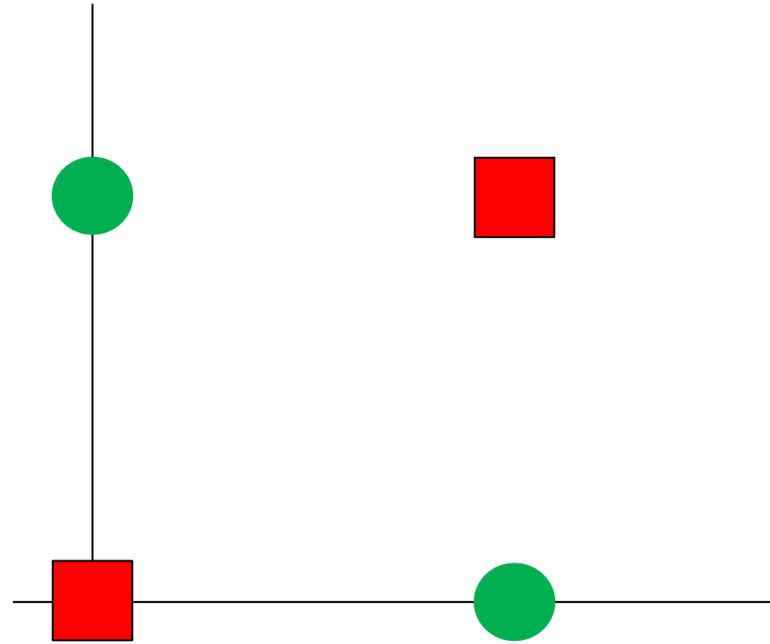
w1: 1
w2: 1
b: 0

AND Classification Problem



<https://foxtrotmike.github.io/CS909/AND-NEURON.html>

But what if we have a linearly inseparable problem?



Machine learning and deep learning involve discovering meaningful representations of input data and then using these representations to partition data for various tasks

Staffing crisis in pathology

SIR,—We are writing to express concern that the methods of teaching pathology to undergraduate medical students are both qualitatively and in outlook inappropriate for current demands. Pathology is the study of disease and as such is the basis of all medical practice. It is the one subject which encourages the strict discipline of the scientific method together with the empirical art of medicine to explain the clinically observed patterns of disease behaviour. There is an increasingly scientific approach to medicine which needs to be reflected in and emphasised by expansion of undergraduate teaching of pathology. The size of undergraduate classes receiving instruction in pathology is generally increasing and there is a serious lack of first-hand experience of practical pathology. At a debate during the recent Pathological Society meeting (9 January) many of the senior academics admitted that large classes were exerting an increasingly detrimental effect on the teaching of pathology. As a direct consequence pathology teaching is in a decline; and the subject is becoming, apparently, more remote from the clinical practice of medicine.

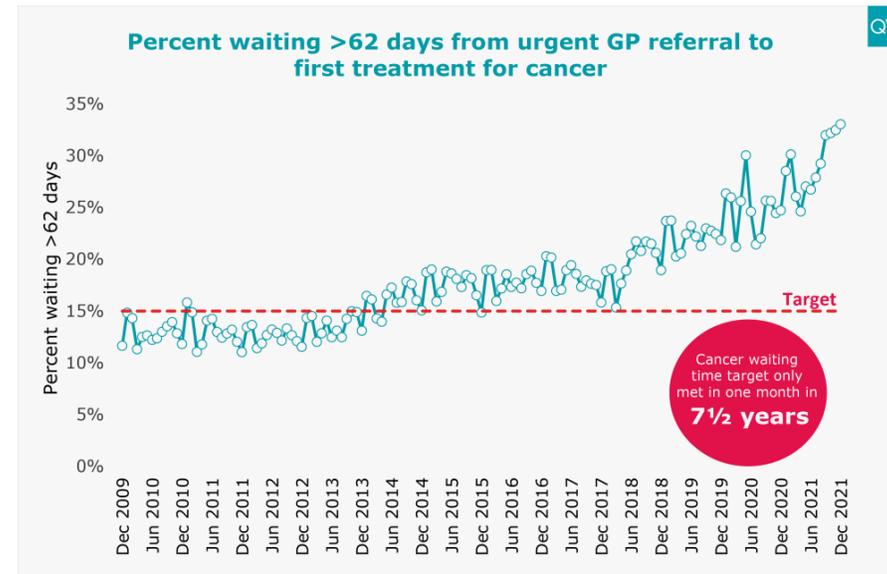
We are concerned that the input to pathology training posts should be not only quantitatively but also qualitatively optimal to meet future demands. We strongly support Professors J R Anderson and J R Tighe (15 November, p 1370), who have pointed out that, although the immediate problem in pathology is a scarcity of suitably trained individuals for the large number of consultant posts shortly to become vacant through retirement, the problem in the future will become compounded by the shortage of SHO and registrar posts to feed the senior registrar grade. However, further factors should be considered and include that undefined proportion of those entering the junior grades who will not complete their higher specialist training in pathology. We also feel that it is extremely important that sufficient posts should be available within pathology departments to provide opportunities for those primarily working in other specialties to benefit from a first-hand experience of pathology. Therefore we further emphasise the urgent need for an increase in posts at the SHO and registrar grade. It is important, however, to avoid focusing on this single measure as being the answer to the future security of pathology. It is evident to us, from personal communications with the directors of pathology in many of the undergraduate teaching departments, that there is widespread concern to provide undergraduates with a proper exposure to pathology. We believe that medical students should have the opportunity to receive practically based tutorials in all branches of pathology. Pathology should be taught as being directly relevant to the clinical situation. In particular, we believe that the general principles of pathology should be introduced from the earliest stage of the undergraduate medical training. In addition, the teaching of pathology should be combined with the teaching of anatomy, histology, physiology, and biochemistry. Symington has stressed the

Why is applying ML to pathology a good idea? Business Need

- Pathologist Recruitment
- Aging workforce
- Subjectivity
- Time & Effort in Quantification

"... workforce census from the Royal College of Pathologists showed that a quarter of all histopathologists are over 55, most of whom are expected to retire by 2023. Furthermore, an all-time low number of trainee doctors are choosing to specialize in pathology with only 3 percent of NHS histopathology departments having enough staff to meet clinical demand."

<https://thepathologist.com/inside-the-lab/beyond-digital>

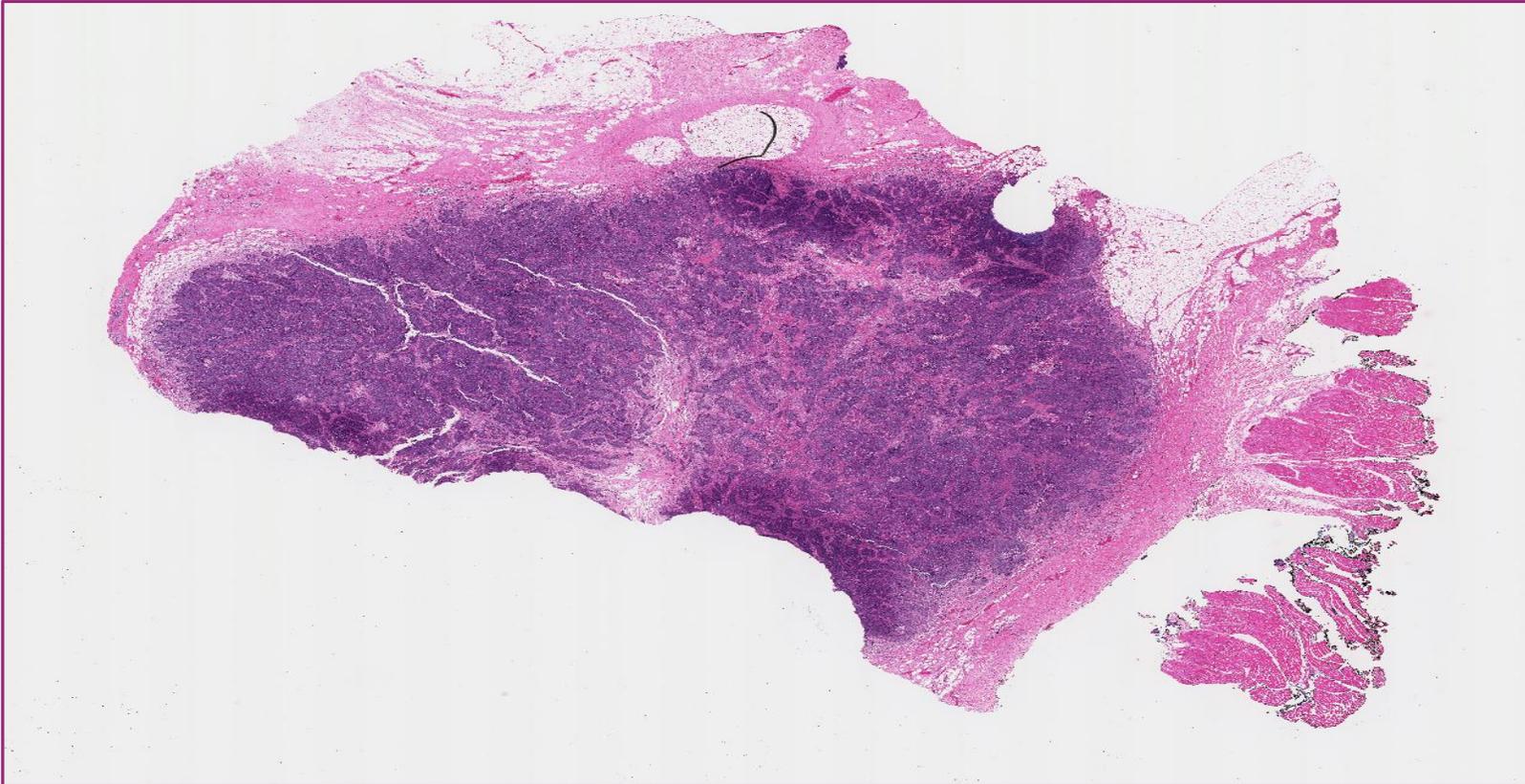


**HISTOPATHOLOGY
WORKFORCE SURVEY**



The Royal College
Pathology: the science

Why is applying ML to pathology a good idea?



Given training data,
Can we predict:

Normal or Abnormal?

Classification

Grade?

Regression

Tumour Regions?

Segmentation

Quantifying cells

Quantification

Receptor Status?

Classification

Mutation Status?

Classification

Survival?

Drug Response?

Similar images?

Retrieval

Biomarkers?

Pattern Discovery

Spectrum of Mundanity



IHC scoring
Finding tumor
Mitotic counting



Outcome based subtyping
Response to therapy



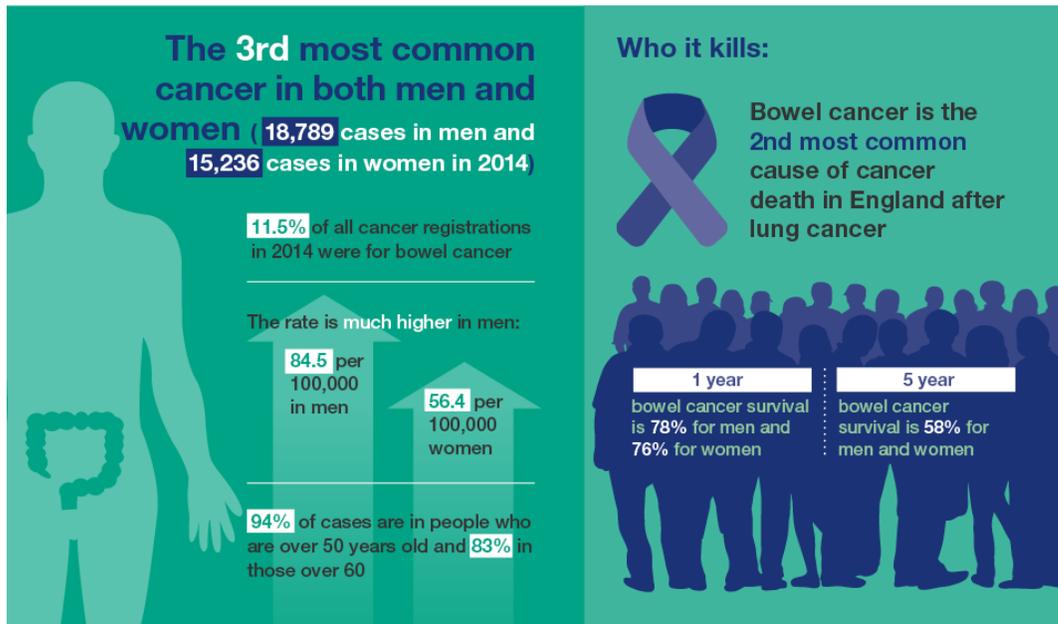
Mutation status
MSI positive
HPV positive

A case study

The problem

- Colorectal Cancer

Who it affects



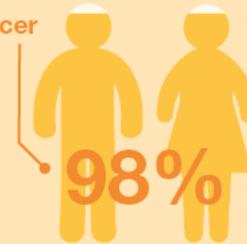
Early detection of bowel cancer

Diagnosis

The earlier a cancer is diagnosed the better the chances of survival:

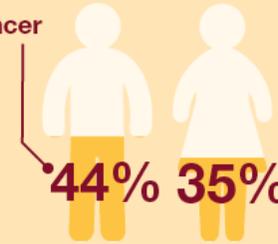
Stage 1 cancer

1 year survival in 2014



Stage 4 cancer

1 year survival in 2014



cancer is **more likely** to be detected at an early stage via screening or via a two week wait GP referral

Routes to diagnosis in 2013

9%

of bowel cancers were diagnosed through screening

23%

through GP referral

30%

through two week wait

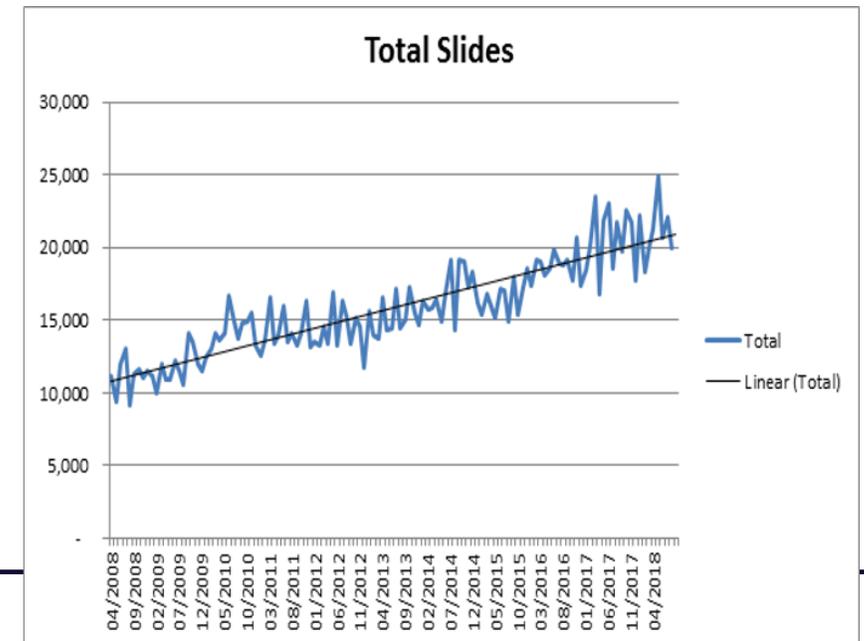
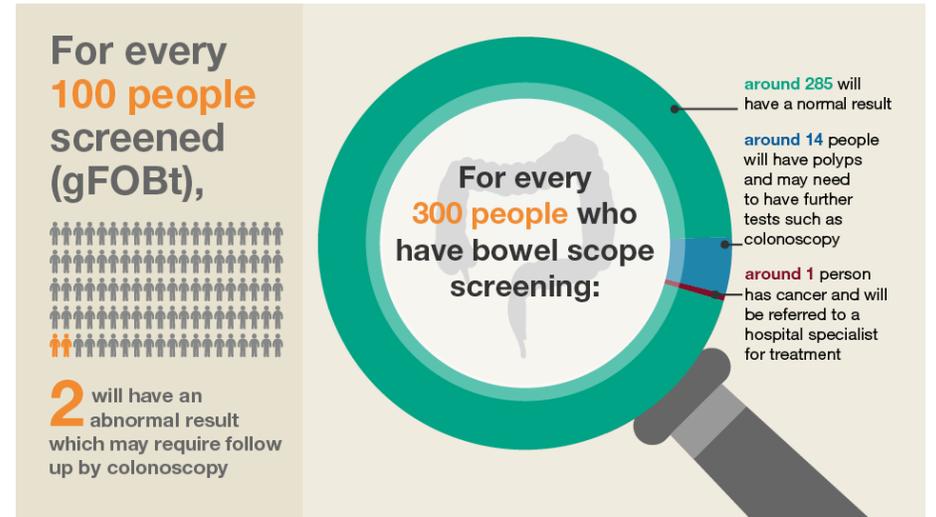
24%

by emergency presentation

Bowel Cancer Screening

- In England, screening is offered every 2 years for men and women aged 60 to 74 (or above on request)
- gFOBt test (replaced by Faecal Immunochemical Test (FIT))
- Followed by Colonoscopy
- Pathology Examination
 - Endoscopic biopsies from the large bowel account for 8% of the requests in NHS cellular pathology laboratories.
 - Of these samples, approximately 40% are reported as normal

What happens after screening?



PathLAKE CoBI Project

- WIBGI
 - *We could “screen out” or “rule out” abnormalities using AI based computational pathology algorithms using whole slide images as input*
- What do we need for this?
 - Training and validation data with labels
 - Develop algorithms



Special thanks to all the funders, researchers and collaborators for support and involvement in the work presented on these slides



CANCER RESEARCH UK



EPSRC
Engineering and Physical Sciences Research Council

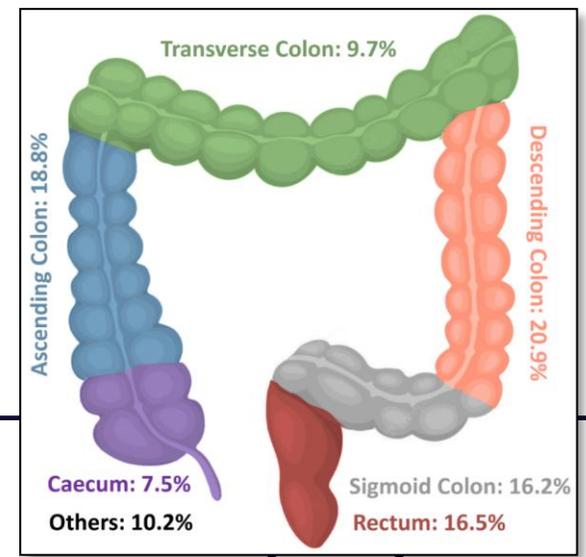
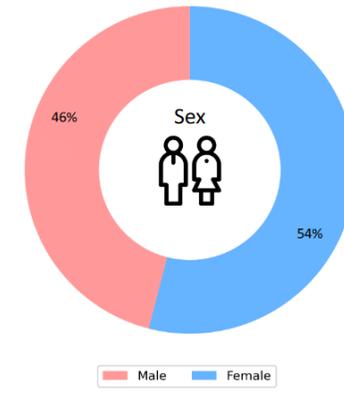
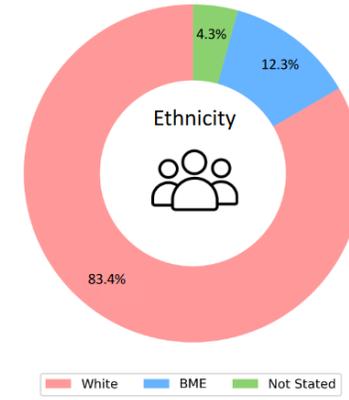
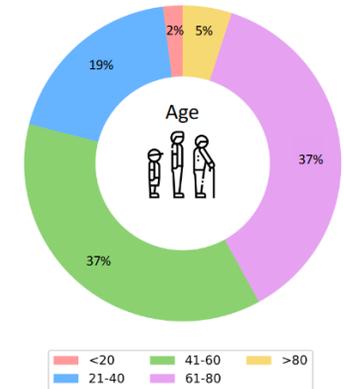
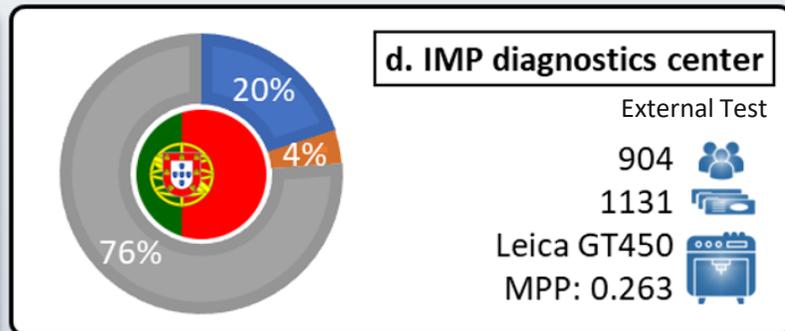
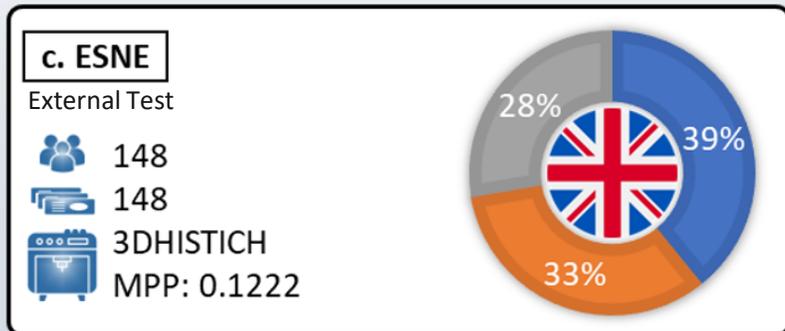
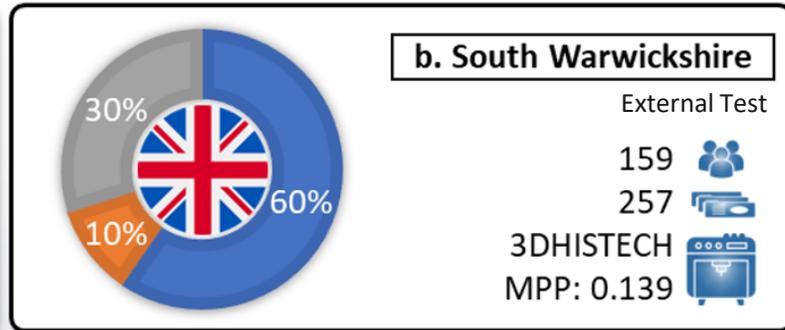
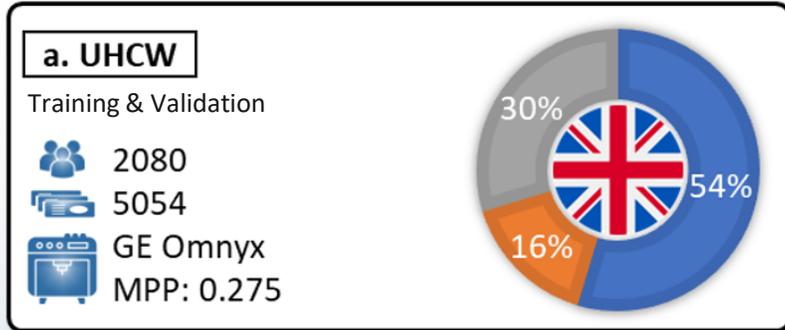


