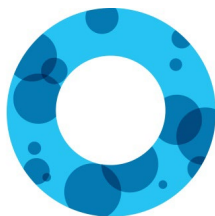


Standardisation of surgical pathology annotation libraries – a proposal

Professor David Snead
Consultant Pathologist UHCW NHS Trust
Professor Pathology University of Warwick
Director PathLAKE
www.pathlake.org

Conflict of interests

Co-founder and shareholder of Histofy



PathLAKE
Computational Pathology Excellence



PathLAKE

Partners in Computational Pathology Excellence

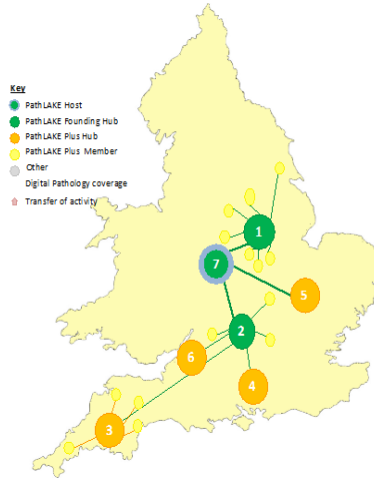
Who we are

- 14 NHS, University and Industry partners
- Original investment >£15m (Jan 2019 to Mar 2023)
 - UKRI grant funding £10m
 - Philips investment £4m
 - Matched funding from industry and academic partners
- OLS Scale-up funding £13.5m (Jan 2020 to Sept 2023)

Our objectives

- Digitise 22 NHS labs
- Create a data lake of annotated WSI & metadata
- Facilitate the development of AI tools
- Implement AI into clinical practice

PathLAKE Plus - Digital Pathology at scale



1. East Midlands 7 m
2. Oxford South 4.2 m
3. Peninsular group 1.7 m
4. Southampton and Dorset 1.5 m
5. Cambridge Norfolk Suffolk 3.1 million
6. Bristol Gloucester Bath pop. 1.8 m
7. Coventry Warwick 0.7M
8. Royal Marsden TBC

- All sites contribute to PathLAKE
- X4 AI solutions for all sites
- Data from AI goes into PathLAKE
- Real-world analysis of AI in practice
- Cohorts for algorithm validation

Digitise workflow and deploy AI

AI procurement

Select use cases

Divide use cases between hub trusts

Case specific specification

Framework mini competition

Each use case transferable across PathLAKE Plus

PathLAKE Plus – AI in practice at scale



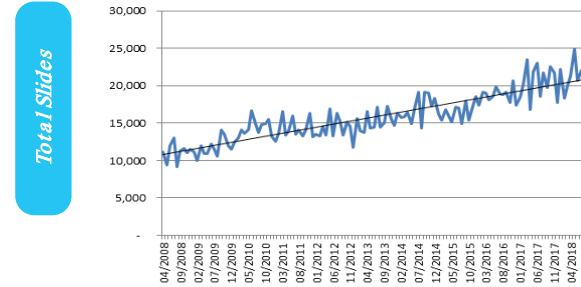
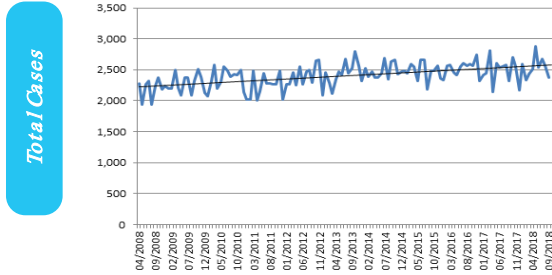
DIGITAL PATHOLOGY SOLUTIONS

HealthTrust Europe's (HTE's) Framework Agreement for Digital Pathology Solutions offers a simple and compliant route to market for the procurement of all equipment, software and services associated with Digital Pathology to suit the needs of each Participating Authority or network. This includes WSI Scanners, Medical Grade Workstations, Gross Imaging Equipment and Software, Microscope Cameras, Image Management Software and AI and Image Analysis Tools as well as other supporting solutions.

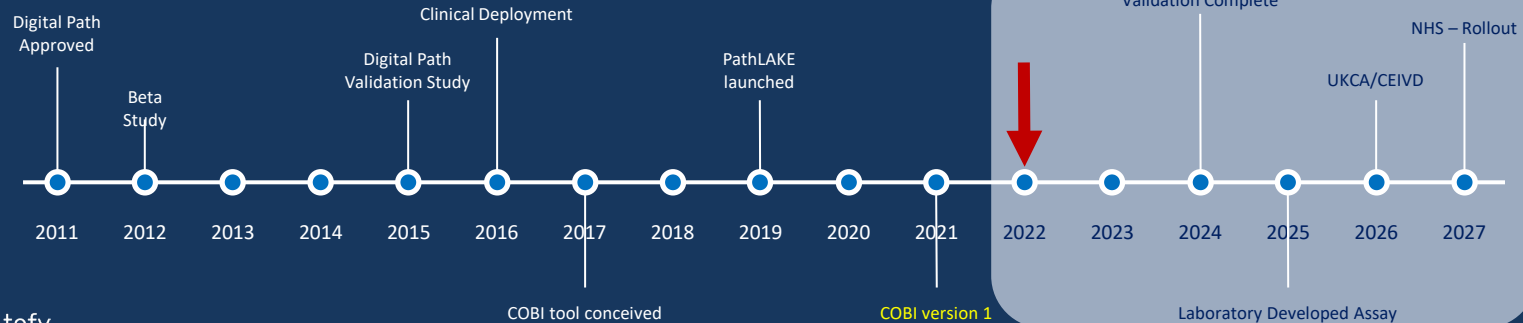
PathLAKE Plus AI requirements

- **Prostate cancer**
 - Tumour detection
 - Tumour grading
 - Tumour volume and number of cores affected
 - Perineural invasion
- **Lymph node metastasis detection**
 - Metastasis detection all solid tumours any node site
 - Micro metastases, and isolated tumour cells
- **Breast cancer**
 - Tumour detection
 - Tumour grading
 - ER, PR, HER2, Ki 67 scoring
- **Gastric cancer**
 - Tumour detection
- **Skin**
 - Melanoma versus naevus detection
 - Common tumour detection
 - Tumour depth, volume, margin clearance
 - Vascular invasion perineural invasion
- **Ki 67**
- **PD-L1**
- **Regulatory approval**
 - MHRA UK/CA, CE IVD, FDA
- **Health economics data**
- **Ethnic group data**
- **Domain shift data**
- **Interoperability between platforms**
- **Deployment plan**
 - Viable deployment plan
 - Expandable to multiple sites
 - Accommodate increasing test requests
 - Aligns with multiple other AI tools
 - Aligns with workflow
- **Equal access**
 - Access for all sites
 - Same terms for grant funded self funders
- **Fixed terms for 6 years**

The COBI Journey



Colon Biopsy Screening Tool Timeline



Machine learning - to annotate or not to annotate

- Most supervised ML algorithms require lots of annotations, often a costly and laborious process
- **Annotation protocols**
 - Slide level
 - Region level
 - Cell level
- Computational ways around:
 - **Weakly supervised learning** (requires only slide- or case-level labels)
 - **AI-assisted annotations**
 - **AI- derived annotations**
 - **Synthetic images, with annotations**