Innovate UK



Data

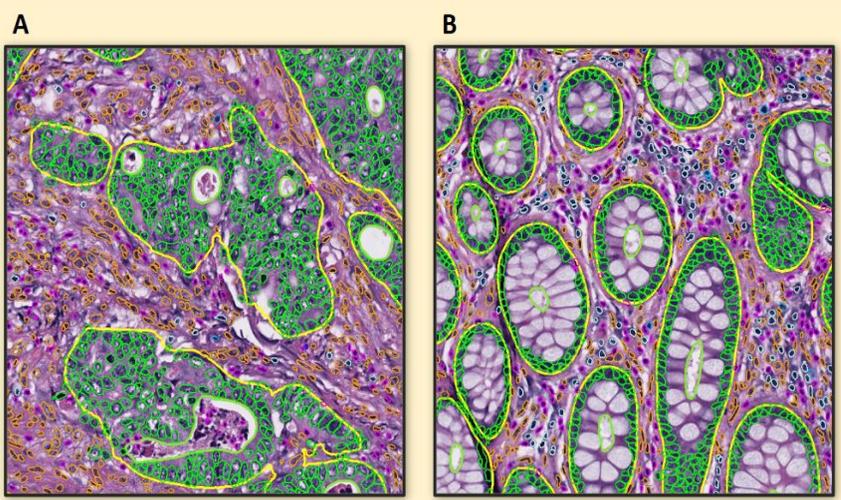
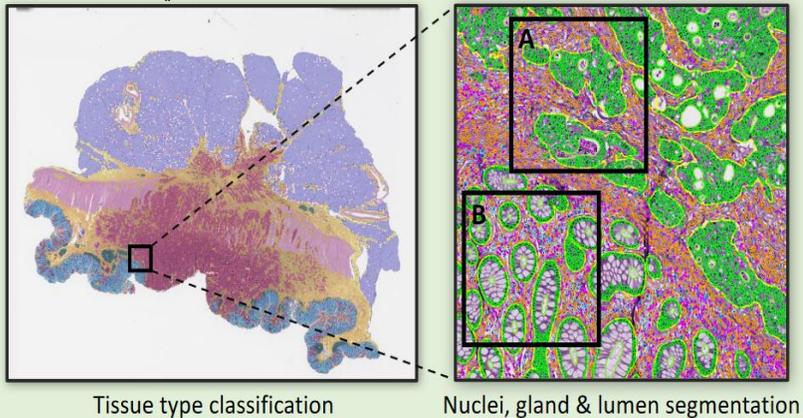
● Normal
 ● Non-neoplastic
 ● Neoplastic
 Number of cases
 Number of Slides
 Scanner details



UHCW used as primary development set colorectal (colonoscopy) biopsies were audited over a 5-year period from 2012 to 2017, reviewed by 13 pathologists with diagnostic consensus

Two Flavours of CPath

Cerberus 



Bottom-up detection & characterization of histological primitives

Pros and Cons (for Top-Down Style)

Lesser annotations needed
Multi-modality integration
Versatility of scientific questions

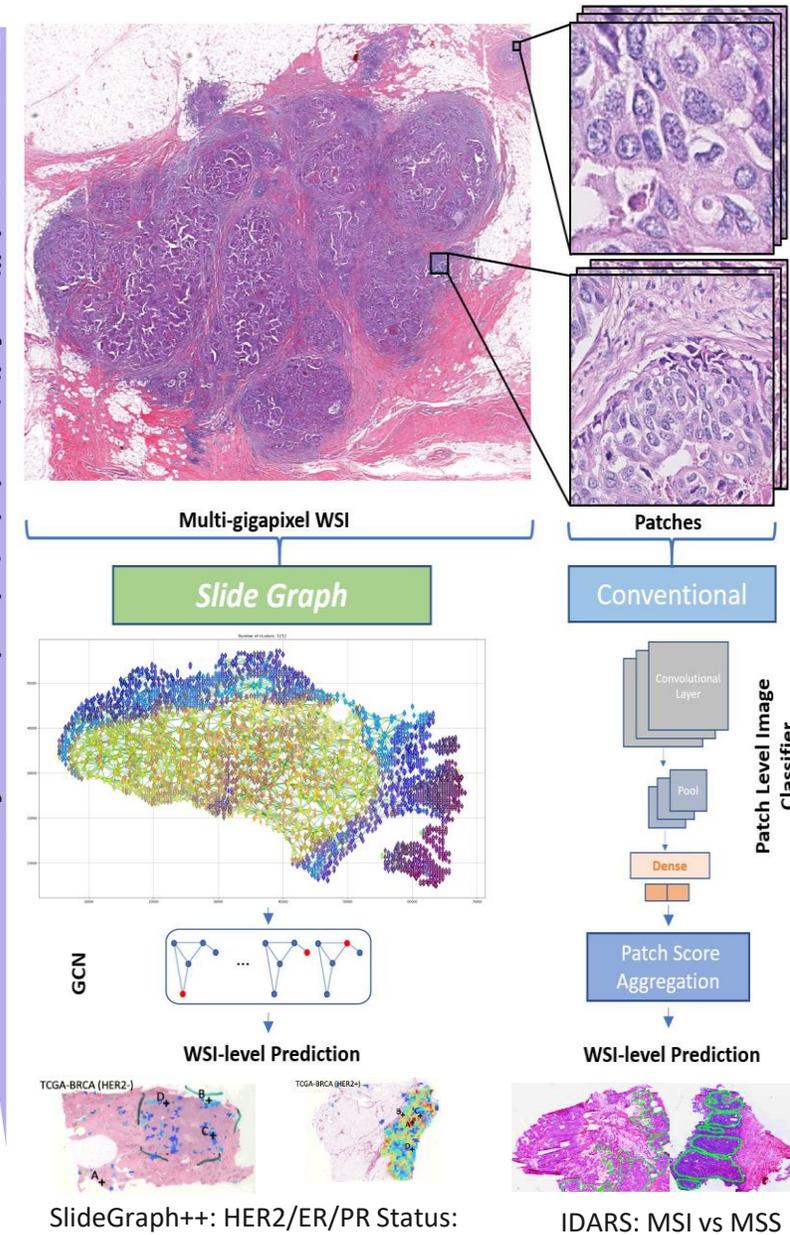
Large(r) amount of WSIs
Poor interpretability & explainability
Robustness
Impact of confounders



- Feature Mining
- Segmenting histological structures
- Detecting nuclei and Classifying cells
- Getting annotations (nuclear, regional)



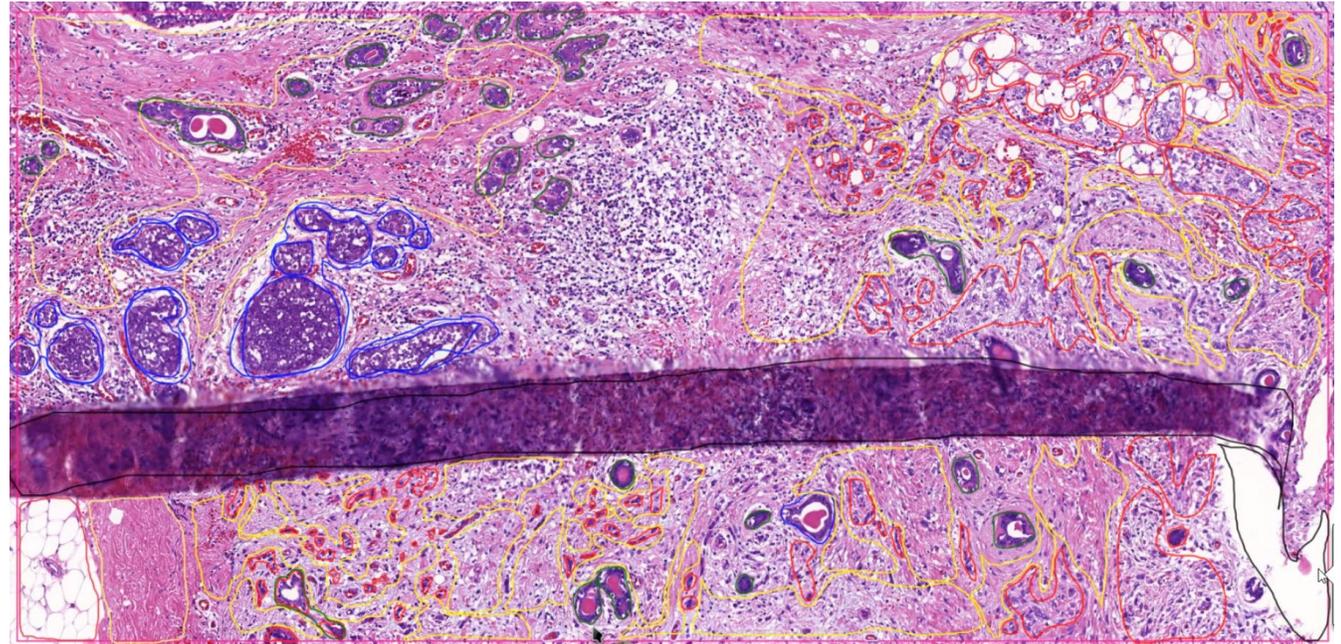
Top-Down Modelling of clinicopathological and response factors



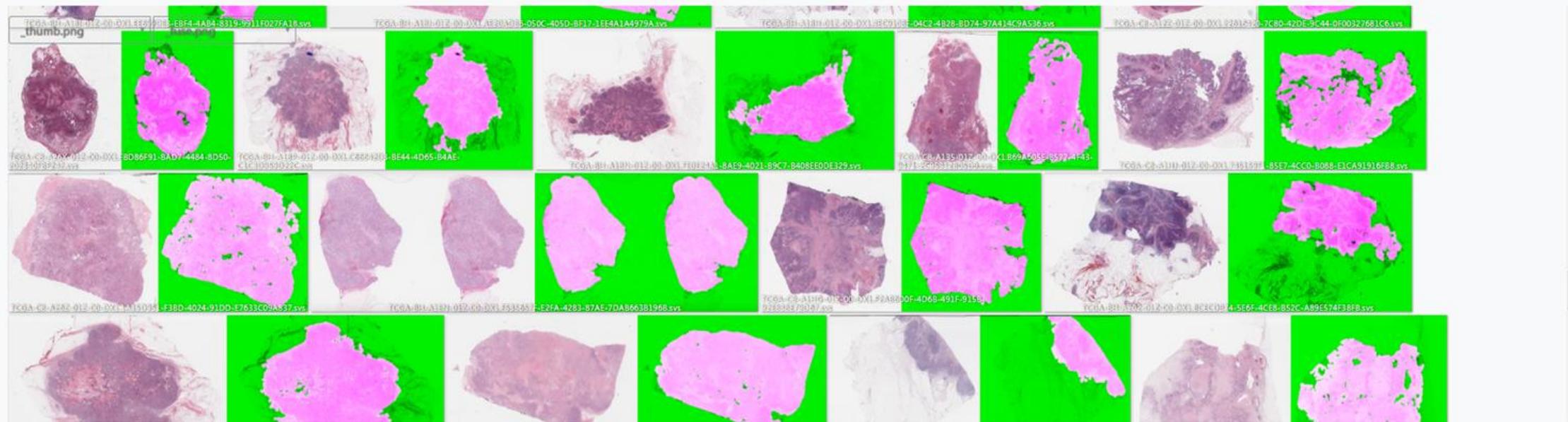
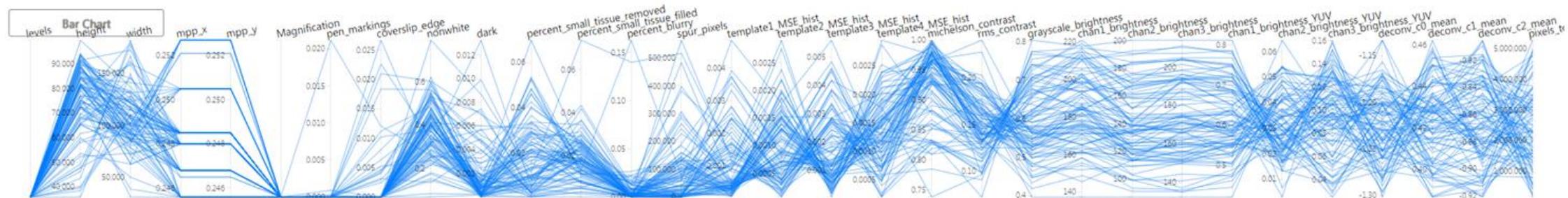
Graham, Simon, Quoc Dang Vu, Mostafa Jahanifar, Shan E. Ahmed Raza, Fayyaz Minhas, David Snead, and Nasir Rajpoot. "One Model Is All You Need: Multi-Task Learning Enables Simultaneous Histology Image Segmentation and Classification." *Medical Image Analysis*, November 11, 2022, 102685. <https://doi.org/10.1016/j.media.2022.102685>.
 Pocock, Johnathan, Simon Graham, Quoc Dang Vu, Mostafa Jahanifar, Sriyaz Deshpande, Giorgos Hadjigeorgiou, Adam Shephard, et al. "TIAToolbox as an End-to-End Library for Advanced Tissue Image Analytics." *Communications Medicine* 2, no. 1 (September 24, 2022): 1–14. <https://doi.org/10.1038/s43856-022-00186-5>.
 Lu, Wengqi, Michael Toss, Emad Rakha, Nasir Rajpoot, and Fayyaz Minhas. "SlideGraph+: Whole Slide Image Level Graphs to Predict HER2status in Breast Cancer." *ArXiv:2110.06042 [Cs]*, October 12, 2021. <http://arxiv.org/abs/2110.06042>.
 Bilal, Mohsin, Shan E. Ahmed Raza, Ayesha Azam, Simon Graham, Mohammad Ilyas, Ian A. Cree, David Snead, Fayyaz Minhas, and Nasir M. Rajpoot. "Development and Validation of a Weakly Supervised Deep Learning Framework to Predict the Status of Molecular Pathways and Key Mutations in Colorectal Cancer from Routine Histology Images: A Retrospective Study." *The Lancet Digital Health* 3, no. 12 (December 1, 2021): e763–72. [https://doi.org/10.1016/S2589-7500\(21\)00180-1](https://doi.org/10.1016/S2589-7500(21)00180-1).

Pre-processing and Image QC

- Stain variation normalization
- Image quality control
 - Blur detection
 - Loss of cover slip
 - Tissue folding
 - Pen Marking



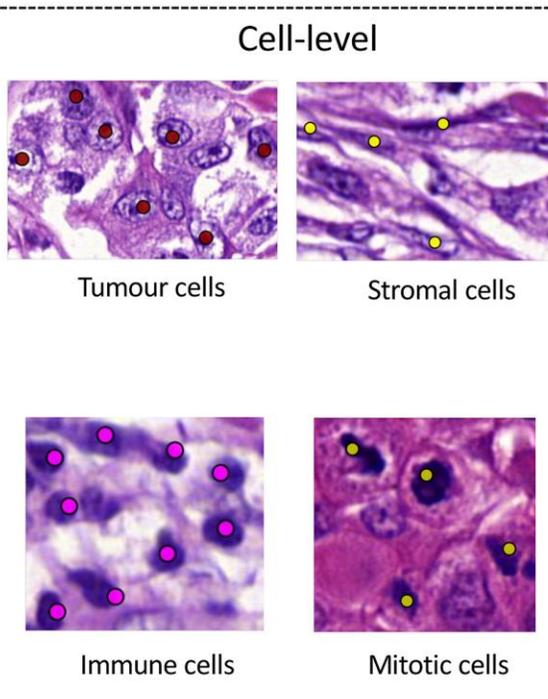
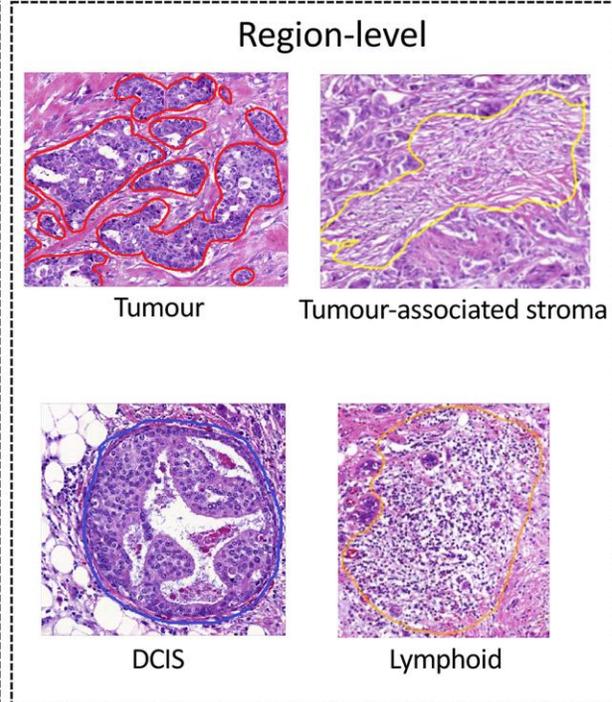
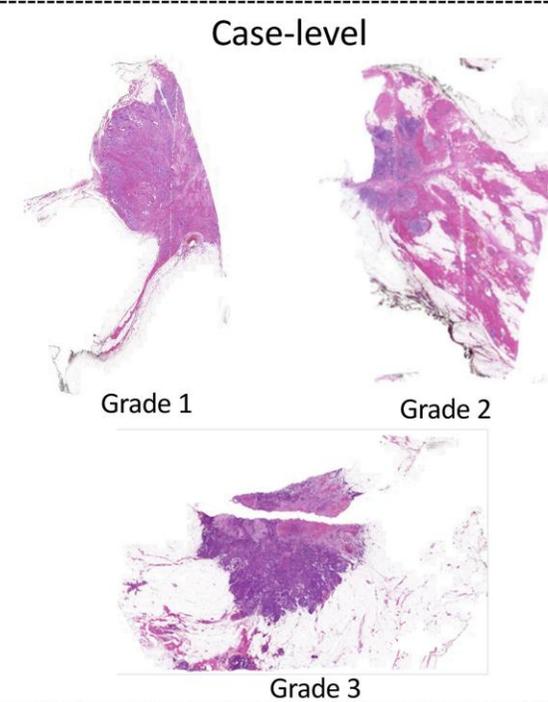
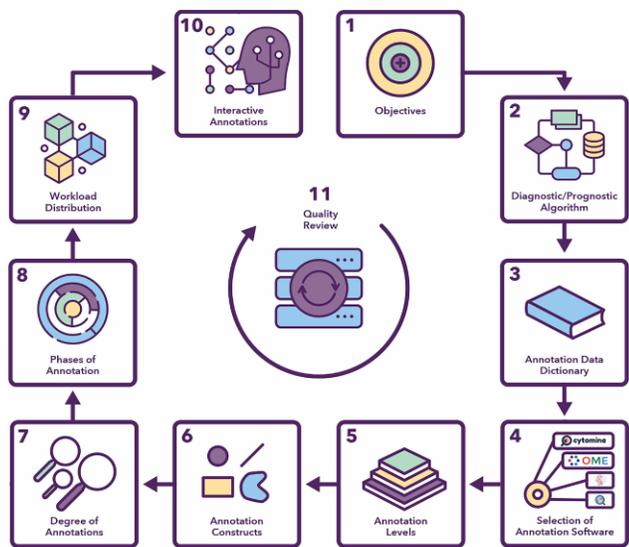
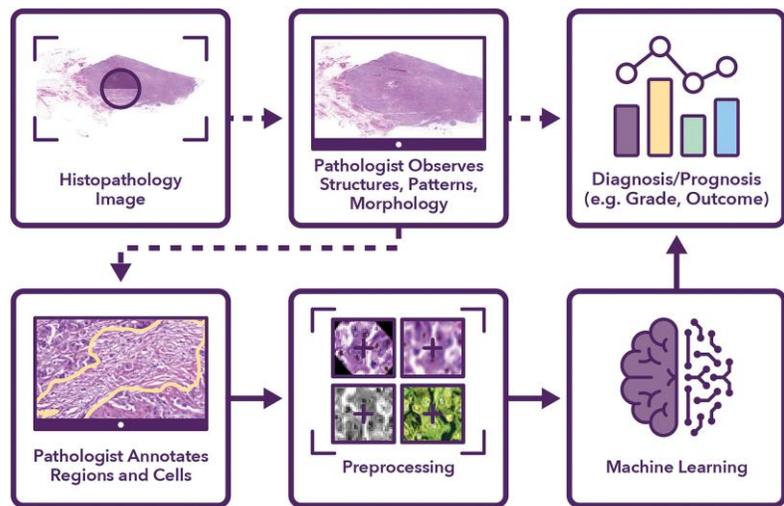
filename	levels	height	width	mpp_x	mpp_y	Magnification	pen_markings	coverslip_edge	nonwhite	dark	percent	Copy
TCGA-D8-A1JJ-01Z-00-DX2.E688E270-26C9-43EC-BC26-86F8B74A31D.svs	4	61772	101591	0.2527	0.2527	40	0	0.010239647505201746	0.37570282730202764	0.0000585859410060789	0.02	Save
TCGA-D8-A1JJ-01Z-00-DX1.a986b48f-b295-4d7a-b778-ce829cdf9c38.svs	4	89863	84727	0.248	0.248	40	0	0	0.40365625618646767	0.00038530934490148993	0.02	Delete
TCGA-D8-A1JJ-01Z-00-DX2.7D20F308-7DC6-4367-9459-3AC4C654E7F7.svs	4	74635	83663	0.2527	0.2527	40	0	0.02097230770845992	0.6022809295769842	0.0069632641758784015	0.02	Deselect



Janowczyk, Andrew, Ren Zuo, Hannah Gilmore, Michael Feldman, and Anant Madabhushi. "HistoQC: An Open-Source Quality Control Tool for Digital Pathology Slides." *JCO Clinical Cancer Informatics* 3 (April 16, 2019).

<https://doi.org/10.1200/CCI.18.00157>.

Semantic Annotation Guidelines



Descriptive and multi-modal Breast core biopsy report

Patient name:	Physician name:
Date of birth:	Clinic:
Record number:	Contact:
Specimen number:	
Date obtained:	
Date processed:	

History: 62-year-old female with a palpable breast mass and suspicious calcification on mammogram.

Body site and procedure: Left breast at 9:00, 5 cm from nipple (ultrasound-guided core needle biopsy).

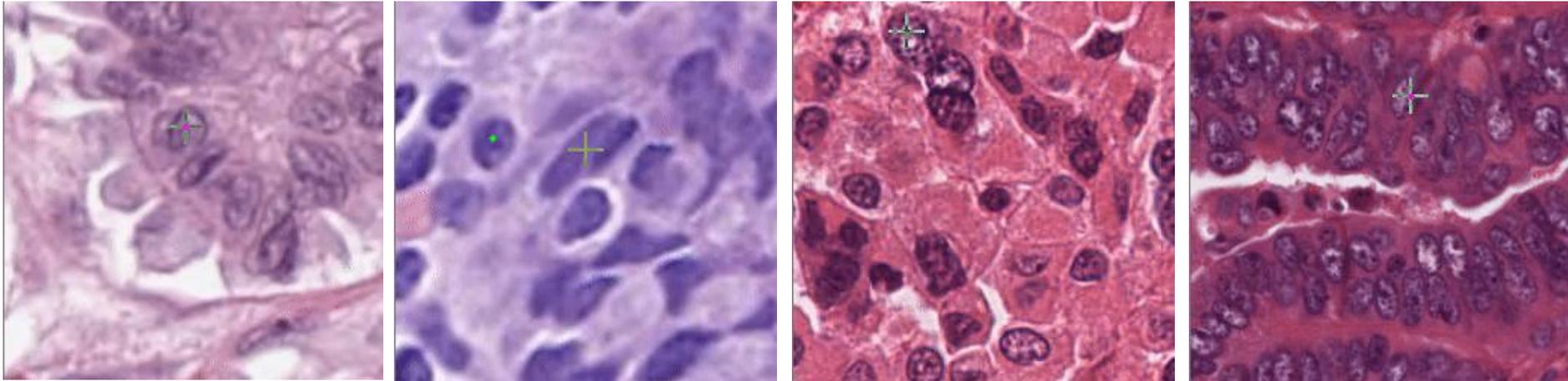
Description: Specimen is received in formalin, labelled with patient's name, and consists of 2 cores of tissue measuring 0.4 cm in width.

Diagnosis: Ductal carcinoma, Nottingham histologic grade 2. The carcinoma is positive for ER (80%) and is negative for HER-2 (IHC score 0).

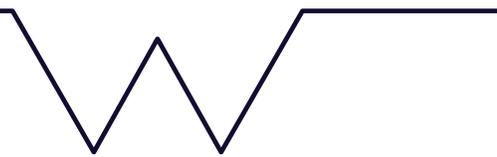
Pathologist:
Contact number:

AI-Assisted Annotations

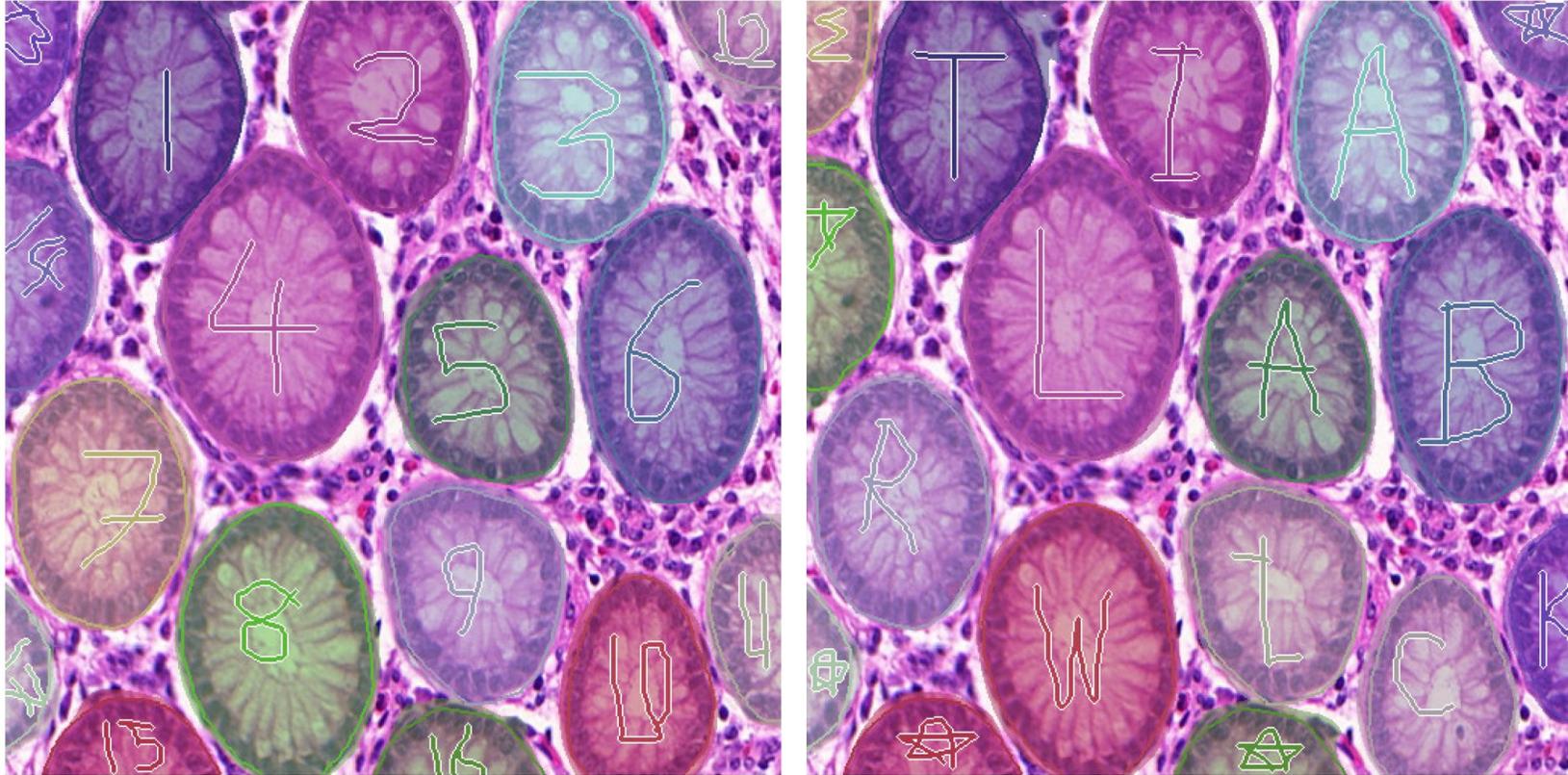
NuClick used for collecting the segmentation masks from point clicks



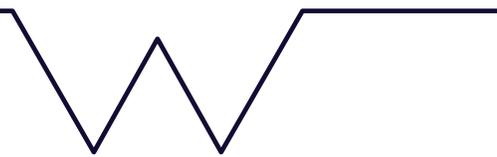
Koohbanani *et al.*, **Medical Image Analysis** (Oct 2020) & Gamper *et al.*, **Arxiv** (Apr 2020)



NuClick+: AI-Assisted Annotation of Glands



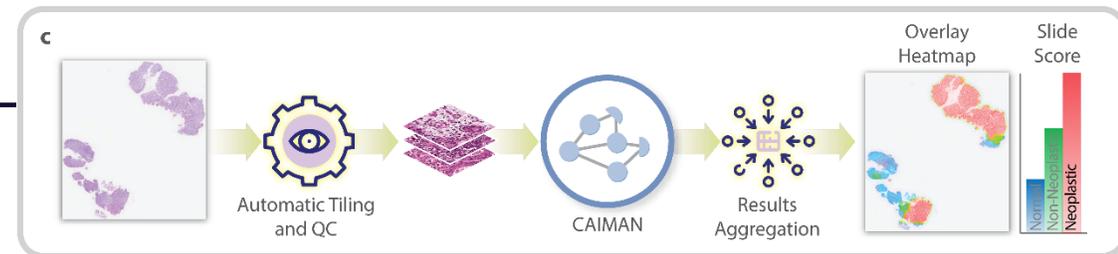
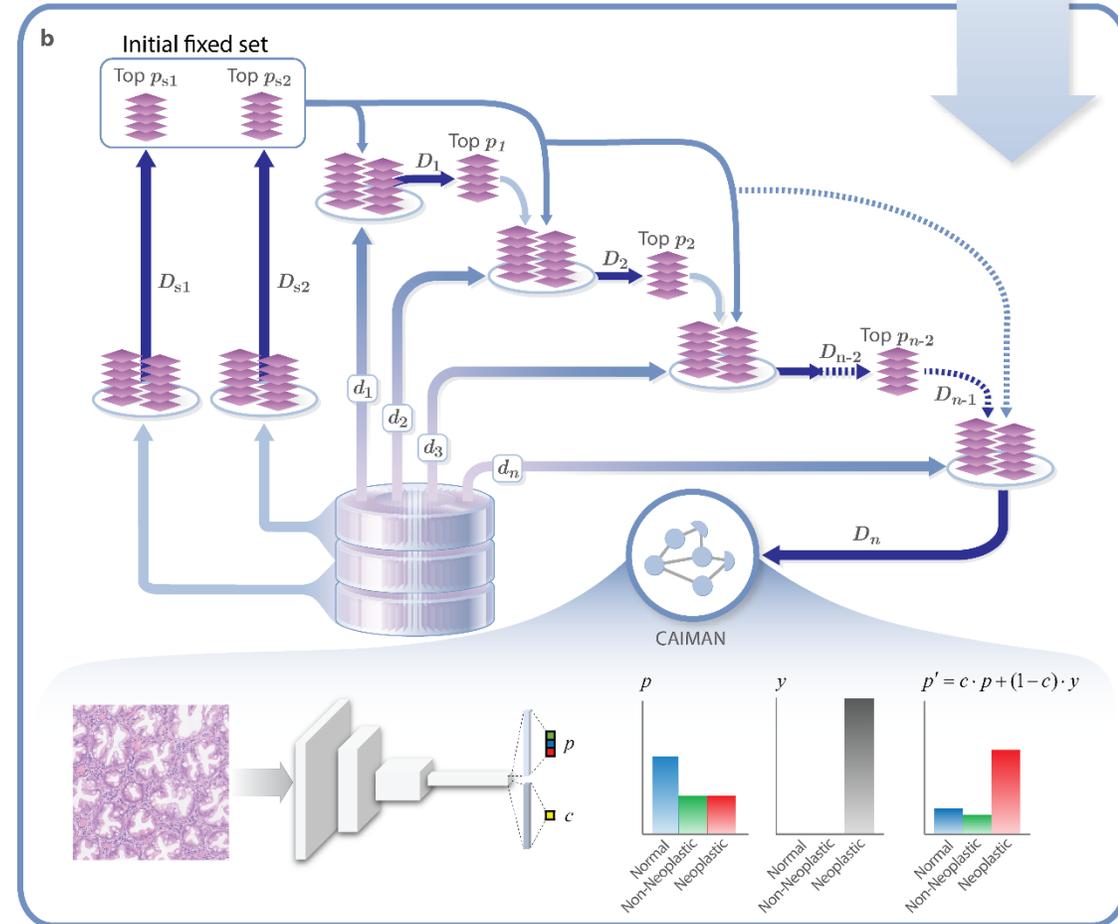
Koohbanani *et al.*, **Medical Image Analysis** (Oct 2020)



CAIMAN

Colorectal AI Model for Abnormality detection

- A weakly supervised top-down approach
 - Using whole slide image level labels for training
 - Based on our previous method for MSI detection in CRC



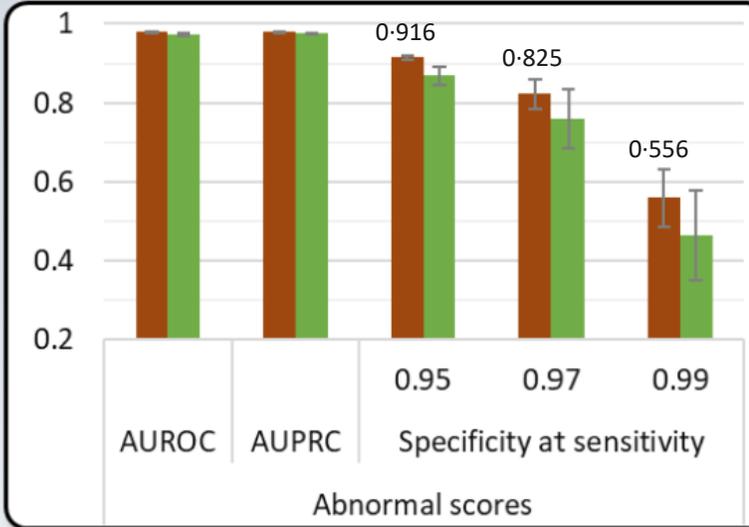
Bilal, Mohsin, Yee Wah Tsang, Mahmoud Ali, Simon Graham, Emily Hero, Noorul Wahab, Katherine Dodd, et al. "Development and Validation of Artificial Intelligence-Based Prescreening of Large-Bowel Biopsies Taken in the UK and Portugal: A Retrospective Cohort Study." *The Lancet Digital Health* 5, no. 11 (November 1, 2023): e786–97. [https://doi.org/10.1016/S2589-7500\(23\)00148-6](https://doi.org/10.1016/S2589-7500(23)00148-6).

Bilal, Mohsin, Shan E Ahmed Raza, Ayesha Azam, Simon Graham, Mohammad Ilyas, Ian A Cree, David Snead, Fayyaz Minhas, and Nasir M Rajpoot. "Development and Validation of a Weakly Supervised Deep Learning Framework to Predict the Status of Molecular Pathways and Key Mutations in Colorectal Cancer from Routine Histology Images: A Retrospective Study." *The Lancet Digital Health*, October 19, 2021. [https://doi.org/10.1016/S2589-7500\(21\)00180-1](https://doi.org/10.1016/S2589-7500(21)00180-1).

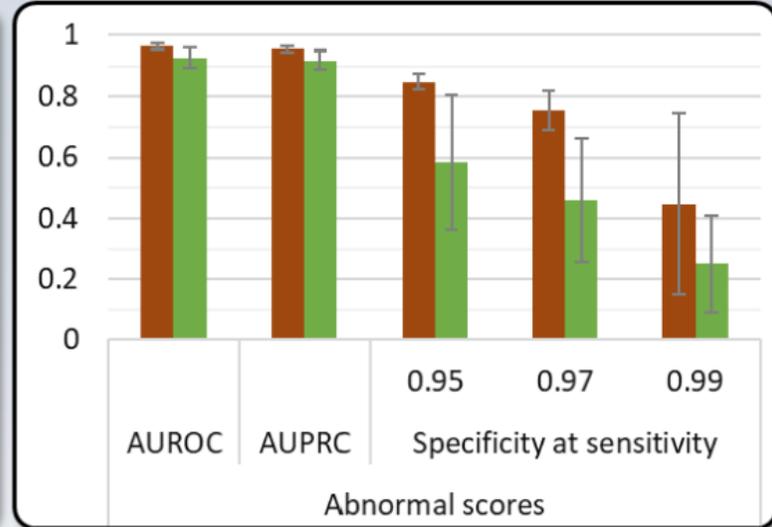
Results

- Cross-validation
- External Testing
- No statistically significant differences in prediction scores across anatomical sites