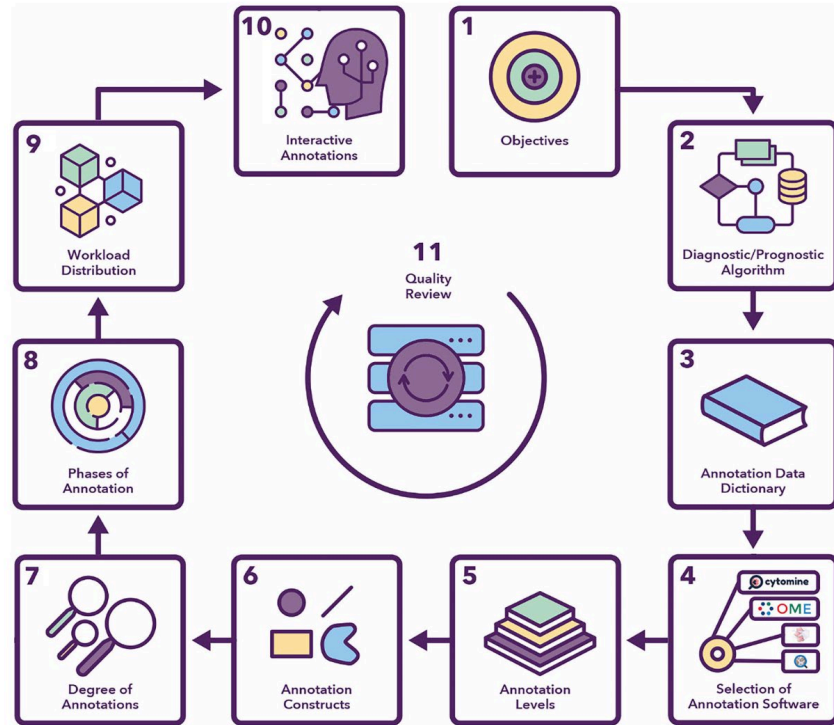
Annotation Protocols

What to annotate?

How to annotate?

How to QC annotations?

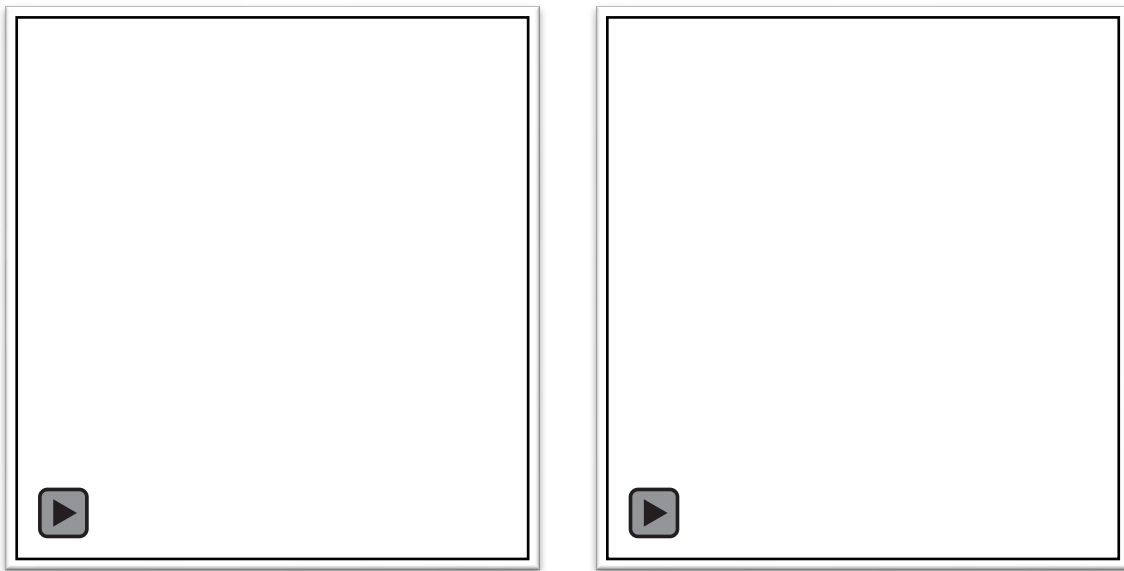
How much to annotate?



Wahab *et al.* (*J Clin. Pathol. Res.* Dec 2021)



Interactive Region Segmentation

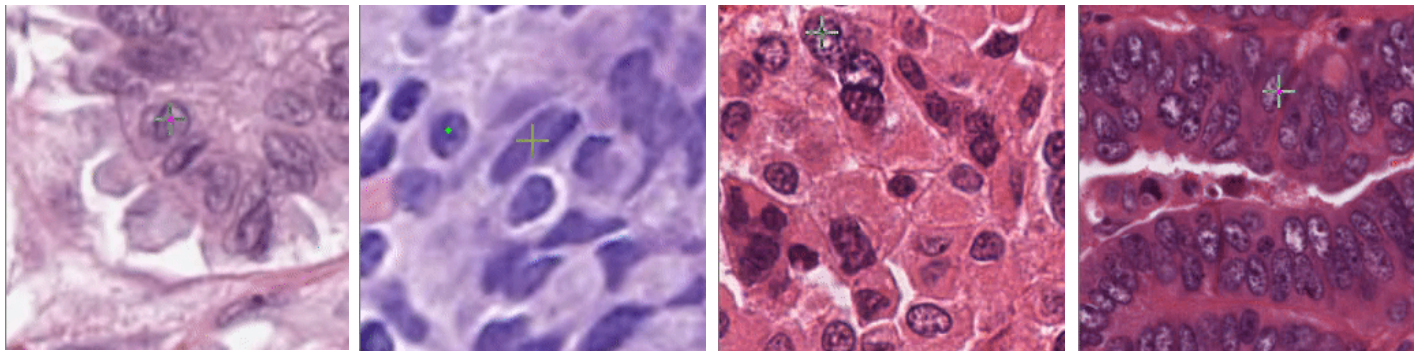


Jahanifar *et al.*, *ICCV CDPATH*, 11 Oct 2021 **Best Paper Award**



NuClick: AI-Assisted Annotations

PanNuke: A dataset of **200K+ pan-cancer nuclei** for nuclear segmentation and classification, where **NuClick** was used for collecting the segmentation masks

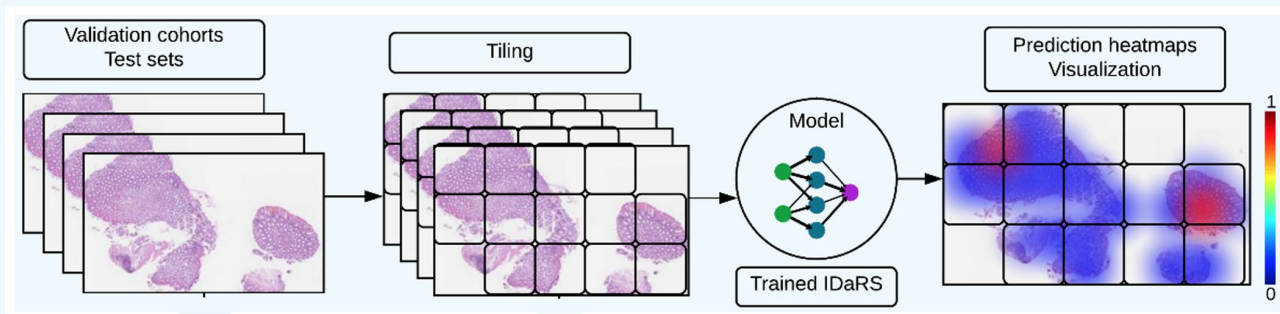


Koohbanani *et al.*, *Medical Image Analysis* (Oct 2020) & Gamper *et al.*, *Arxiv* (Apr 2020)

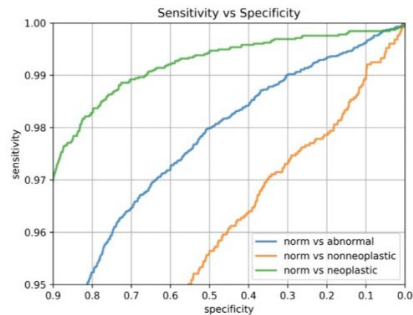
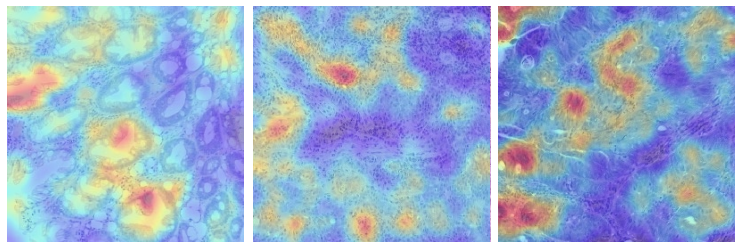
https://warwick.ac.uk/fac/cross_fac/tia/software/

COBix version weakly supervised

Iterative Draw and Ranks Sampling (IDARS)



IDaRS (Bilal *et al.*, *Lancet Digital Health*, Oct 2021)