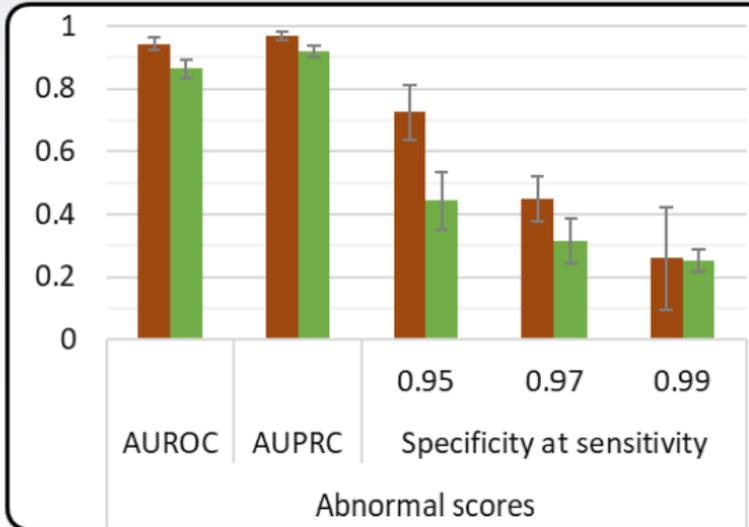
e. UHCW



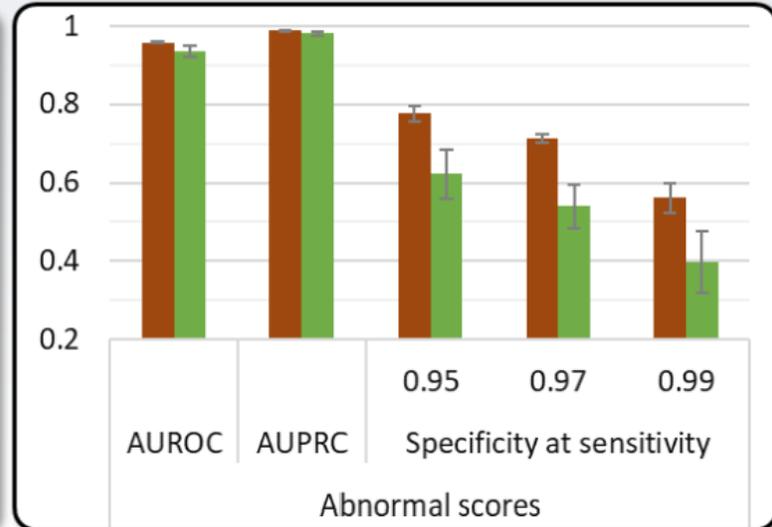
f. South Warwickshire



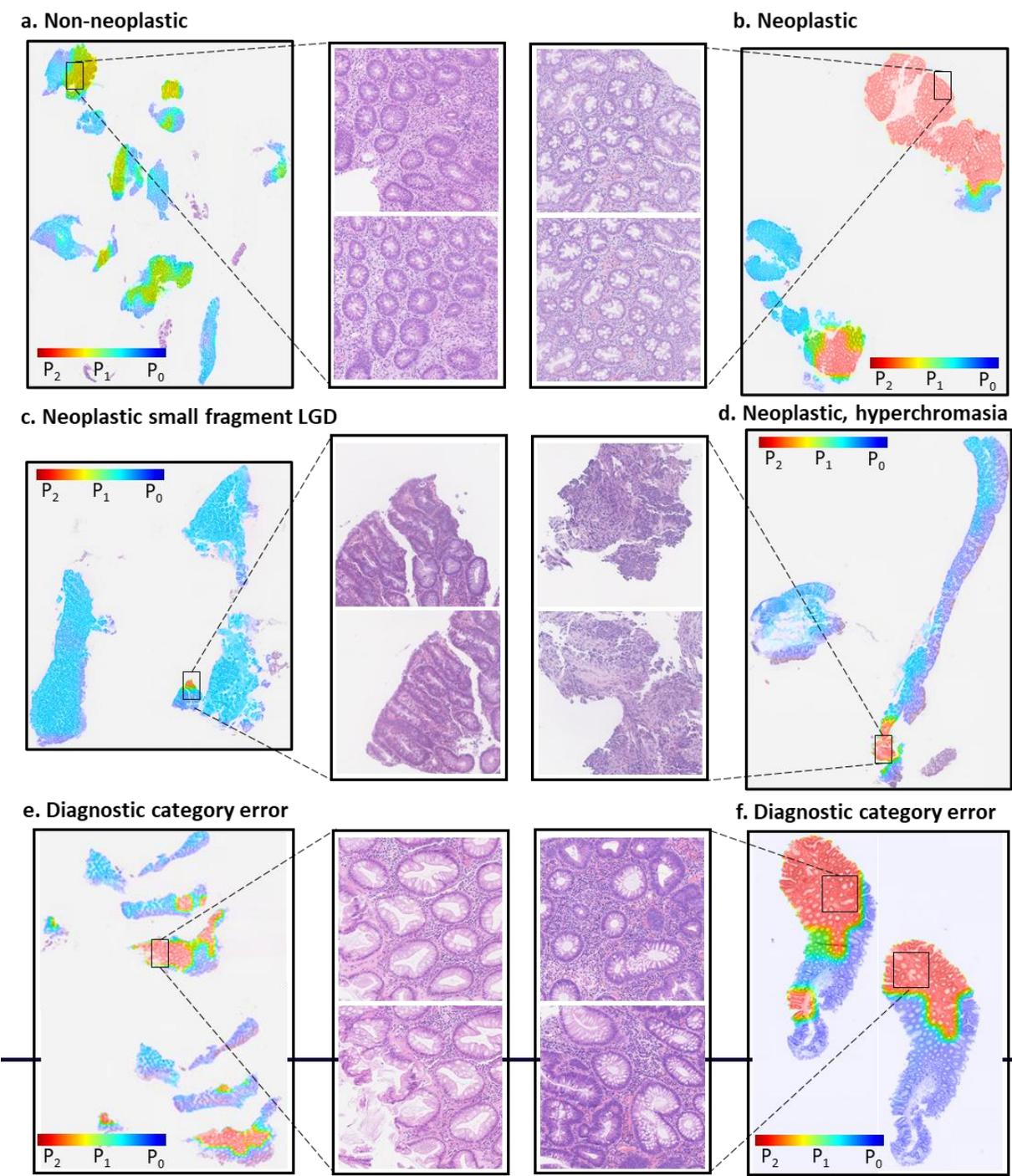
g. ESNE



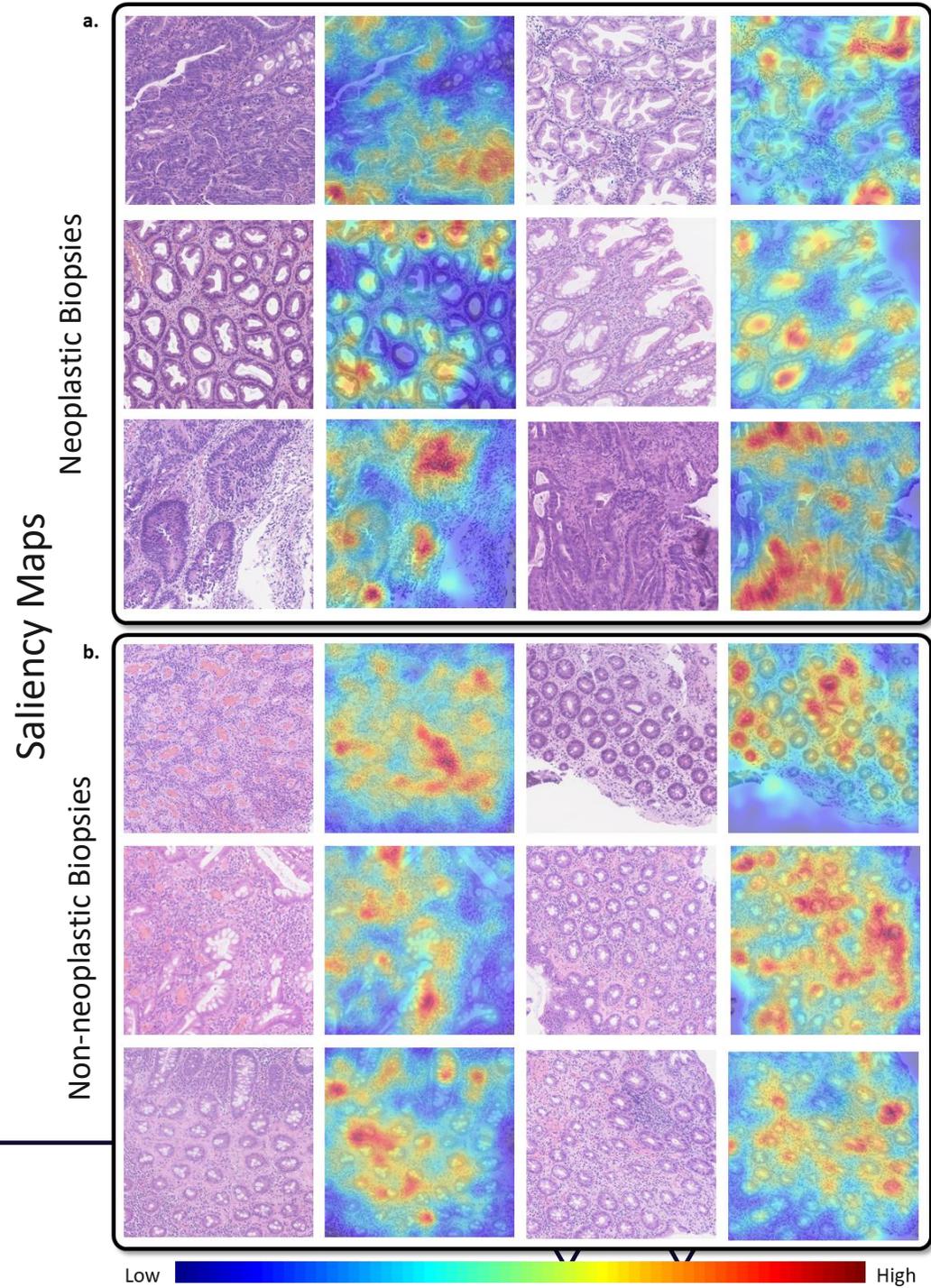
h. IMP diagnostic center



CAIMAN IDaRS



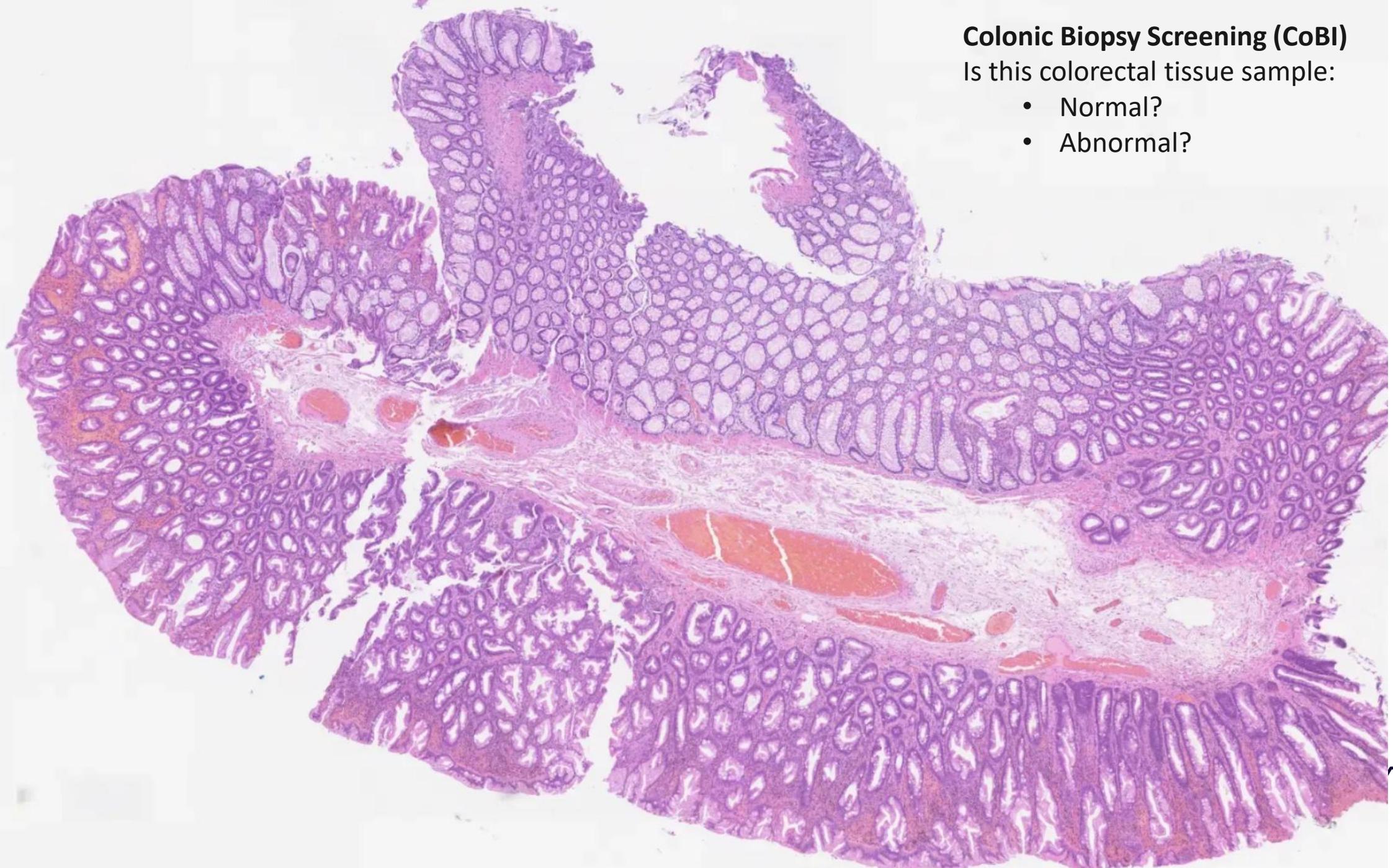
WSI Results



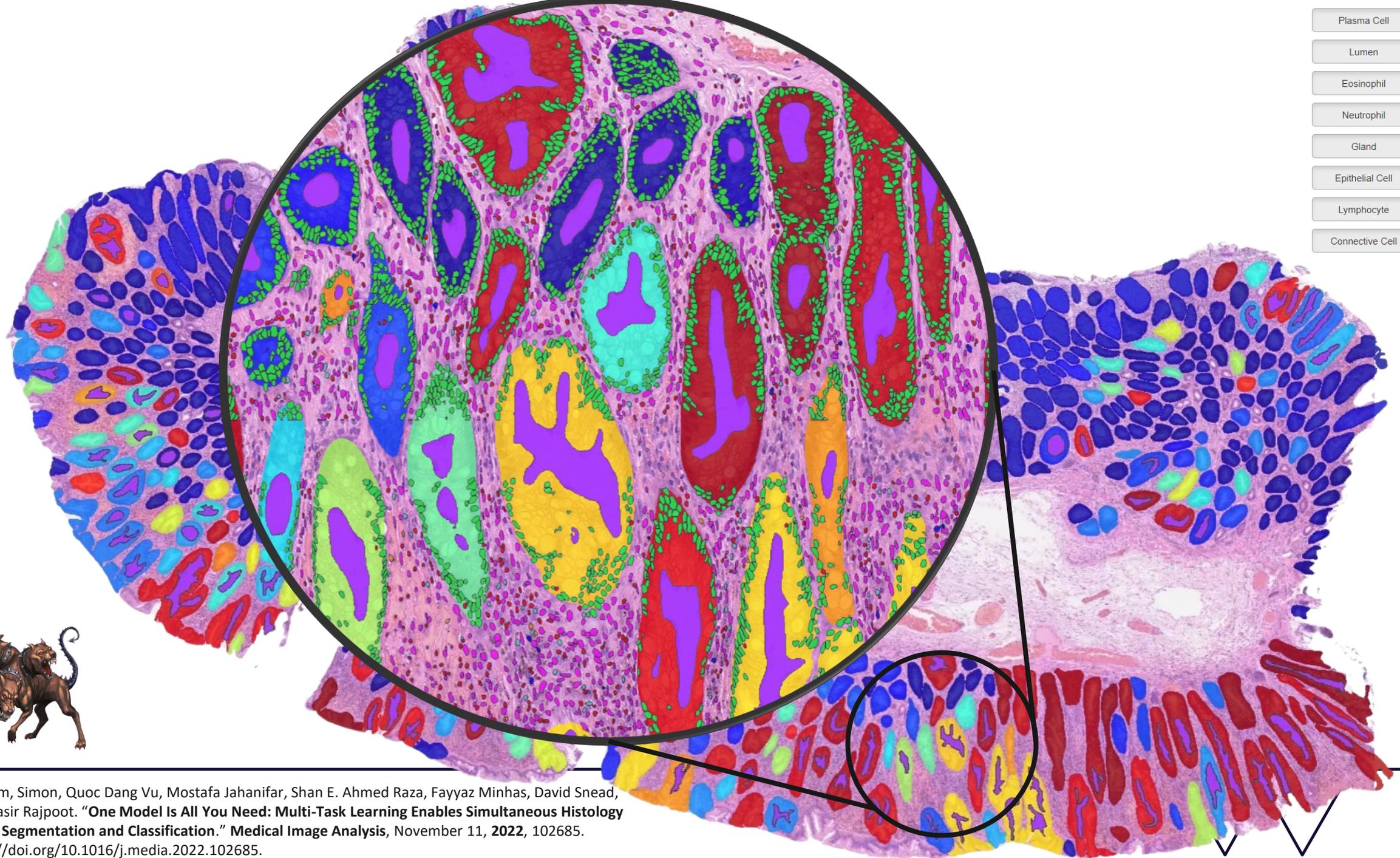
Colonic Biopsy Screening (CoBI)

Is this colorectal tissue sample:

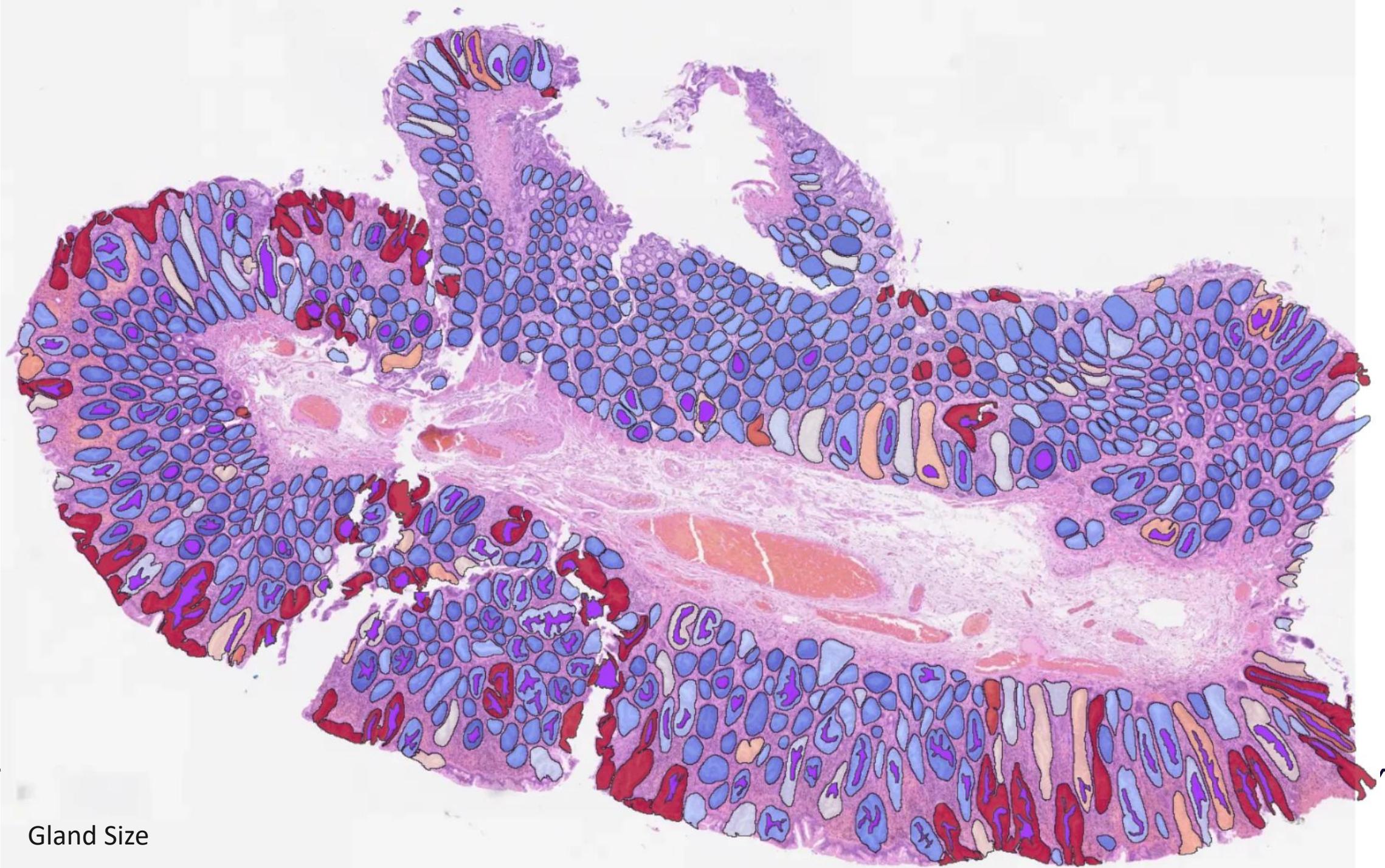
- Normal?
- Abnormal?



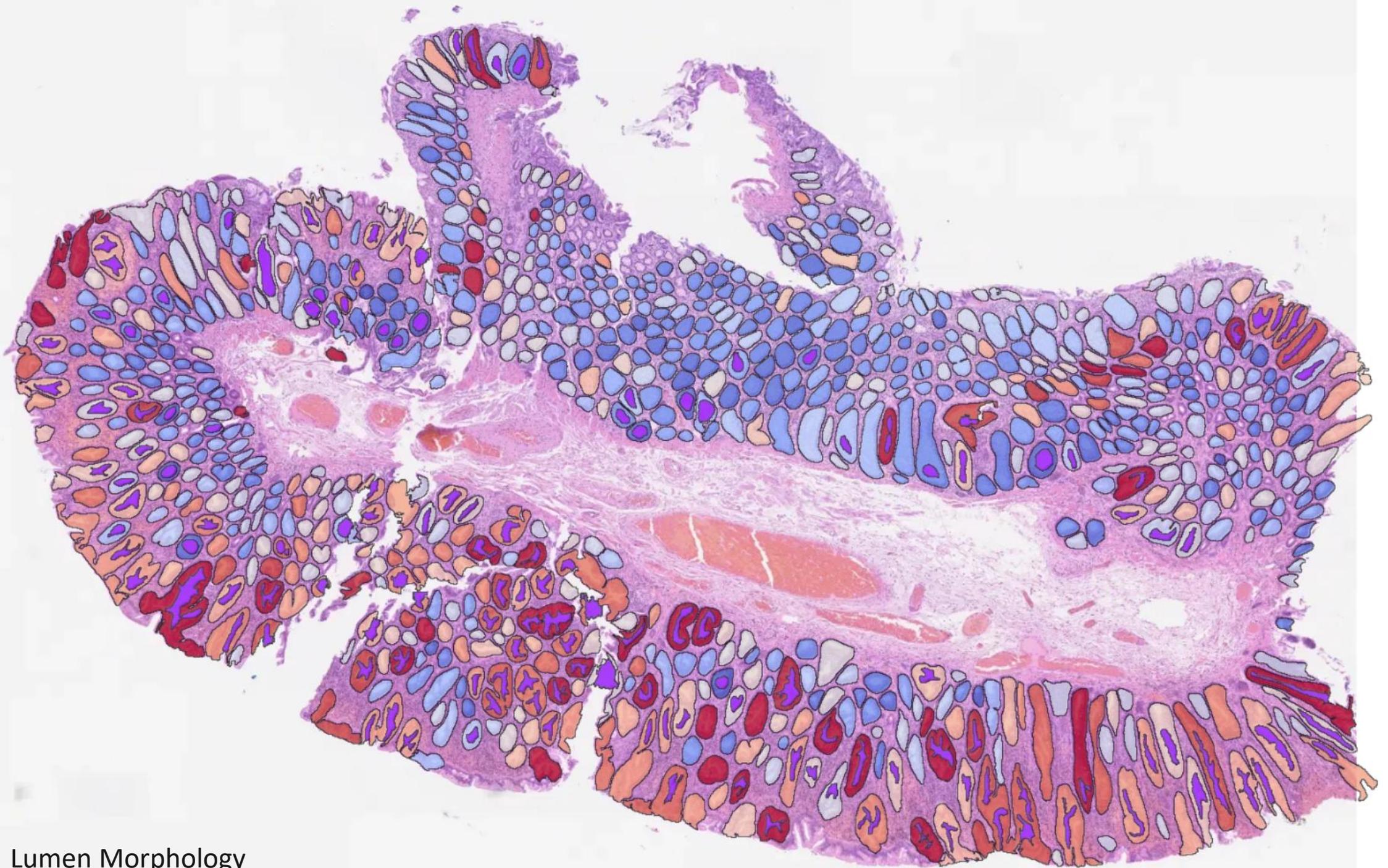
Plasma Cell	
Lumen	
Eosinophil	
Neutrophil	
Gland	
Epithelial Cell	
Lymphocyte	
Connective Cell	



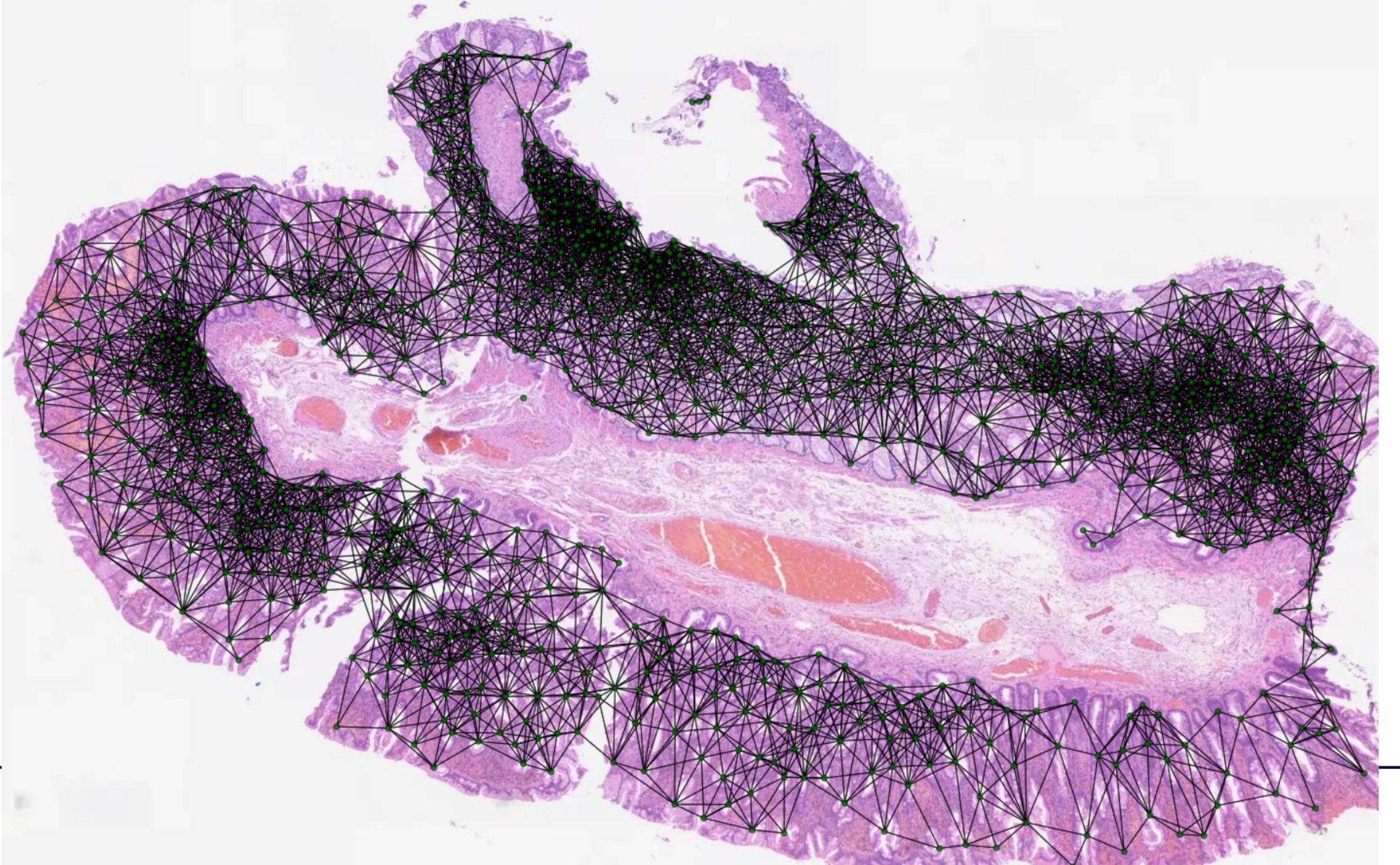
Graham, Simon, Quoc Dang Vu, Mostafa Jahanifar, Shan E. Ahmed Raza, Fayyaz Minhas, David Snead, and Nasir Rajpoot. "One Model Is All You Need: Multi-Task Learning Enables Simultaneous Histology Image Segmentation and Classification." *Medical Image Analysis*, November 11, 2022, 102685. <https://doi.org/10.1016/j.media.2022.102685>.