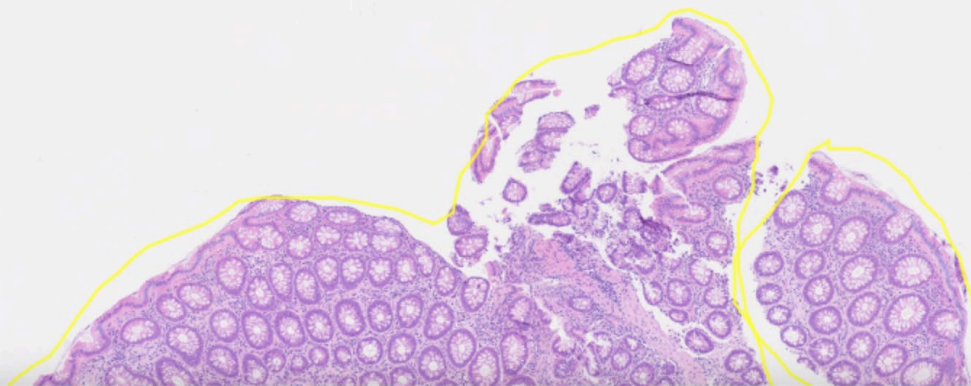


AUROC values

Normal vs Abnormal: **0.96**

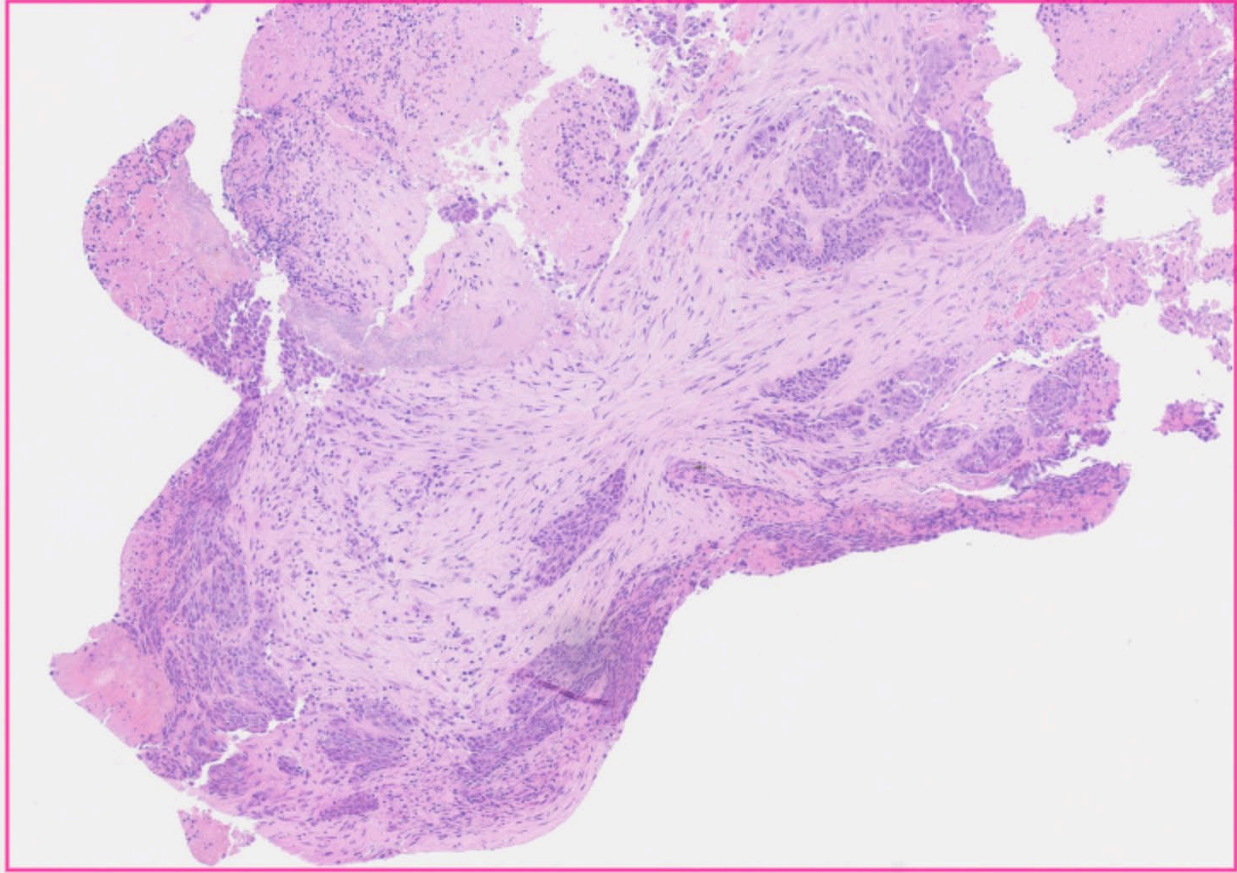
Normal vs Neoplastic: **0.99**

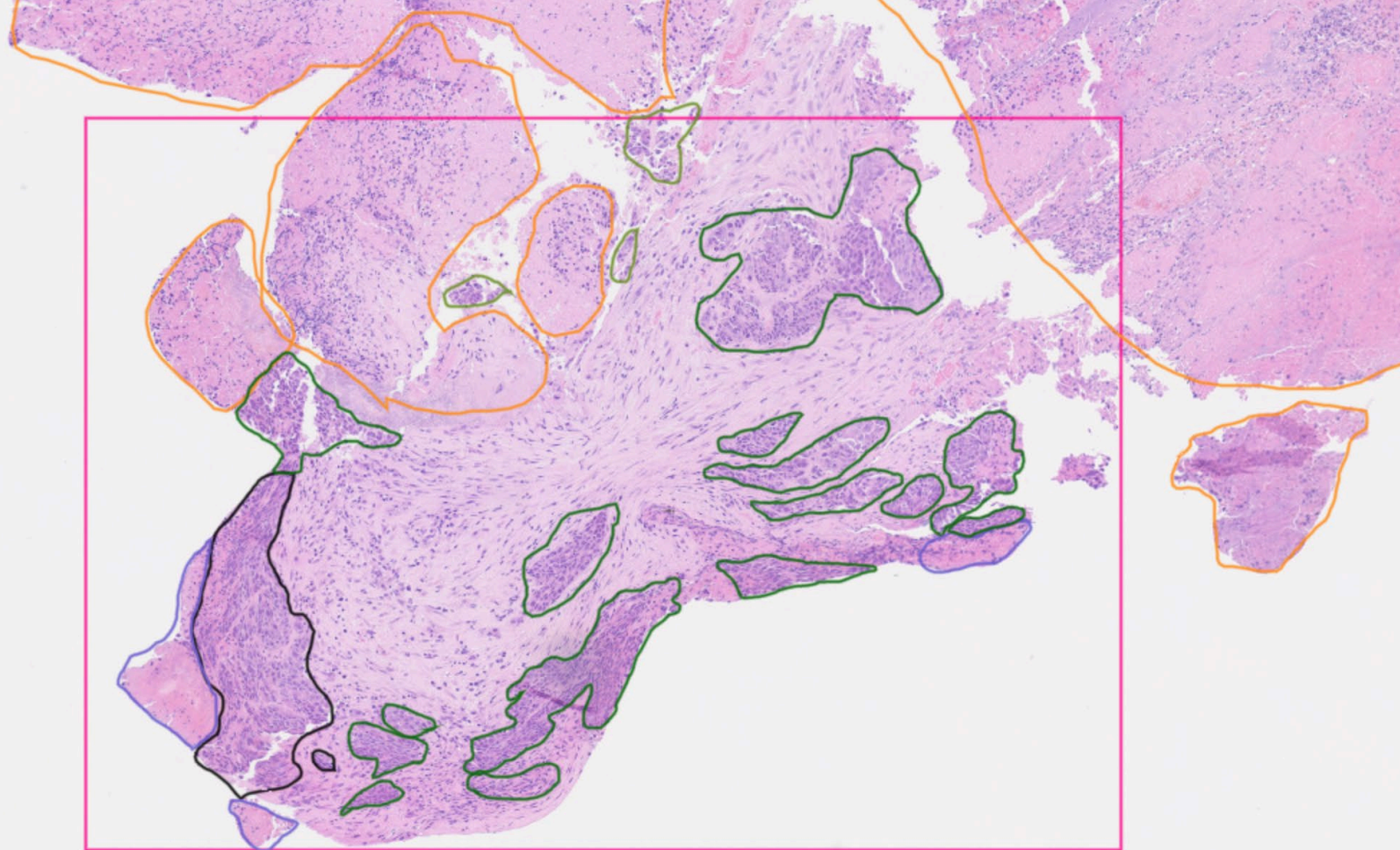
Normal vs Non-Neoplastic: **0.94**

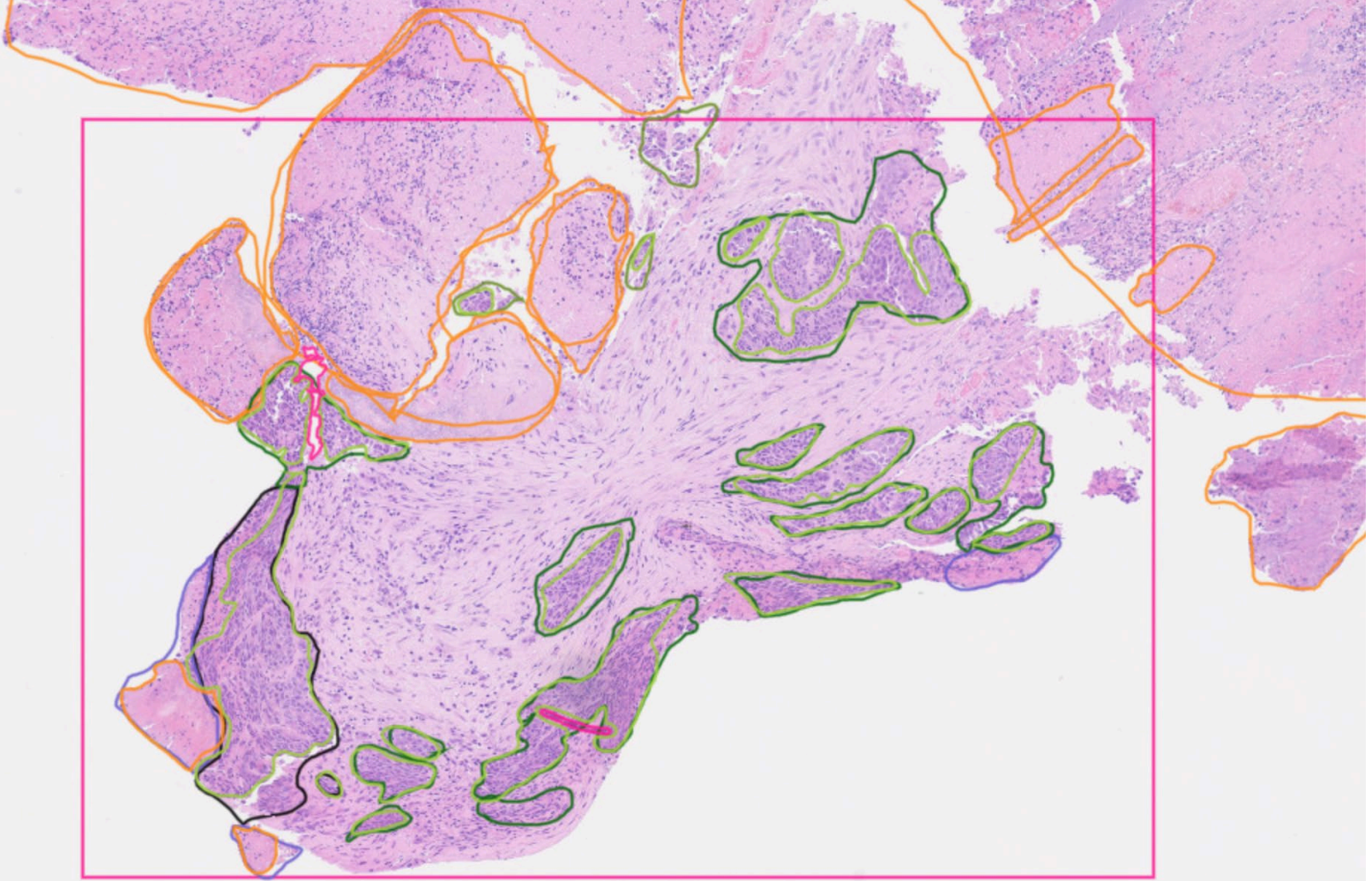


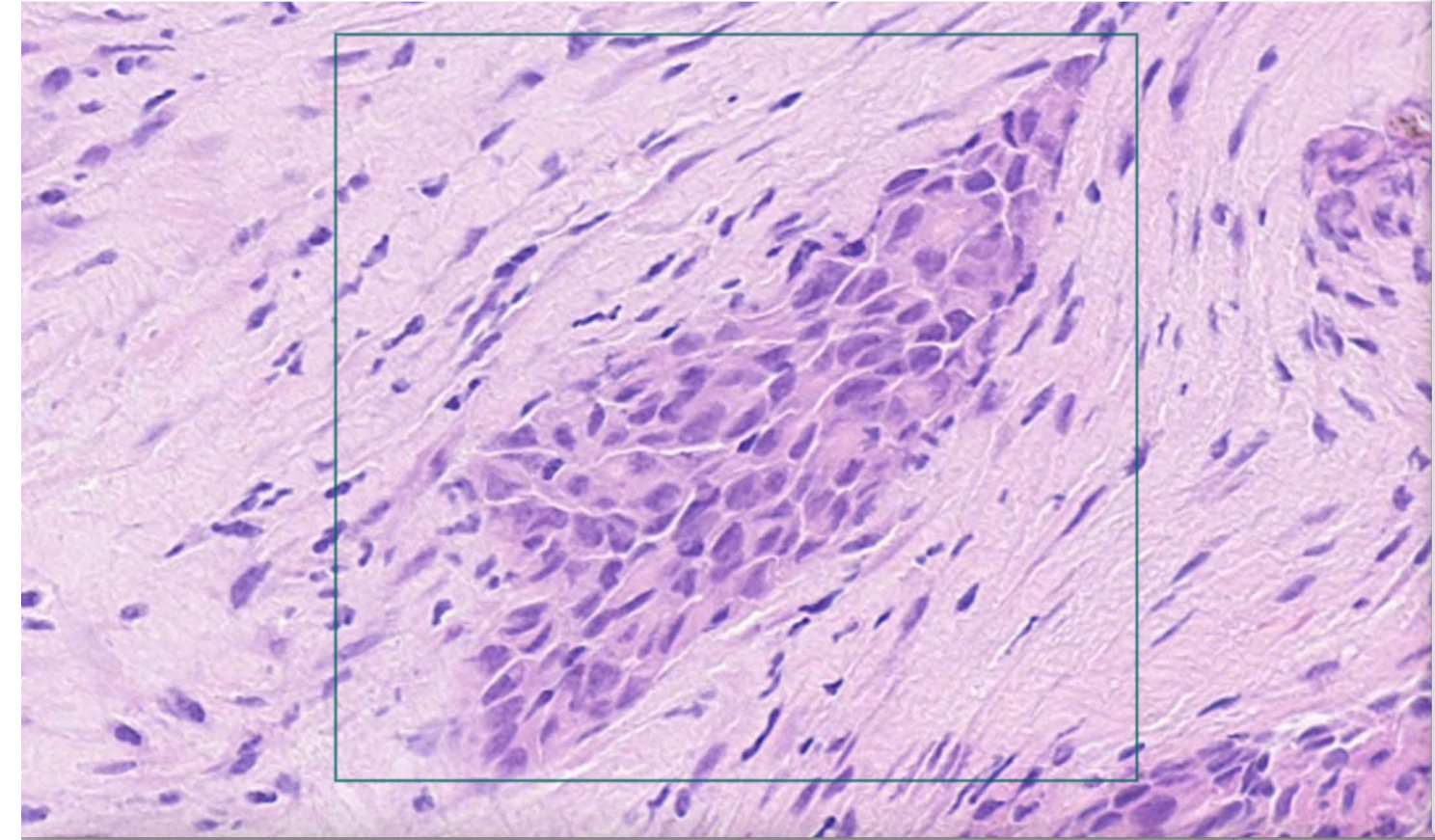
- box_special_5X
- cell_Amoeba
- cell_Apoptosis
- cell_Artifact
- cell_Bacteria
- cell_Endothelial
- cell_Eosinophils
- cell_Epithelial_hyperplastic
- cell_Epithelial_neoplastic
- cell_Epithelial_normal
- cell_Epithelial_reactive
- cell_Fibroblasts
- cell_Ganglion
- cell_Giant
- cell_Lipofuscin_within_macrophages
- cell_Lymphocytes
- cell_Macrophages
- cell_Mast cells
- cell_Mitosis
- cell_Neoplastic(other)
- cell_Neuroendocrine
- cell_Neutrophils
- cell_Paneth
- cell_Plasma
- cell_Red_blood
- cell_Signet ring
- cell_Smooth_muscle
- cell_Squamous
- cell_Unknown
- cell_Viral_inclusions**
- ✓ default
- region_Adenocarcinoma