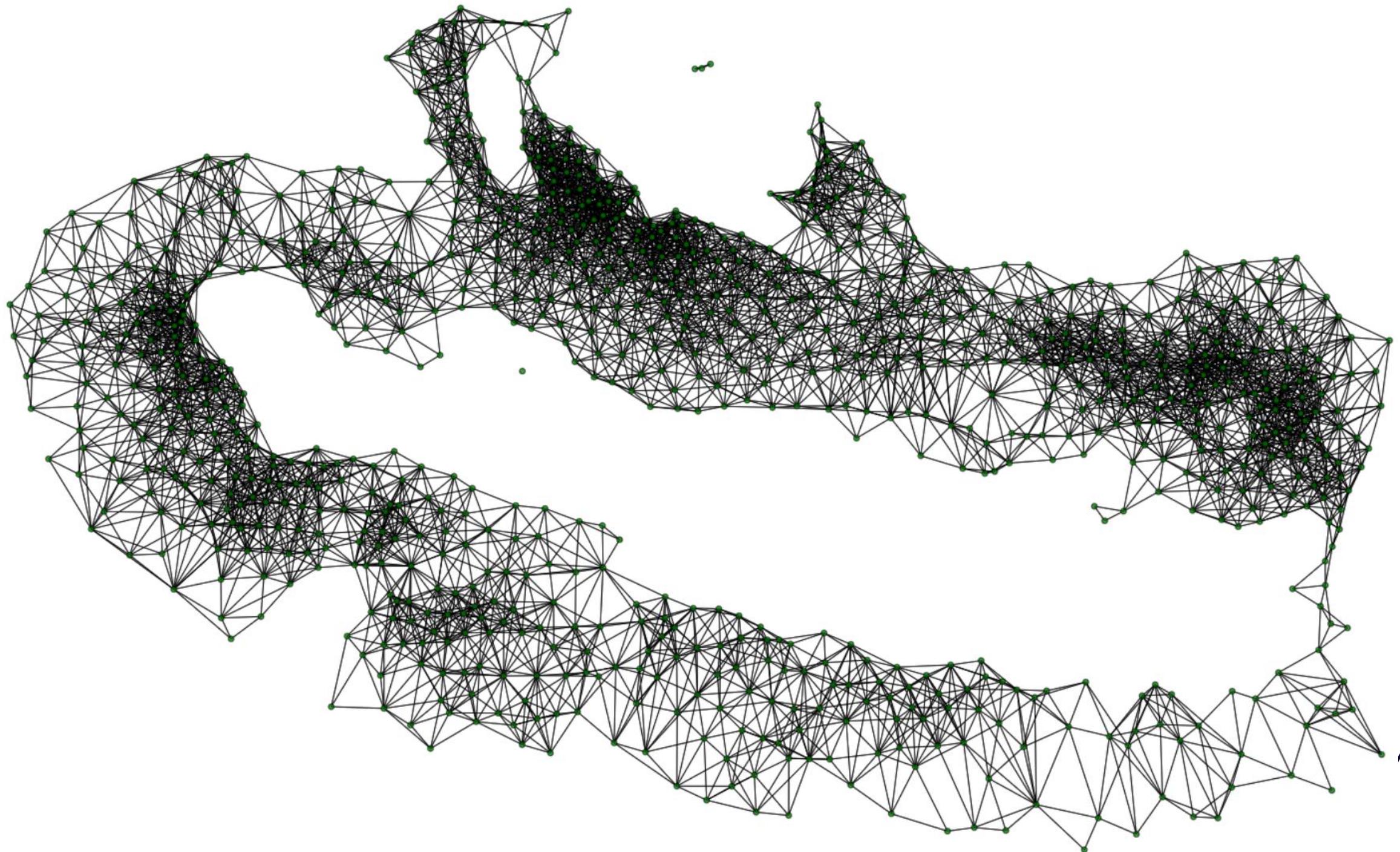


Gland Size



Lumen Morphology

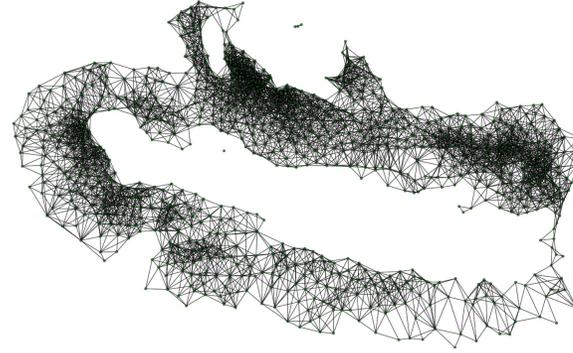




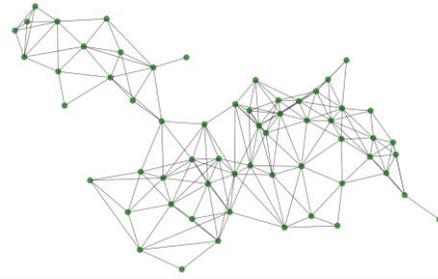
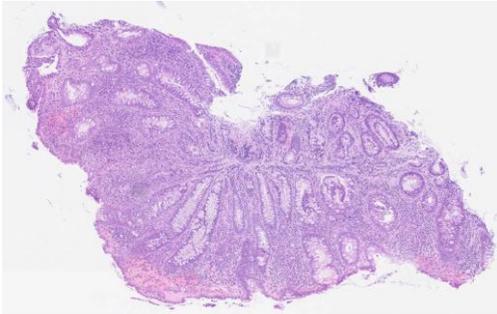
WSIs

Gland Graphs

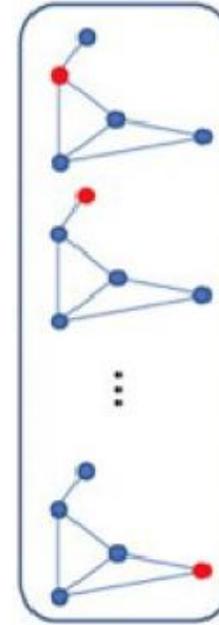
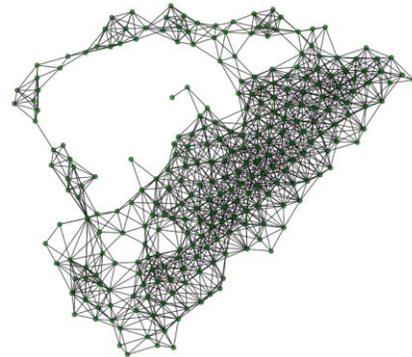
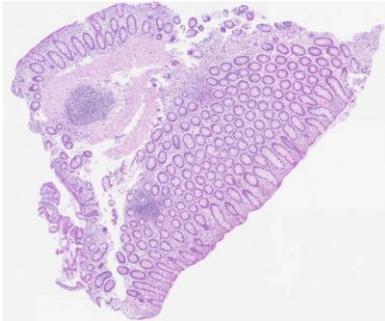
Abnormal: Hyperplastic



Abnormal: Inflammatory



Normal

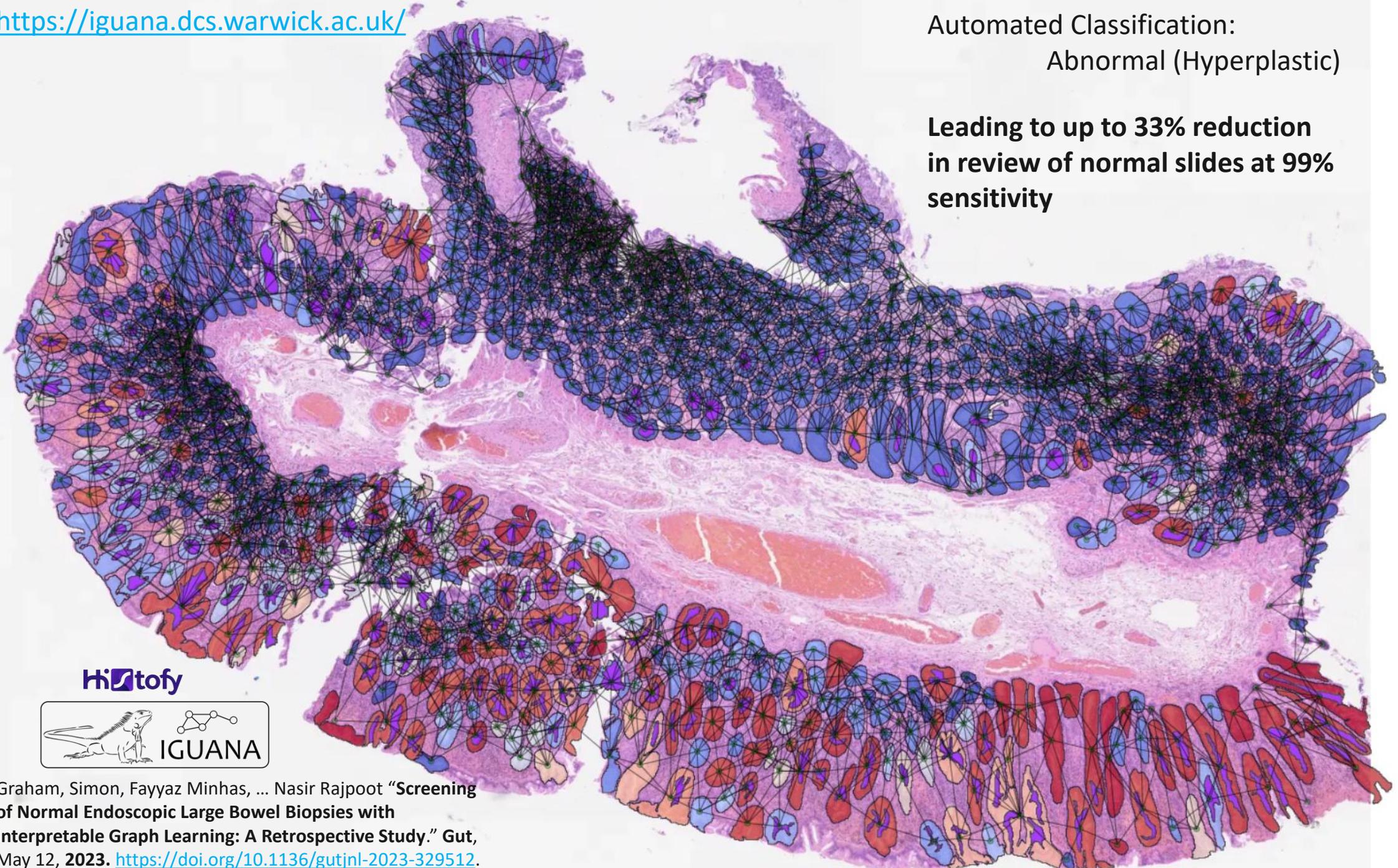


Prediction

Graph Neural Network

Automated Classification:
Abnormal (Hyperplastic)

Leading to up to 33% reduction
in review of normal slides at 99%
sensitivity



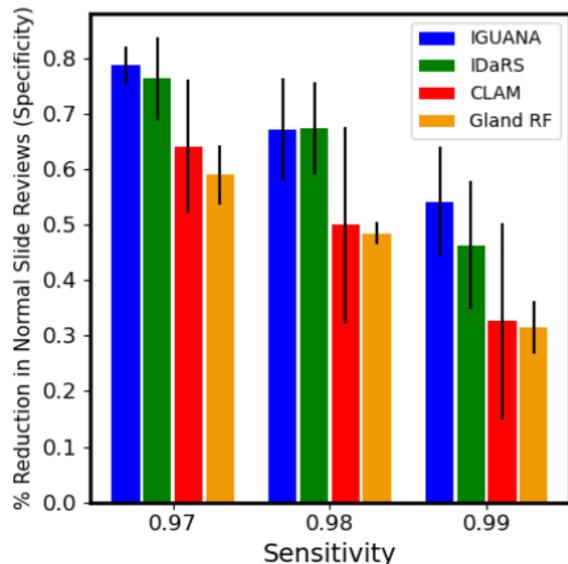
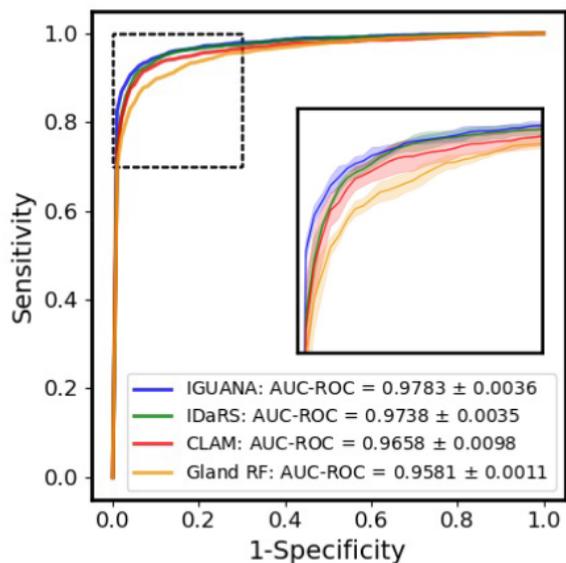
Hi tofy



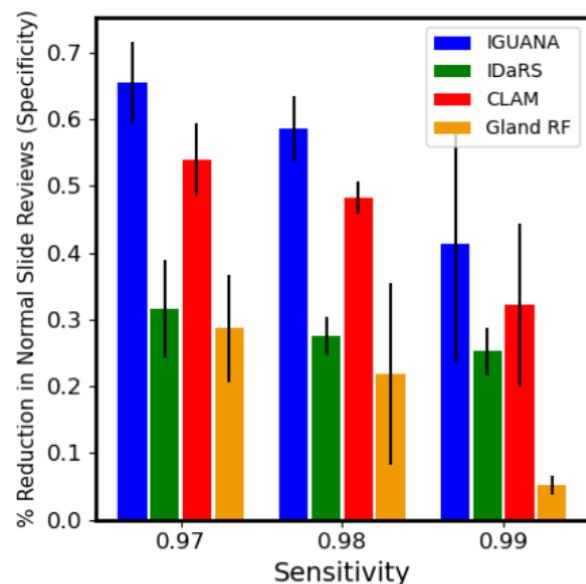
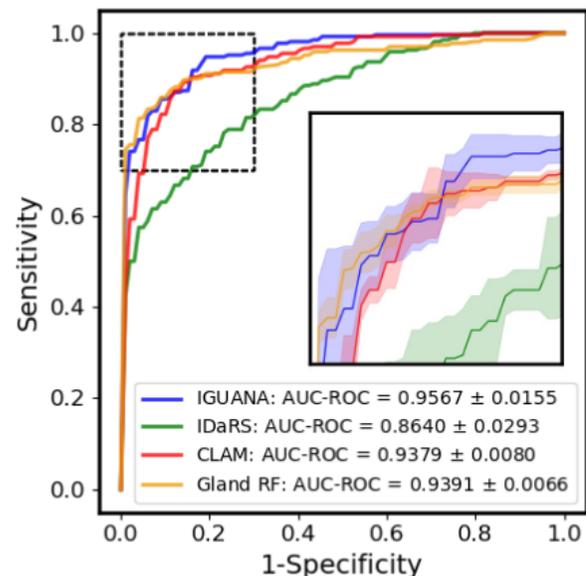
Graham, Simon, Fayyaz Minhas, ... Nasir Rajpoot "Screening of Normal Endoscopic Large Bowel Biopsies with Interpretable Graph Learning: A Retrospective Study." Gut, May 12, 2023. <https://doi.org/10.1136/gutjnl-2023-329512>.

Internal Cross Validation

UHCW

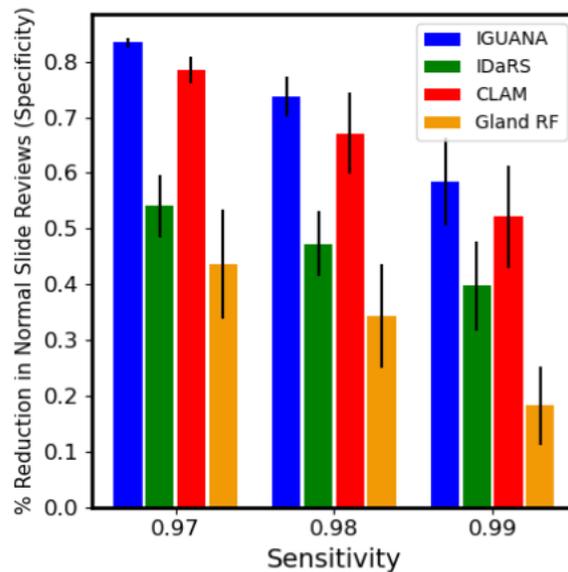
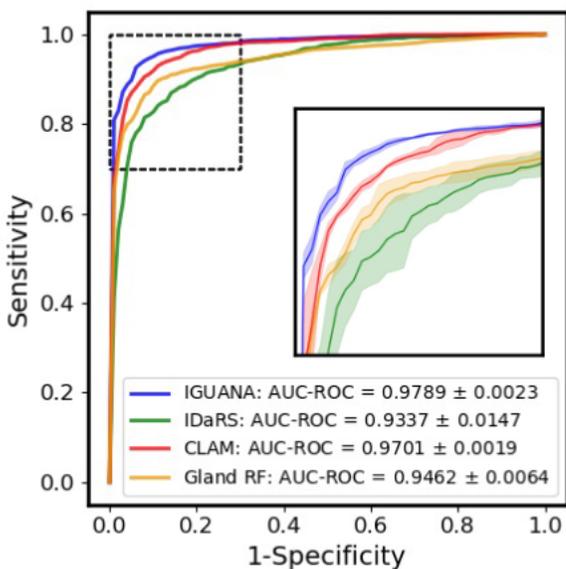


ESNE

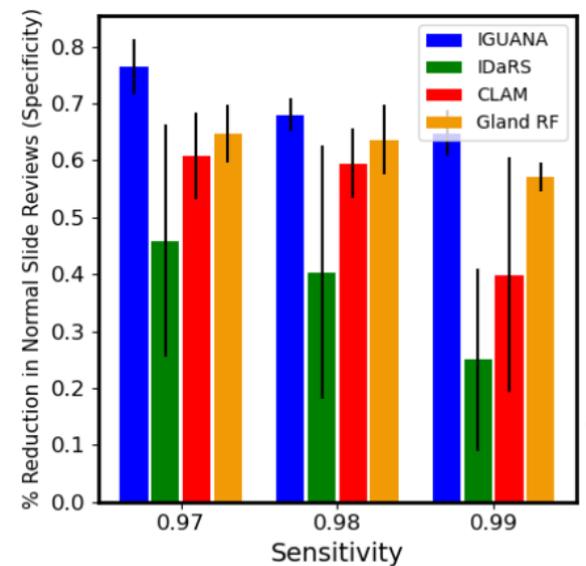
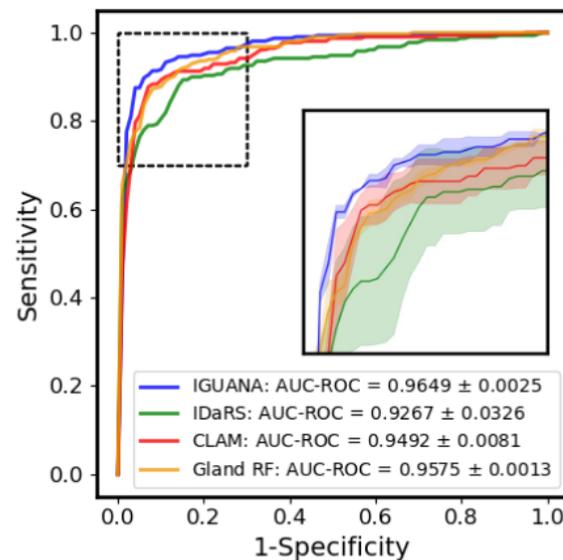


Independent Validation

IMP Diagnostics

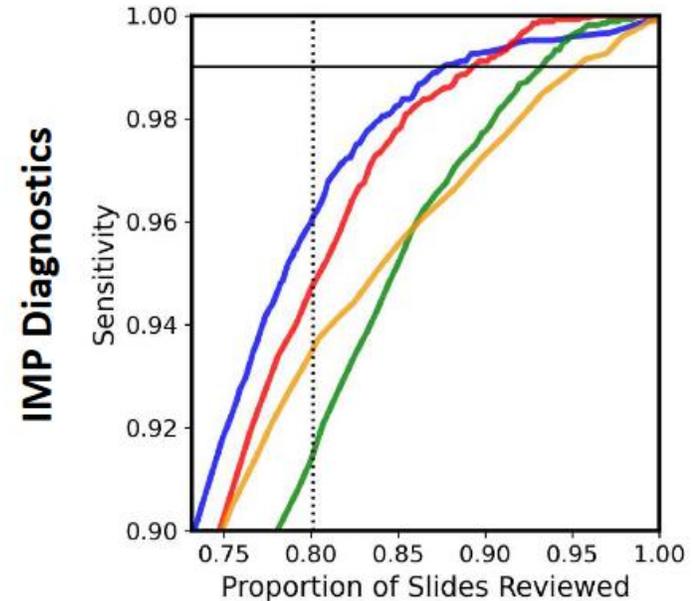
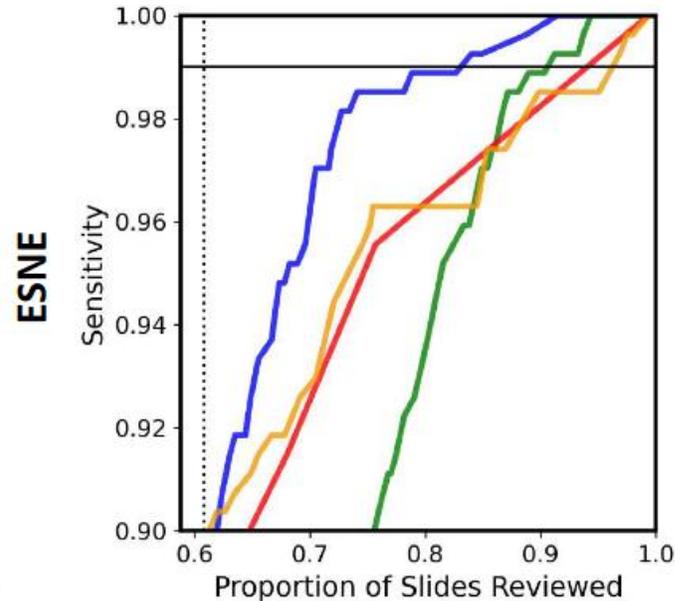
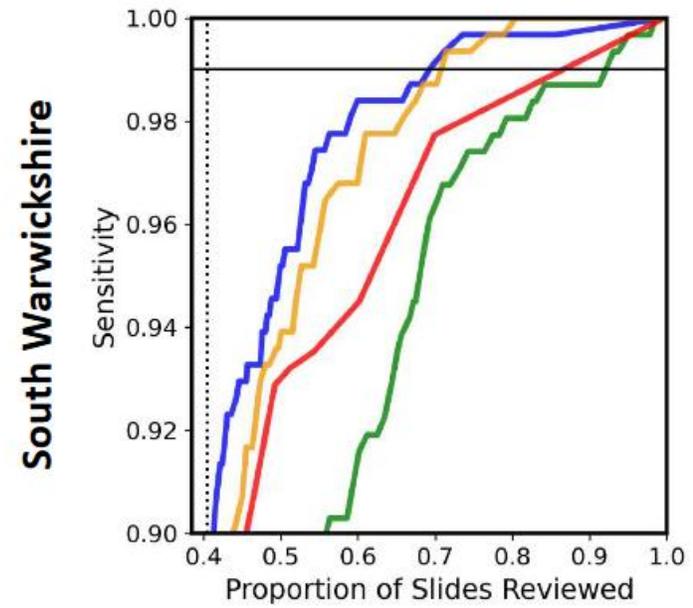
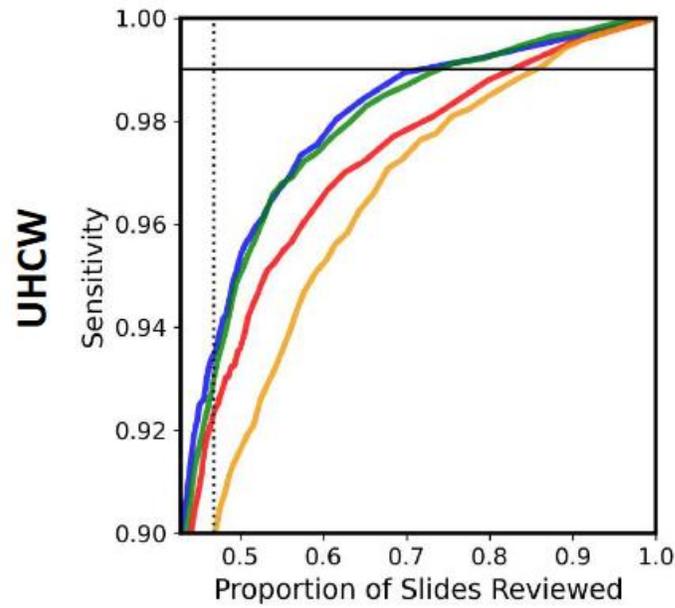


South Warwickshire



So what?

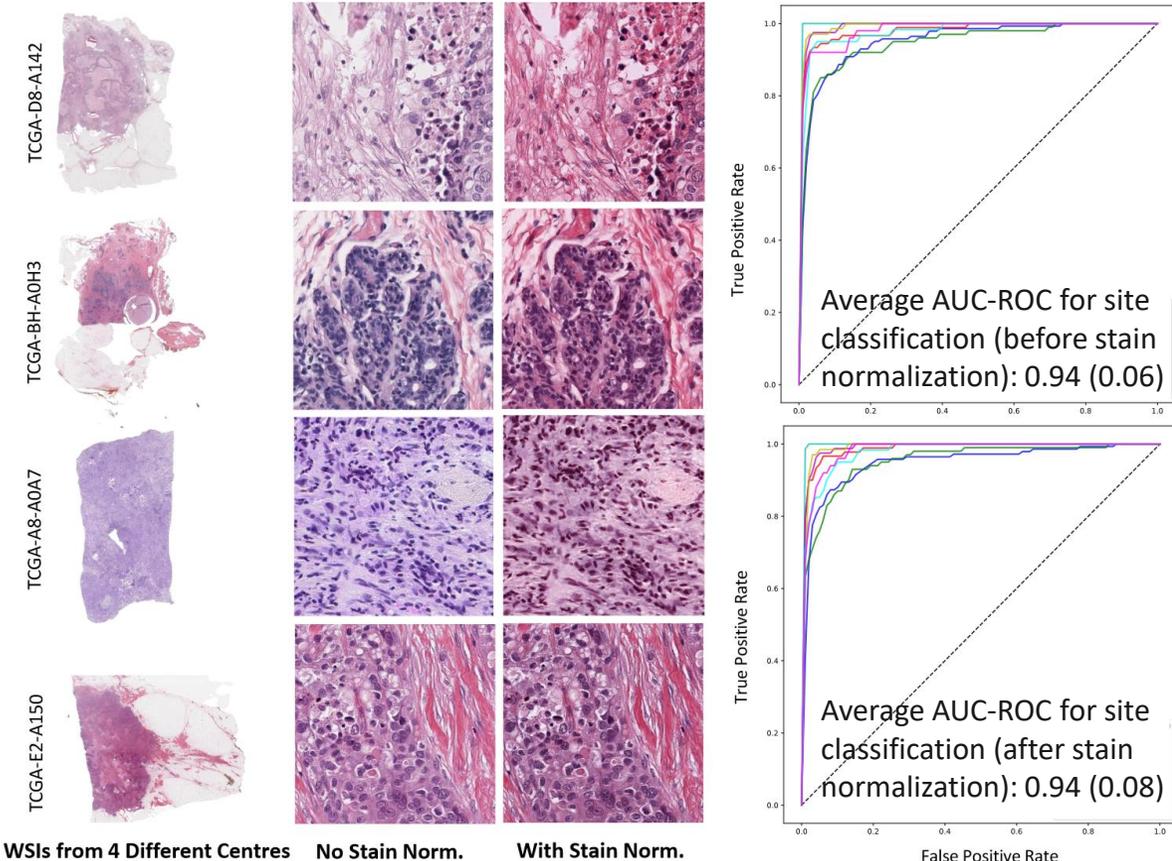
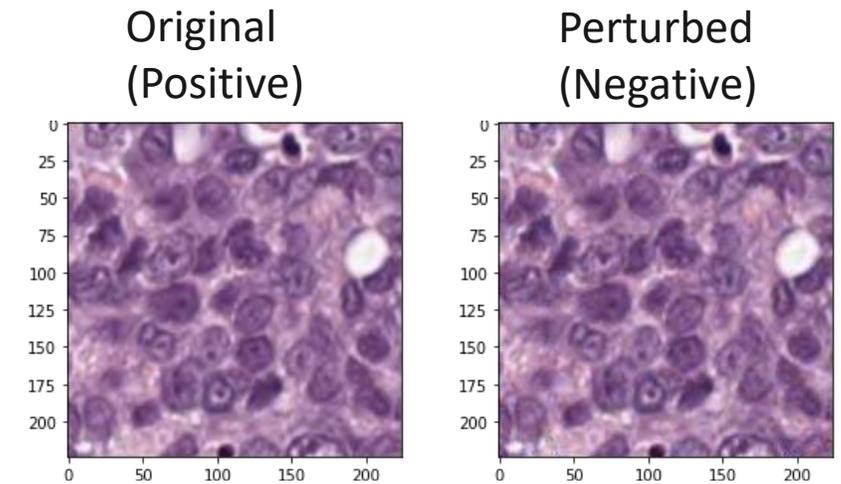
- How many slides need to be reviewed (to get a sensitivity of 99%)?



Next Steps: CoBIx

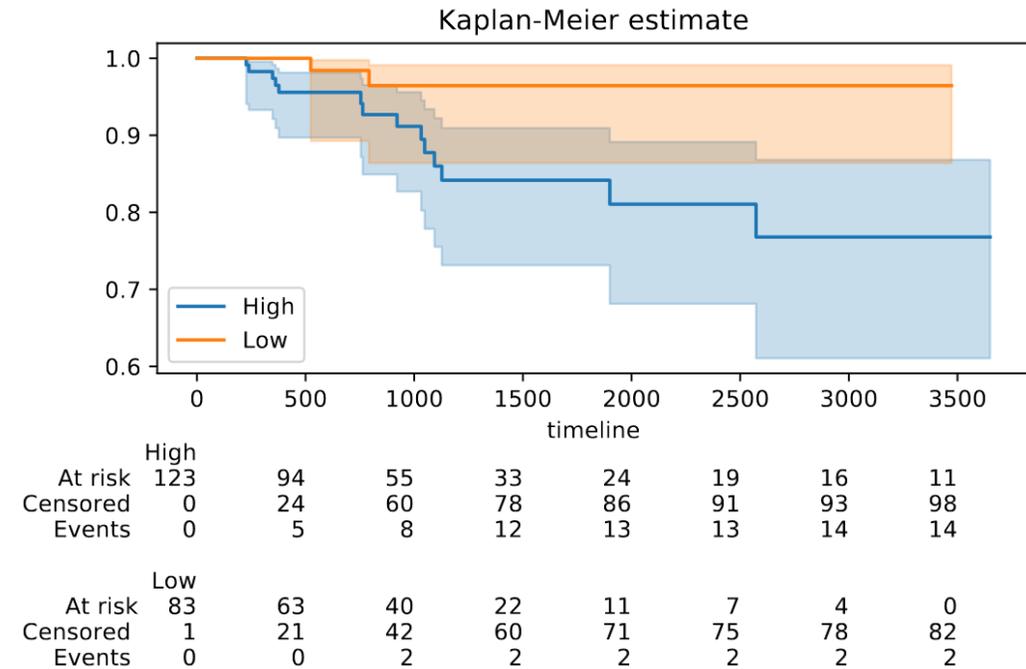
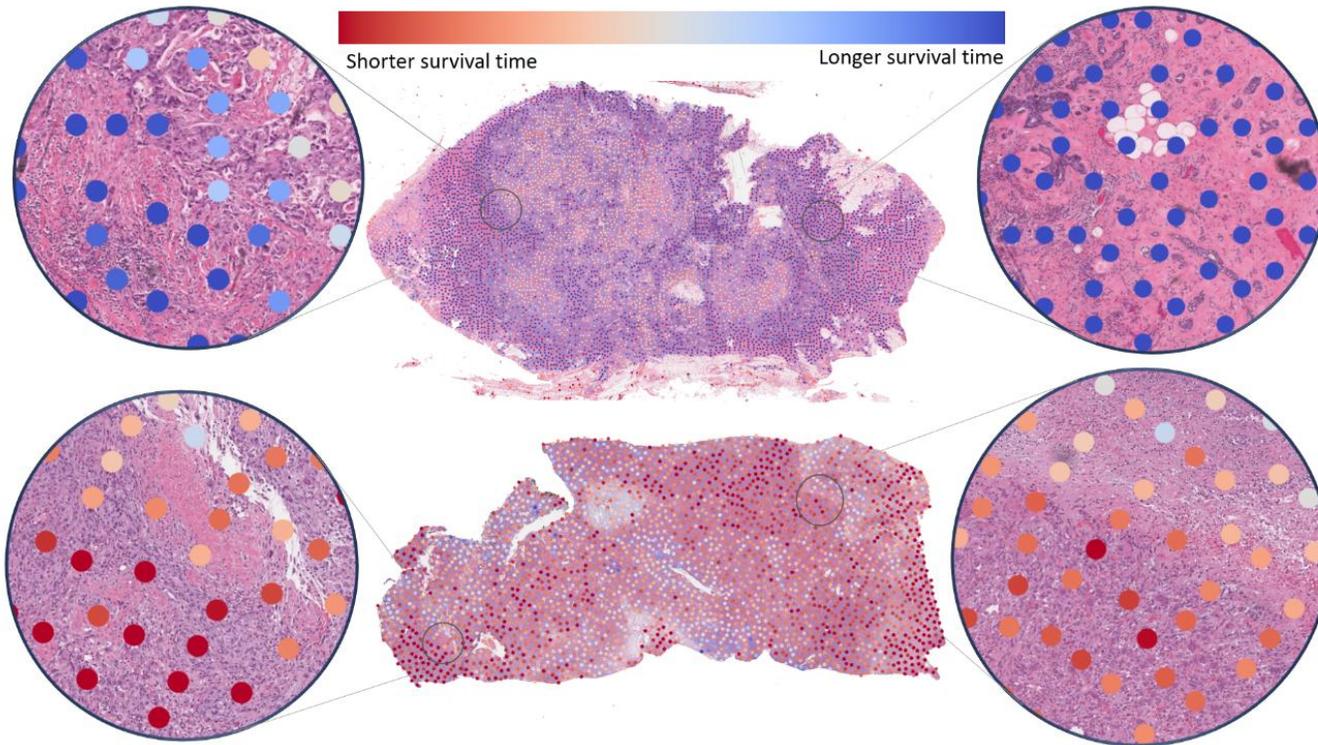
NIHR-COBIx: Multi-site validation of automated AI tool for screening of large bowel endoscopic biopsy slides

- Larger validation
 - Data Quality
 - Ethnicity
 - Batch effects & Source site signatures
- Analysing robustness
- Improving transparency & explainability
- Health Economics
- Regulatory approvals
- Productization



Footnote: Foote, Alex, Amina Asif, Nasir Rajpoot, and Fayyaz Minhas. "REET: Robustness Evaluation and Enhancement Toolbox for Computational Pathology." *Bioinformatics*, May 9, 2022, btac315. <https://doi.org/10.1093/bioinformatics/btac315>.
Dawood, Muhammad, Piotr Keller, and Fayyaz ul Amir Afsar Minhas. "Do Tissue Source Sites Leave Identifiable Signatures in Whole Slide Images beyond Staining?," *ICLR-W TML4H*, 2023. <https://openreview.net/forum?id=flfJ1OwD-FD>. Video: <https://www.youtube.com/watch?v=lq9g7xmZYPA>

WSI graphs to predict patient survival

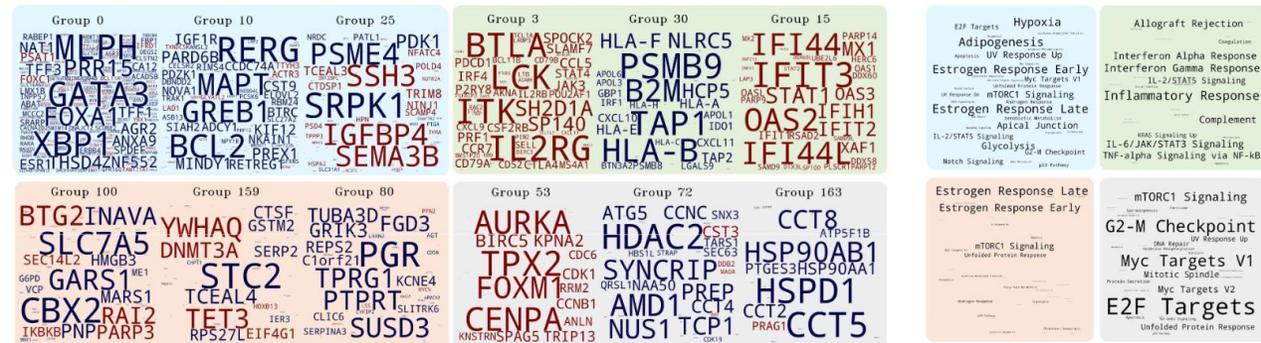


Engineering and
Physical Sciences
Research Council

Mackenzie, Callum Christopher, Muhammad Dawood, Simon Graham, Mark Eastwood, and Fayyaz ul Amir Afsar Minhas. "Neural Graph Modelling of Whole Slide Images for Survival Ranking." In *Proceedings of the First Learning on Graphs Conference*, 48:1-48:10. PMLR, 2022.

<https://proceedings.mlr.press/v198/mackenzie22a.html>.

Predicting gene expression from WSIs

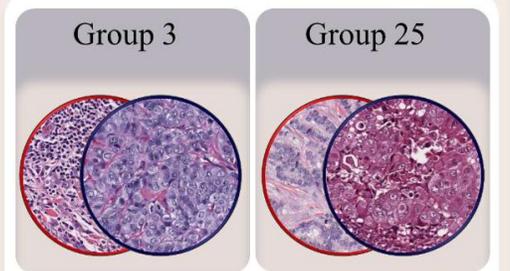


Patient Data

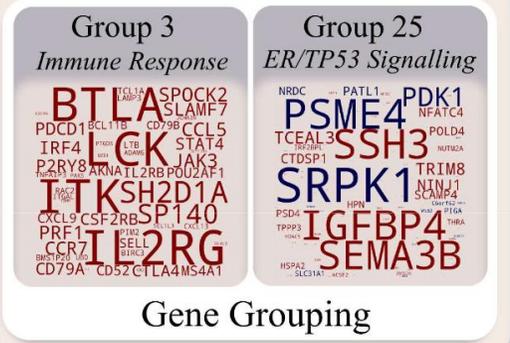
Diagnostic Slide

Gene Expression Profile

Phenotype to Genotype

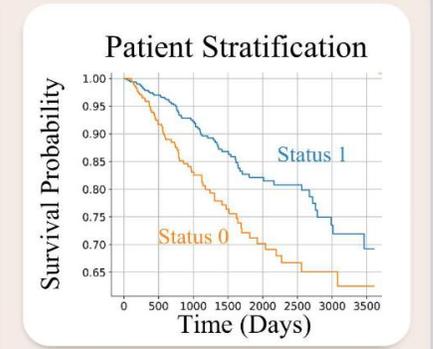


Histological Correlates

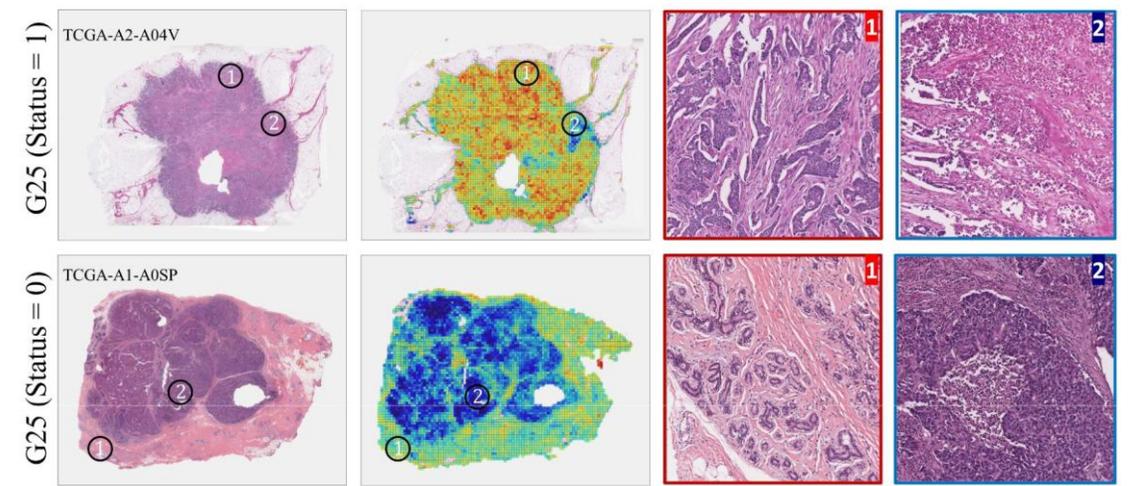
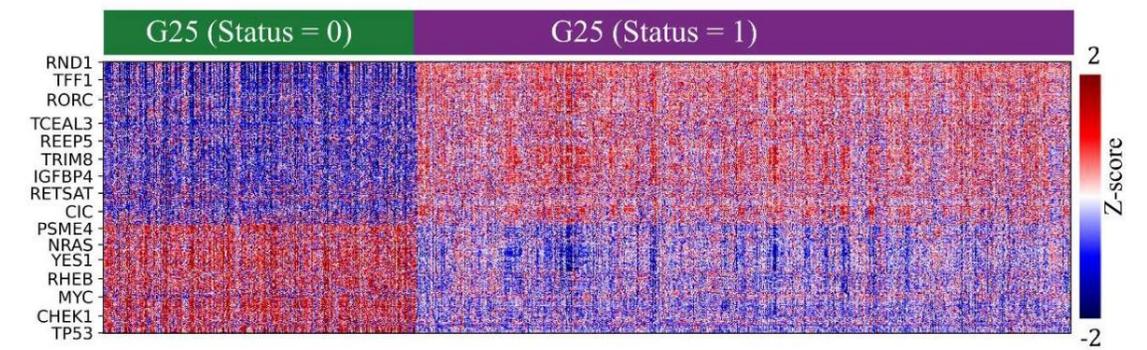