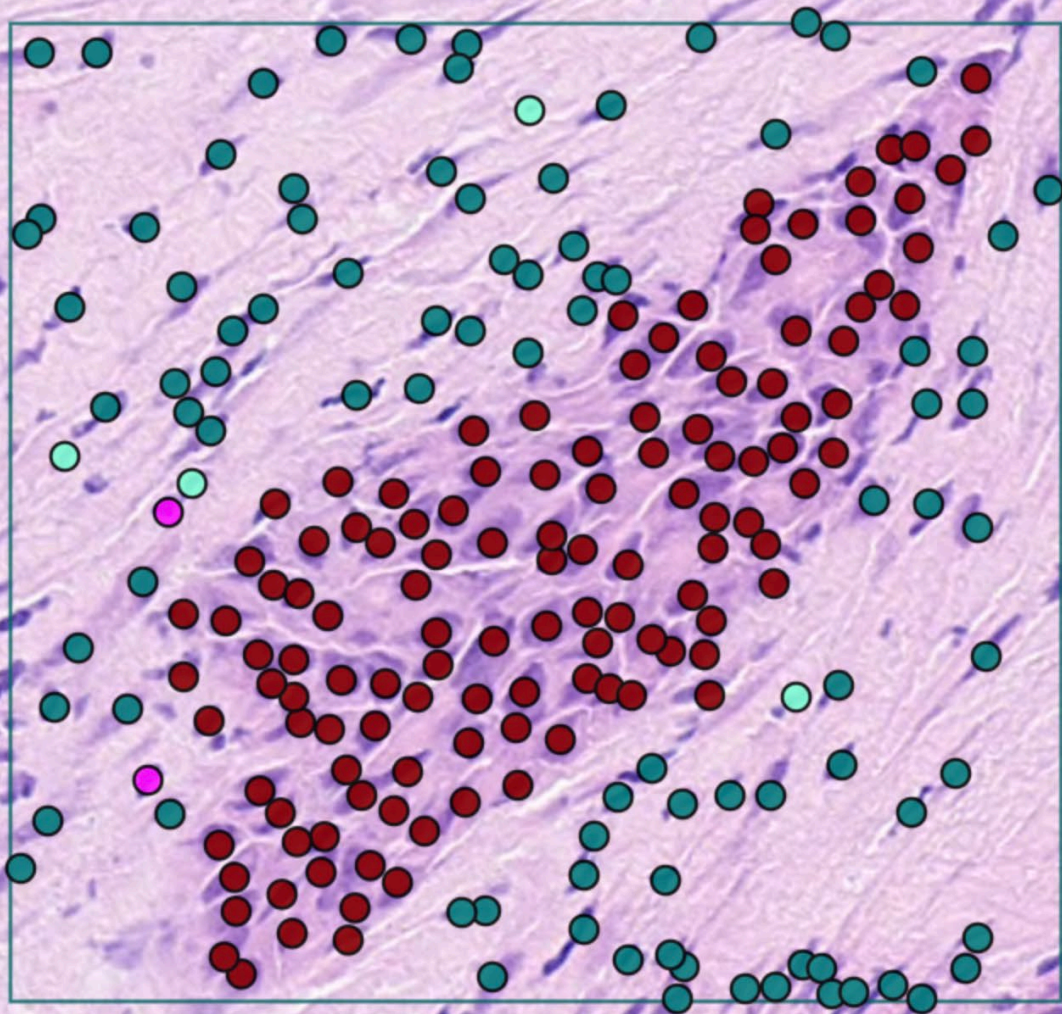


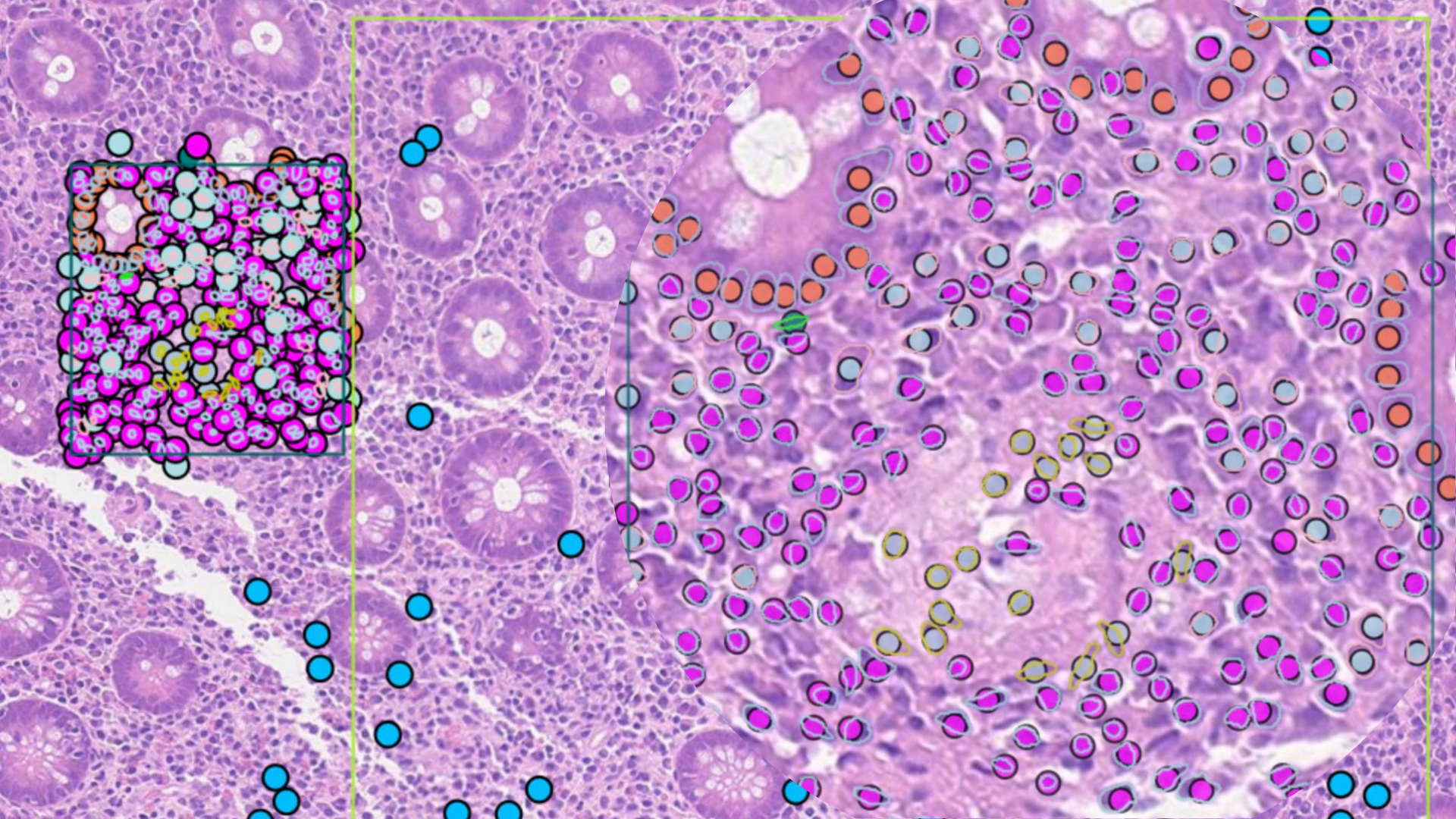


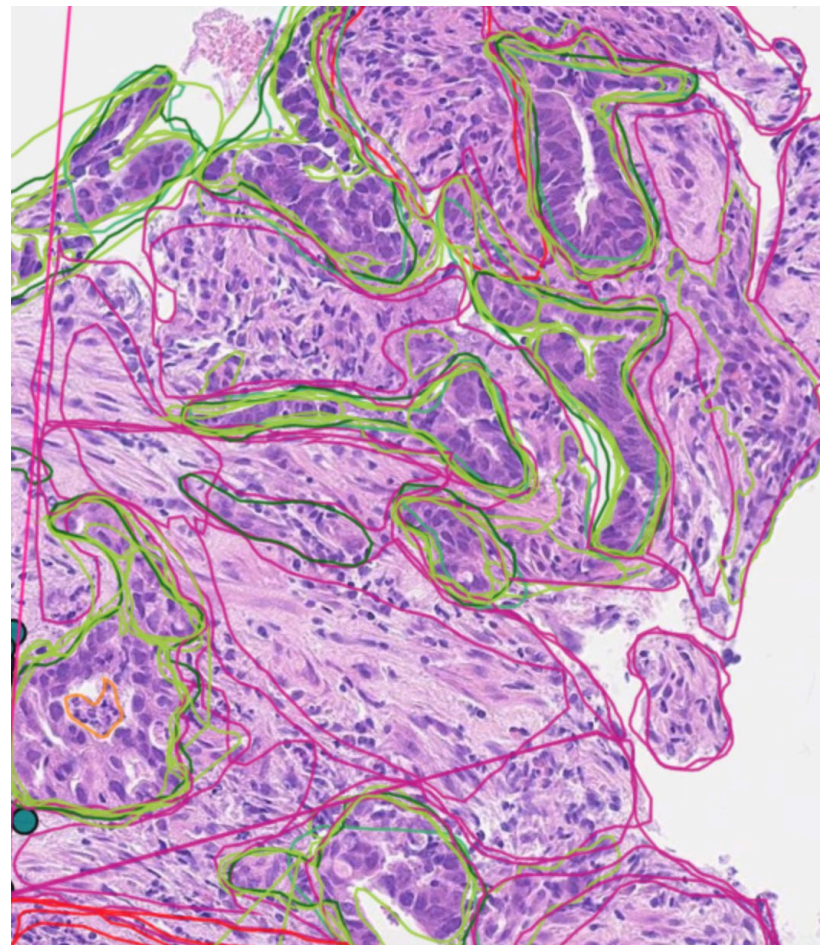
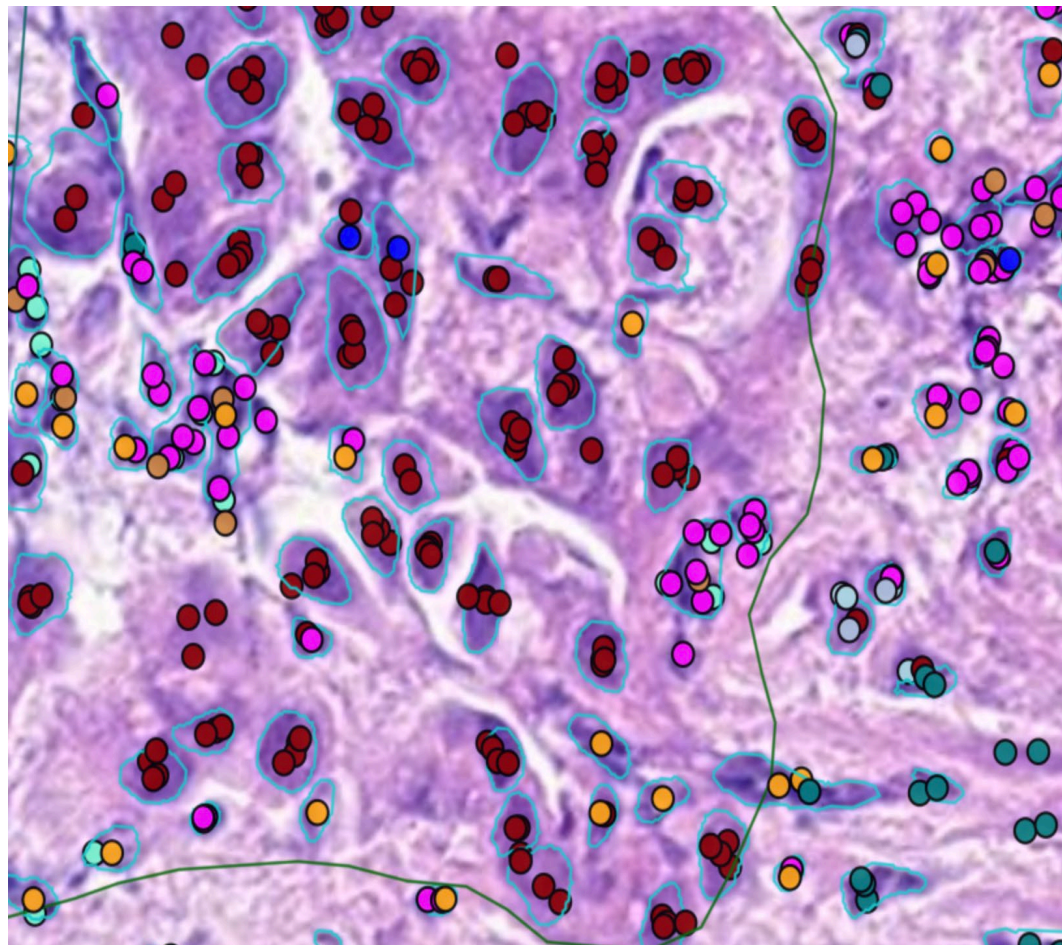




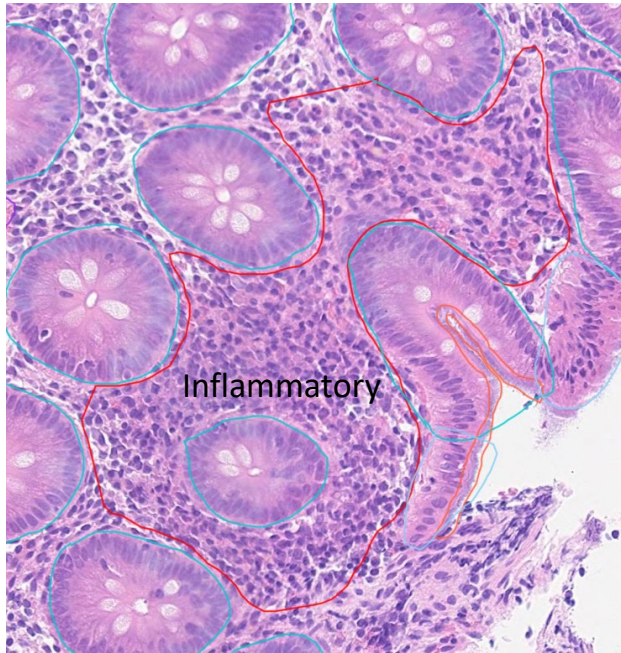
- box_ALL_special_20X
- box_ALL_special_5X
- box_cell_20X
- box_region_5X
- box_special_20X
- box_special_5X
- cell_Amoeba
- cell_Apoptosis
- cell_Artifact
- cell_Bacteria
- cell_Endothelial
- cell_Eosinophils
- cell_Epithelia_hyperplastic
- cell_Epithelia_neoplastic
- cell_Epithelia_normal
- cell_Epithelia_reactive
- cell_Fibroblasts
- cell_Ganglion
- cell_Giant
- cell_Lipofuscin_within_macrophages
- cell_Lymphocytes
- cell_Macrophages
- cell_Mast cells
- cell_Mitosis
- cell_Neoplastic(other)
- cell_Neuroendocrine
- cell_Neutrophils
- cell_Paneth
- cell_Plasma
- cell_Red_blood
- cell_Signet ring
- cell_Smooth_muscle
- cell_Squamous
- cell_Unknown
- cell_Viral_inclusions
- default
- region_Adenocarcinoma