Gene Grouping

Therapeutic Implication



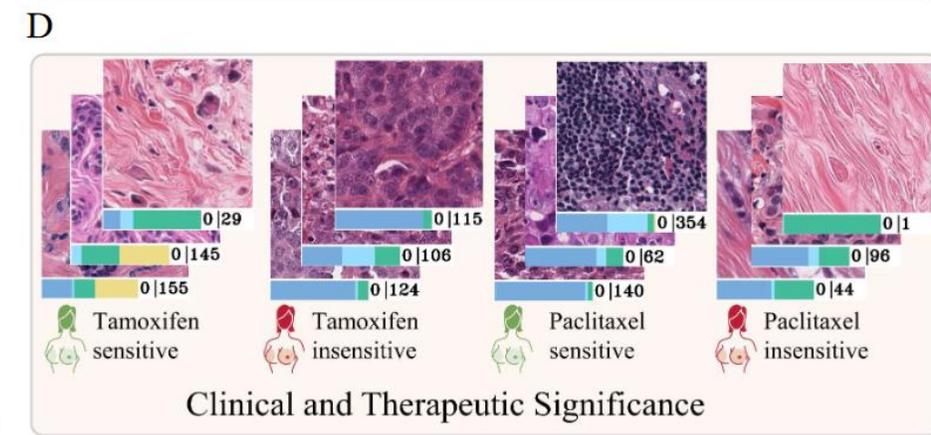
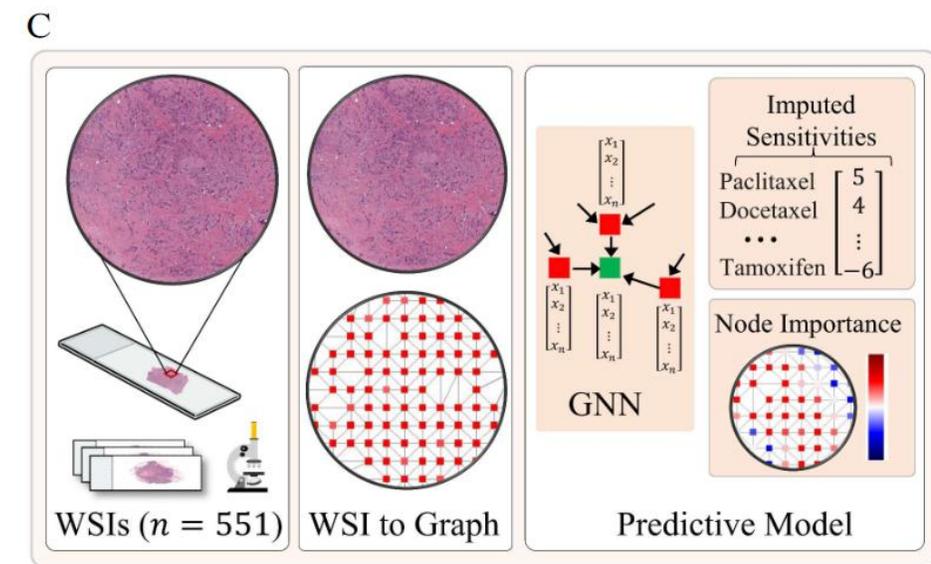
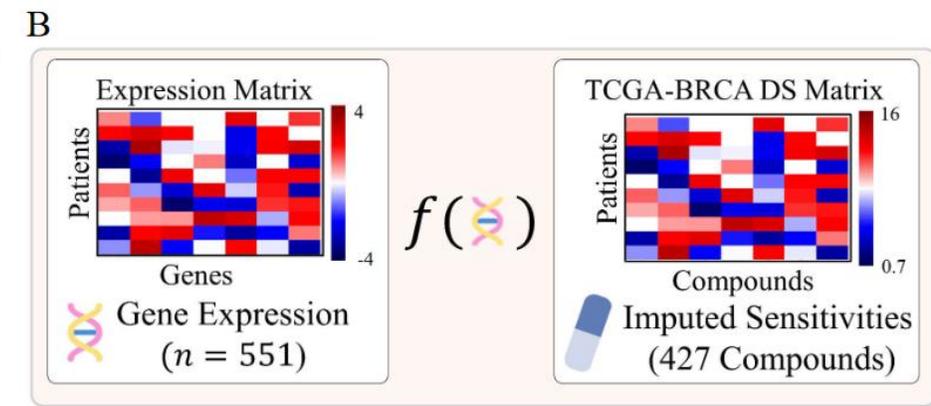
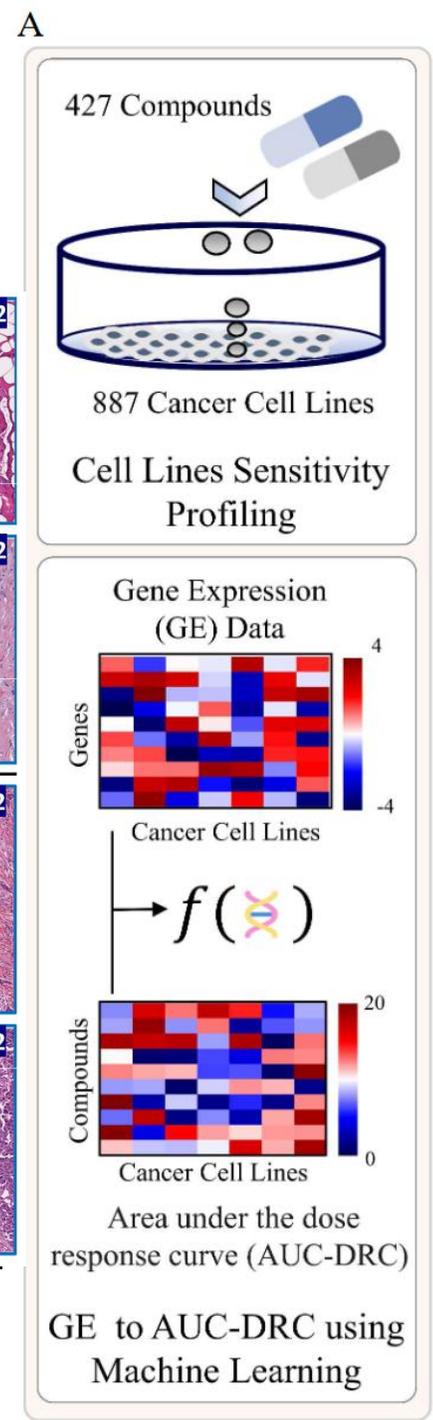
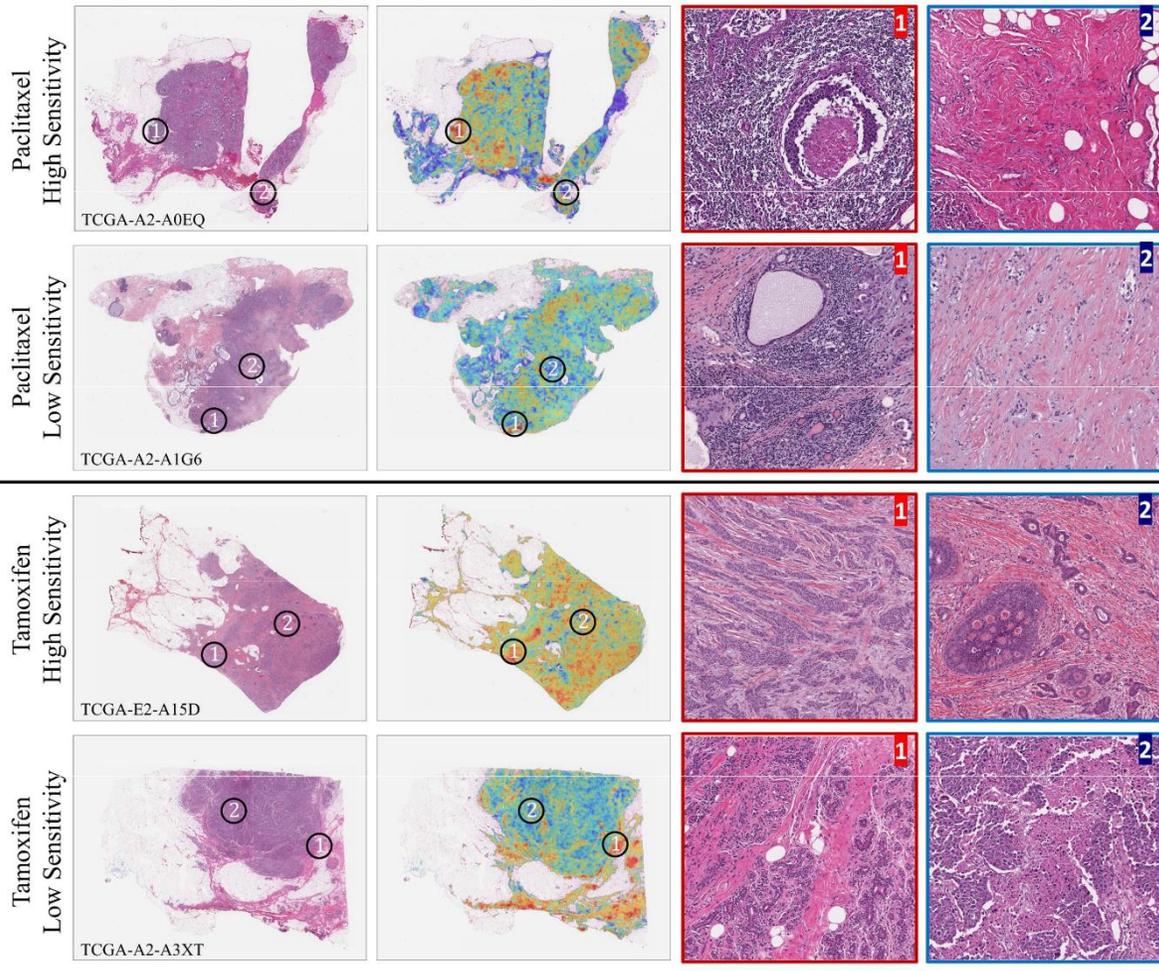
Therapy Planning



Dawood, Muhammad, Mark Eastwood, Mostafa Jahanifar, Lawrence Young, Asa Ben-Hur, Kim Branson, Louise Jones, Nasir Rajpoot, and Fayyaz ul Amir Afsar Minhas. "Cross-Linking Breast Tumor Transcriptomic States and Tissue Histology." *Cell Reports Medicine* 4, no. 12 (December 19, 2023). <https://doi.org/10.1016/j.xcrm.2023.101313>.



WSI graphs to predict drug sensitivities



Dawood, Muhammad, Quoc Dang Vu, Lawrence S. Young, Kim Branson, Louise Jones, Nasir Rajpoot, and Fayyaz ul Amir Afsar Minhas. "Cancer Drug Sensitivity Prediction from Routine Histology Images." *Npj Precision Oncology* 8, no. 1 (January 6, 2024): 1–13. <https://doi.org/10.1038/s41698-023-00491-9>.

Using Large Language Models

- For Standardization of Reports: LABIEB

Pathology Report

Diagnosis:
This is a moderately to poorly differentiated adenocarcinoma of the rectosigmoidal junction of grade G2 to 3, with tumor infiltration of the parietal layers into the subserous fatty connective tissue, with a regional lymph node metastasis (1/14) and tumor-free overview slices from all other resection material portions described.

Tumour stage of pT3, pN1 (1/14), MX, L1, R0.

Prompts



Extracted Fields

```
{
  "Site": "Rectosigmoid",
  "Grade": "Low",
  "Local Invasion": "pT3",
  "Size": "35 mm",
  "Resection": "R0",
  ...
}
```

JSON

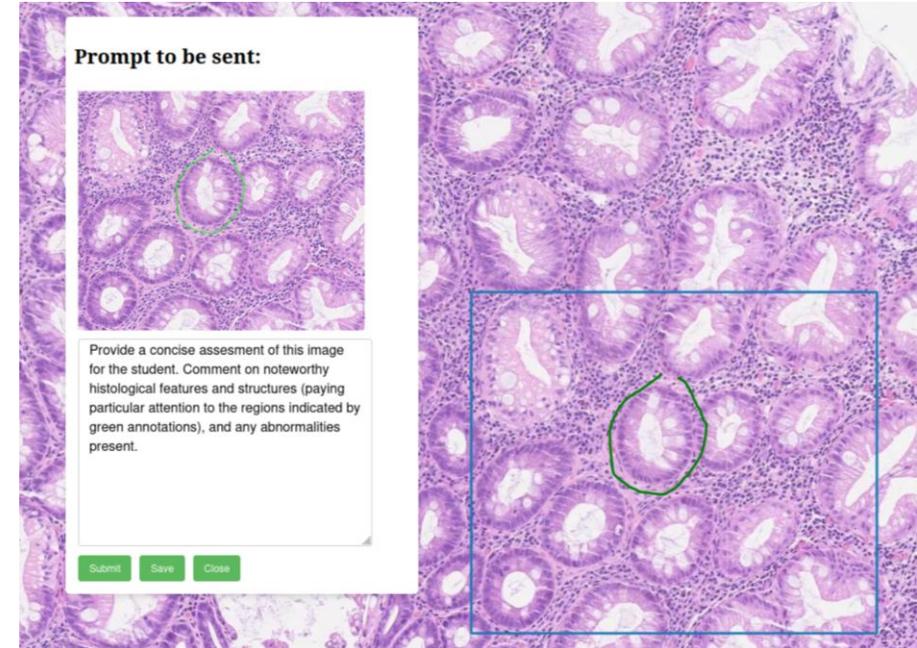
LMM

Standardised Pathology Report

Reporting proforma for colorectal carcinoma resection specimens	
Site of Tumour: Rectosigmoid	Number of lymph nodes: 14
Tumour type: Adenocarcinoma	Number of involved lymph nodes: 1
Differentiation by worst area: Low	Lymph node status: pN1a
Local invasion: pT3	Pathologically confirmed distant metastatic disease: pMX
TNM (8 th edition) pT3 , pN1a , pMX	Resection status: R0

<https://labieb.dcs.warwick.ac.uk/>

- Image Interpretation: TIAViz



GPT-vision response:

This is an H&E (hematoxylin and eosin) stained histological image that appears to show a section of the gastrointestinal tract, most likely the colon or small intestine due to the presence of goblet cells, which are mucus-secreting cells. The epithelial layer is columnar with regular nuclei, which is typical for this part of the digestive tract. The green annotation circles a gland in the epithelium.

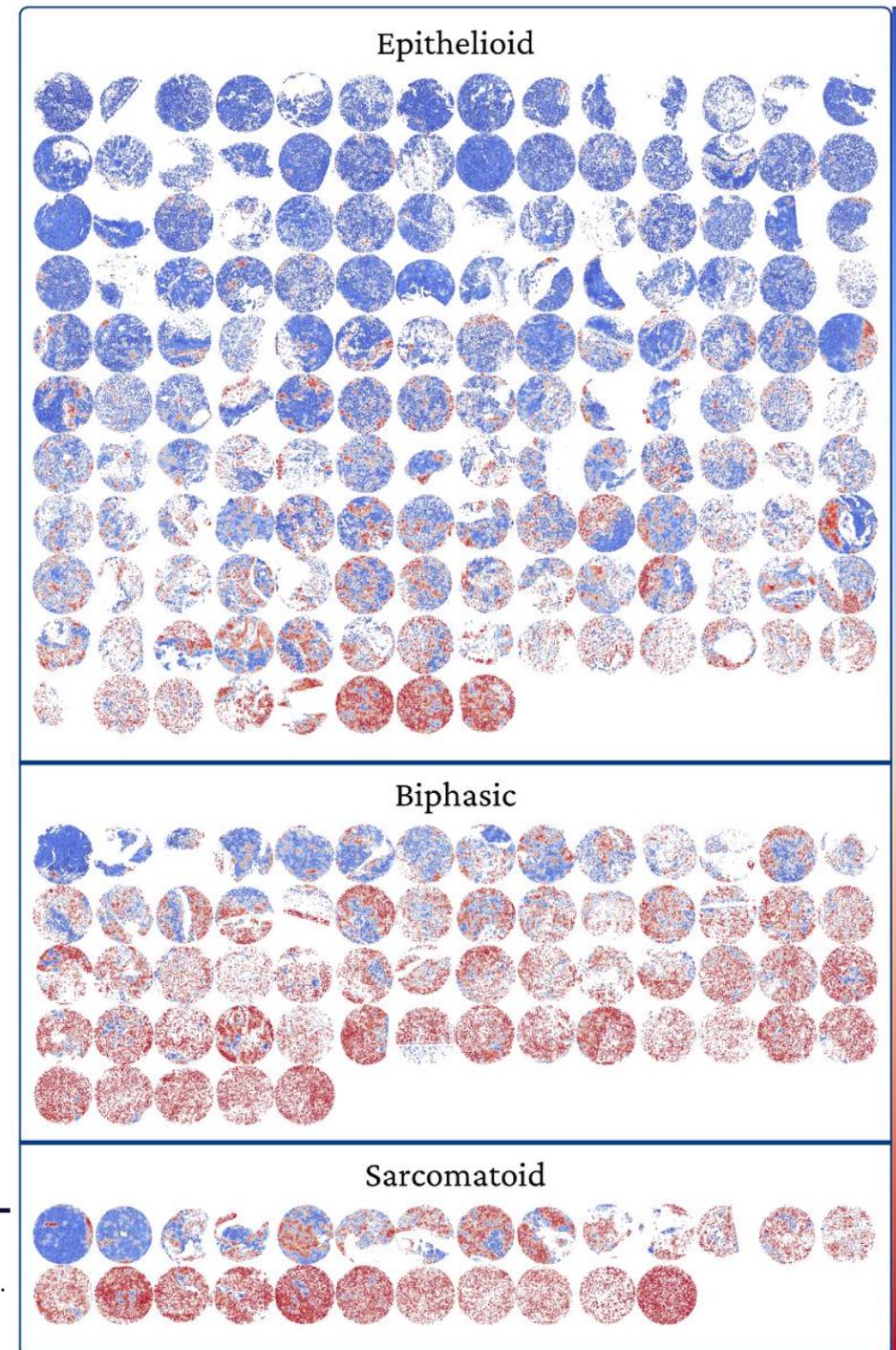
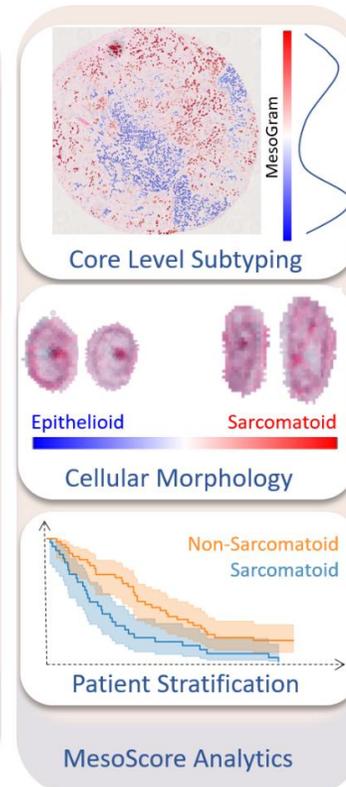
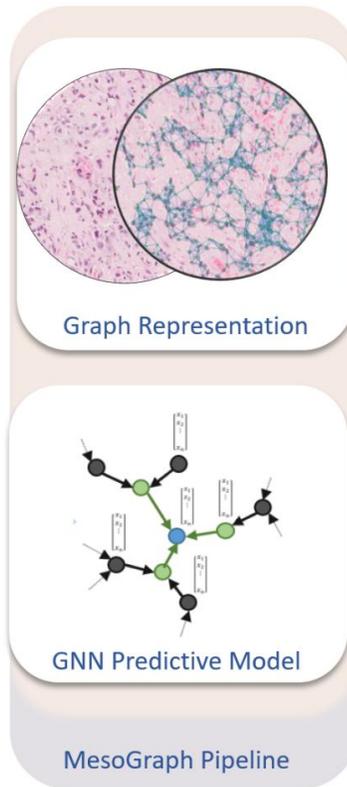
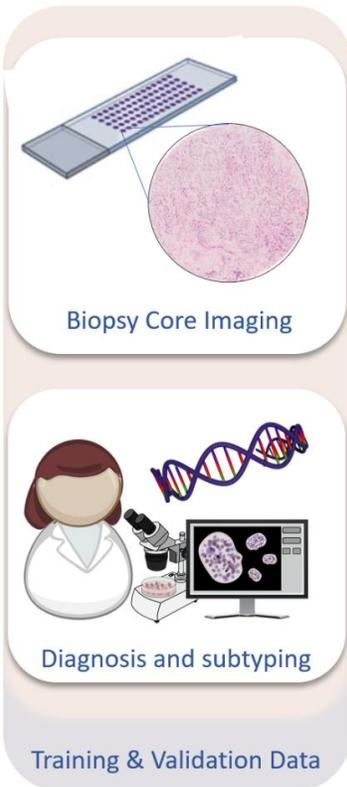
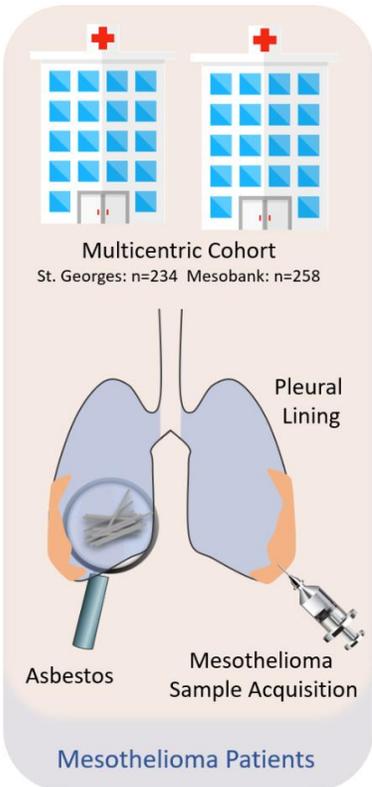
What's noteworthy in this region:

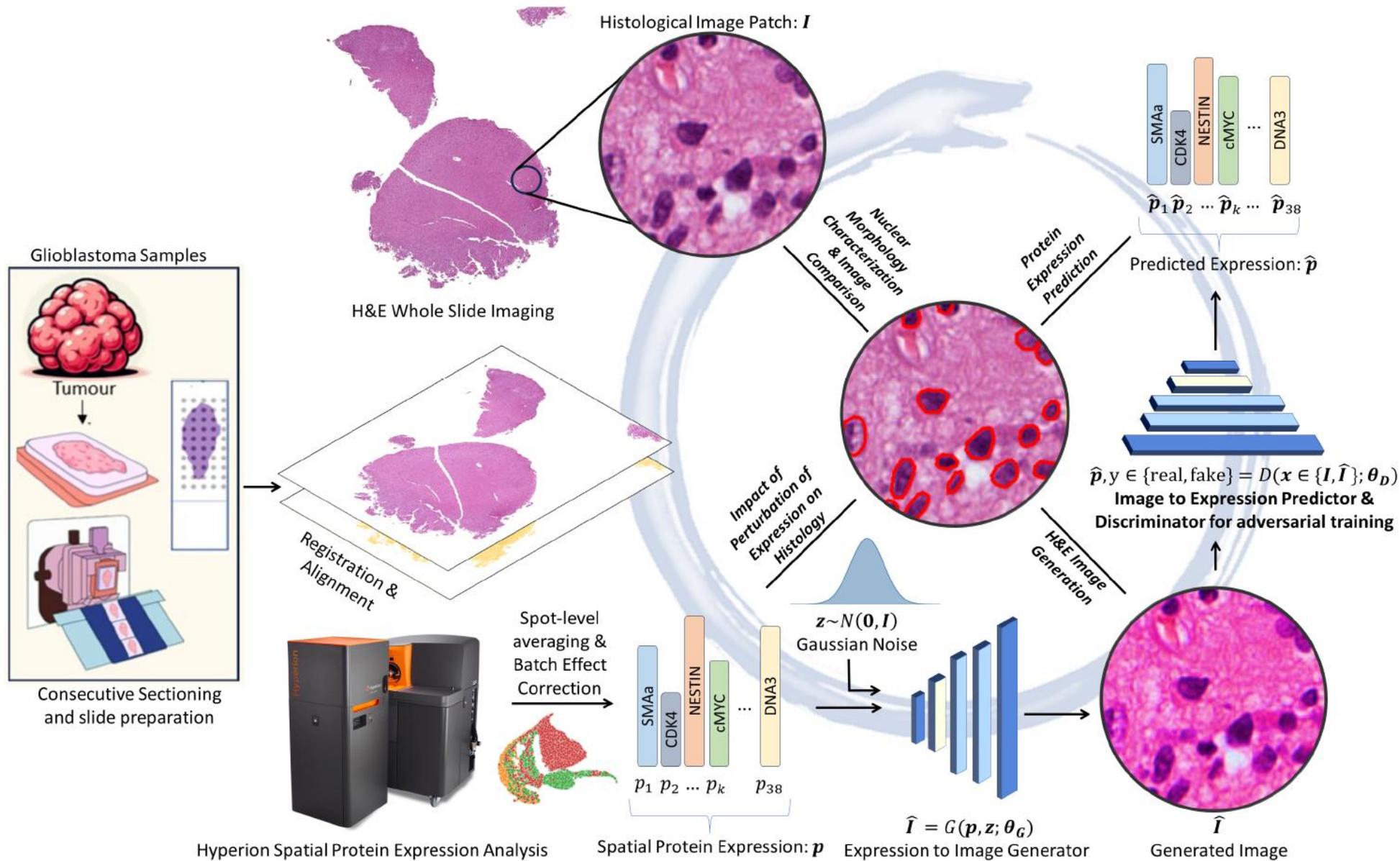
1. In the annotated area, there is an intestinal gland (crypt of Lieberkühn), which looks to be lined with regularly spaced columnar epithelial cells with basal, round to oval nuclei. The presence of goblet cells interspersed between the absorptive cells is typical for normal intestinal mucosa.
2. Within the crypt, there is no evidence of dysplasia, which is characterized by variations in cell size (anisocytosis), shape (pleomorphism), nuclear enlargement, hyperchromasia, and increased mitotic activity. The cells retain uniformity and there's no apparent crowding or stratification of nuclei, which is reassuring.
3. The surrounding stroma contains lamina propria, which shows a moderate number of inflammatory cells which can be normal, but an increased number could indicate mild inflammation. However, there doesn't appear to be significant infiltrate to suggest active inflammation or an inflammatory bowel disease.
4. The crypt architecture is maintained and there doesn't appear to be any evidence of cryptitis or crypt abscesses, which would suggest more specific inflammatory conditions such as ulcerative colitis.

Abnormalities: No overt abnormalities can be determined from this single region of the image, as the annotated crypt seems to have retained its normal histological architecture. Pathologists must, however, assess multiple fields and levels to establish a definite diagnosis and determine the absence of pathology accurately.

For educational review, always correlate microscopic findings with macroscopic observations and clinical data

MesoGraph: subtyping mesothelioma

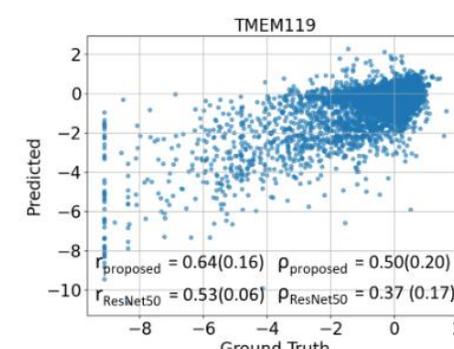
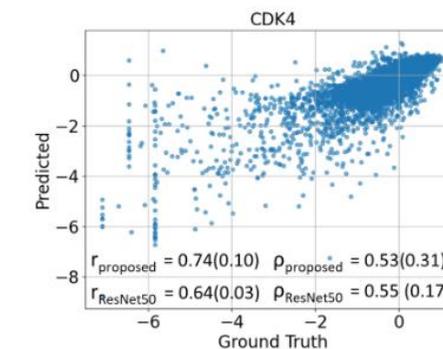
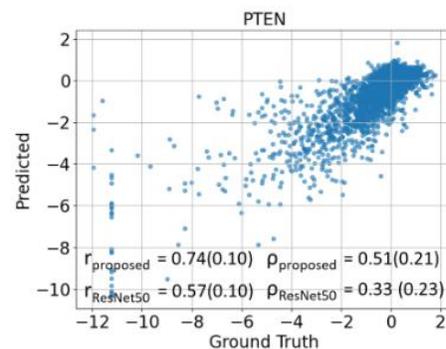
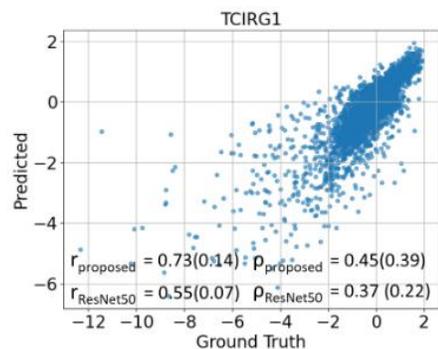
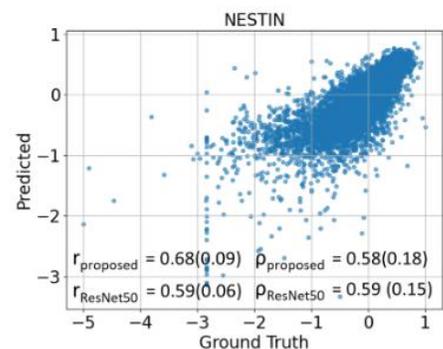
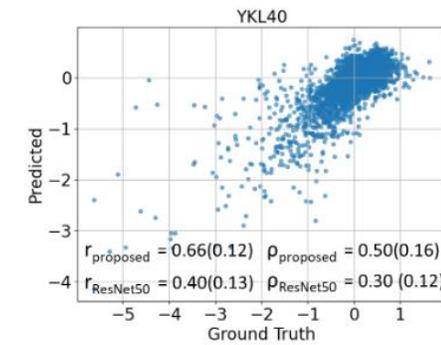
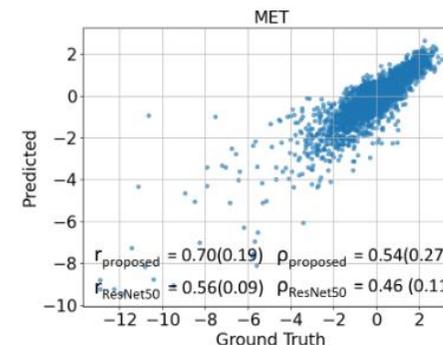
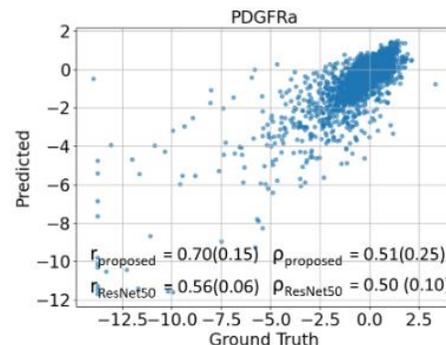
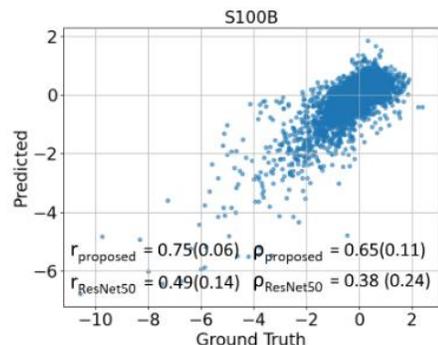
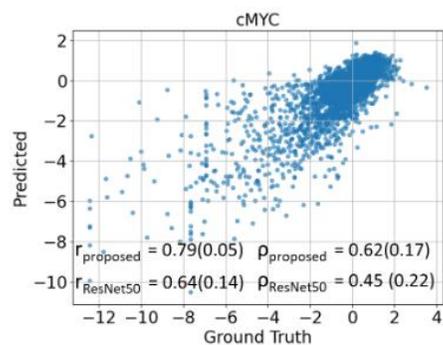




Can we predict spatial expression from imaging?

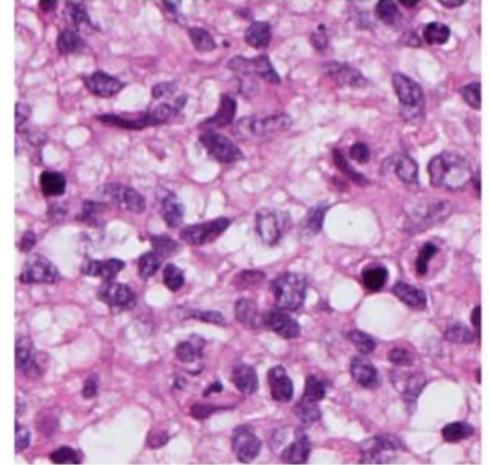
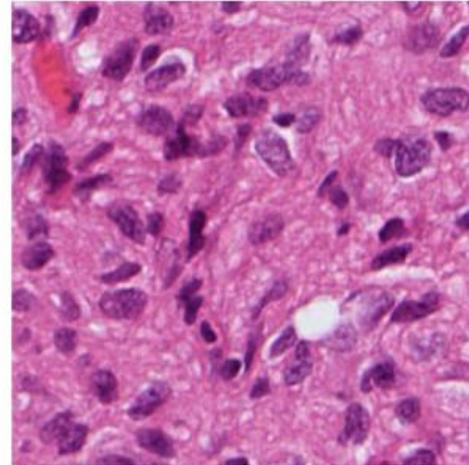
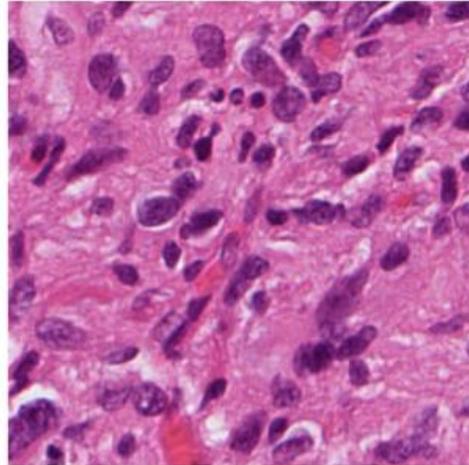
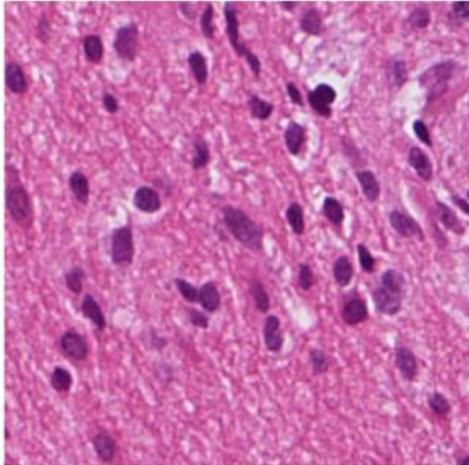
Can we generate images from expression profiles?

SpecTx: Predicting protein expression from Imaging

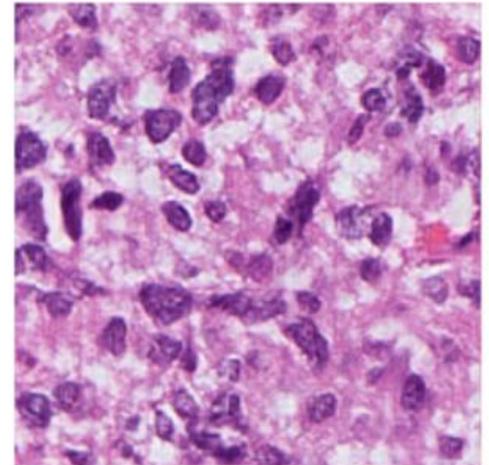
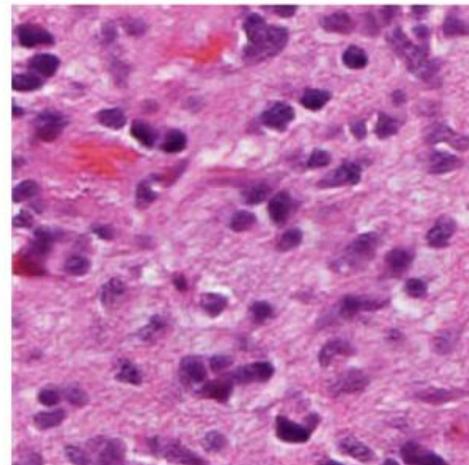
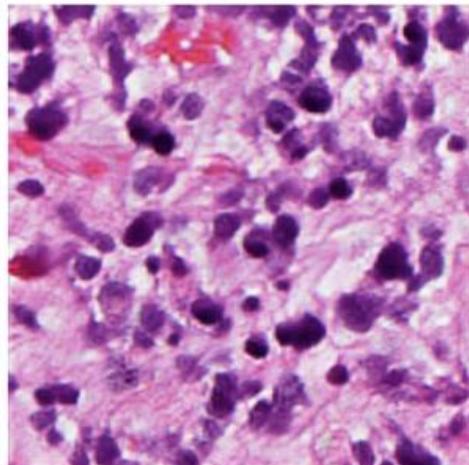
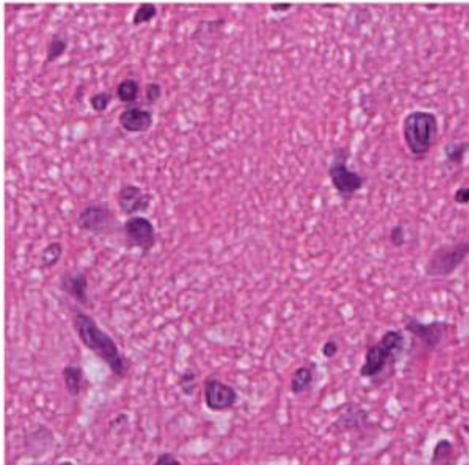


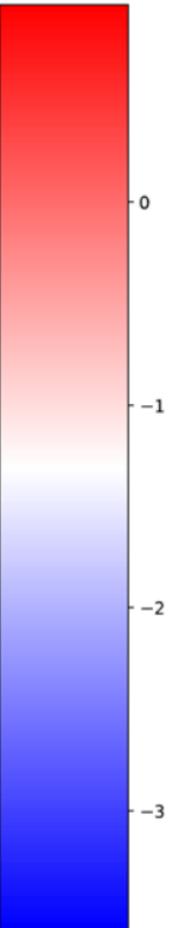
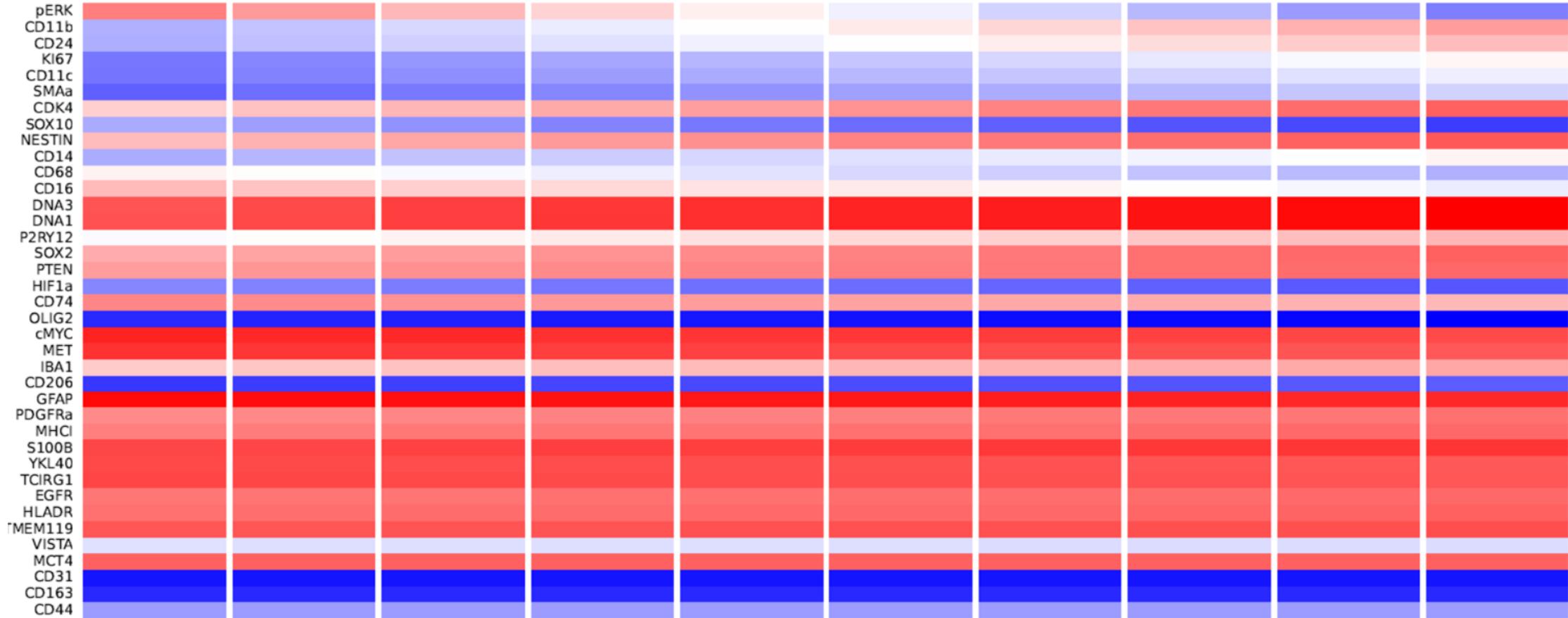
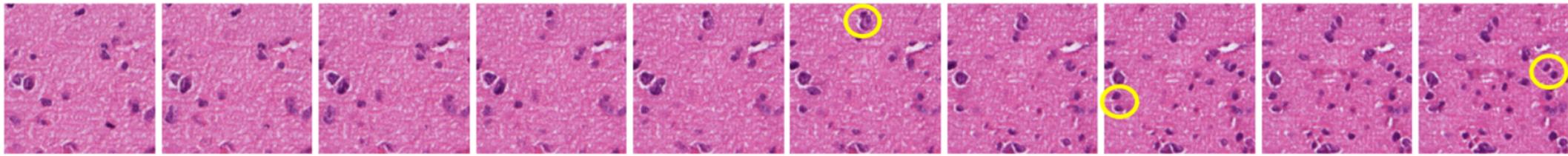
SpecTx: Generating images from protein expression

Real
Images



Synthetic
Images





The path for Cpath ahead



Biology easily has 500 years of exciting problems to work on. It's at that level. (Knuth)

- **5 Pathologist Nobel Prize Winners [1]**

- **Johannes Fibiger** received the Nobel Prize in Physiology or Medicine in 1926 for the experiments in which he produced gastric carcinoma in rats by feeding them *Spiroptera*-infected cockroaches.
- **George Whipple** was awarded the 1934 Nobel Prize in Physiology or Medicine for his discovery that a diet rich in liver cured pernicious anemia.
- **Renato Dulbecco** was awarded the 1975 Nobel Prize for discovering by molecular techniques that the genetic material of viruses was incorporated into the genetic material of the transformed cells and represented the first phase in carcinogenesis.
- **Baruj Benacerraf** received the 1980 Nobel Prize in Physiology or Medicine for his discovery of immune response genes.
- **John Warren** received the Nobel Prize in Physiology or Medicine in 2005 in recognition of this discovery of the role of *H. Pylori* in ulcers.



[1] <http://www.annclinlabsci.org/content/39/2/196.long>

[2] Bernal, J *The Social Importance of Science* 1939 p. 9