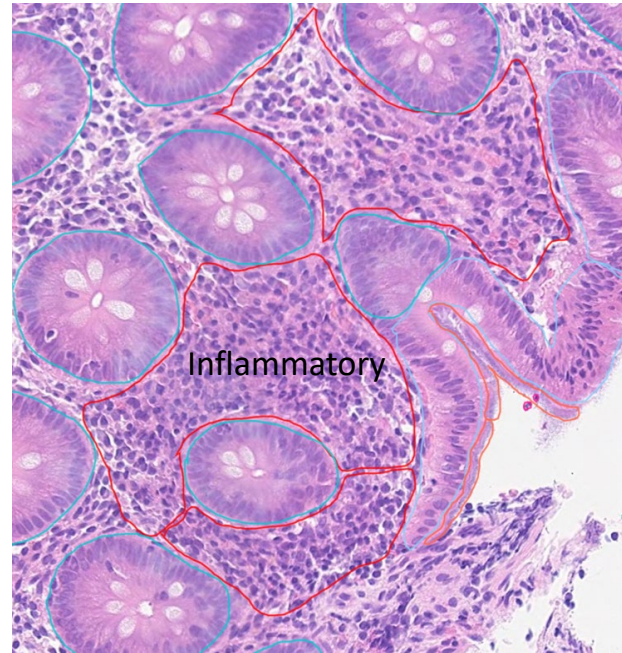




Good concordance

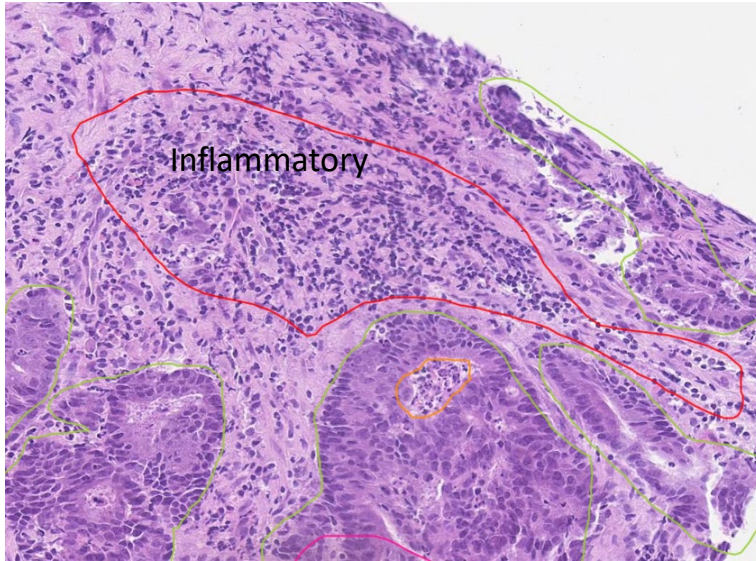


Pathologist 1

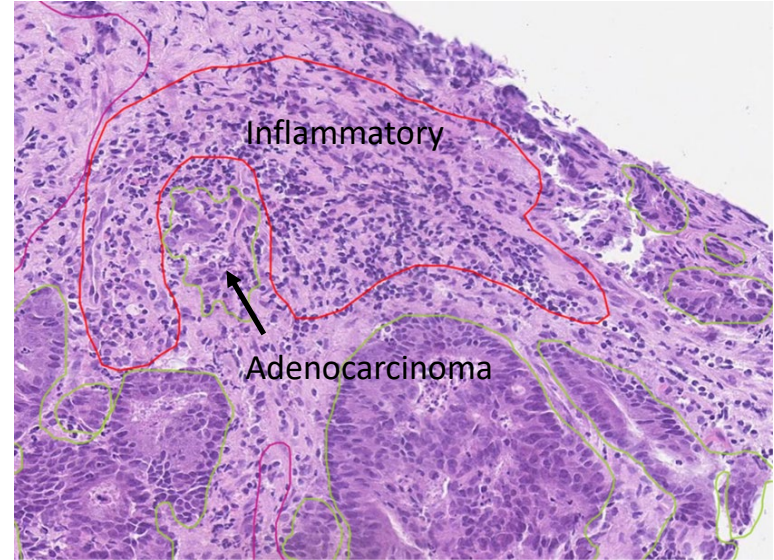


Pathologist 2

Poor concordance

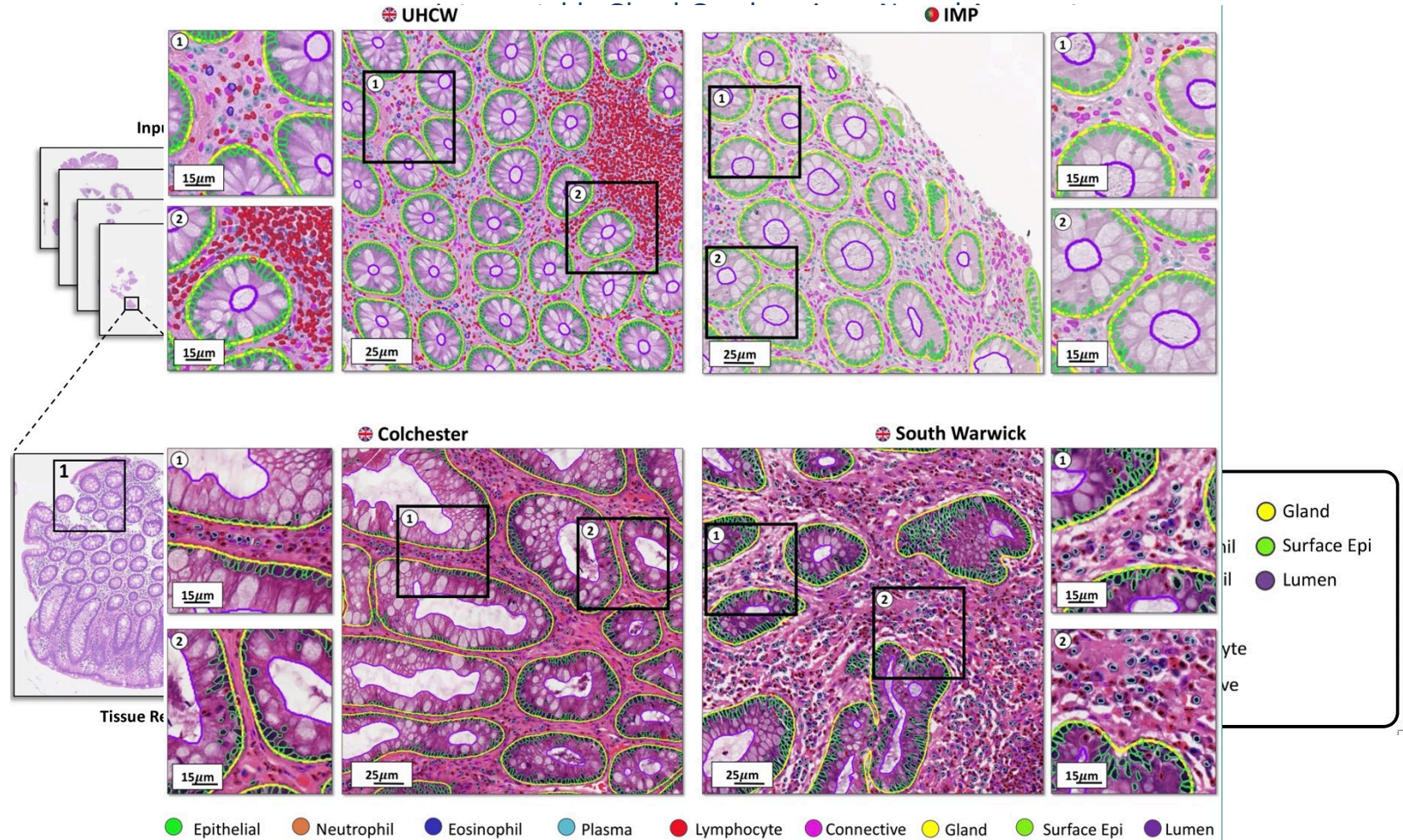


Pathologist 1

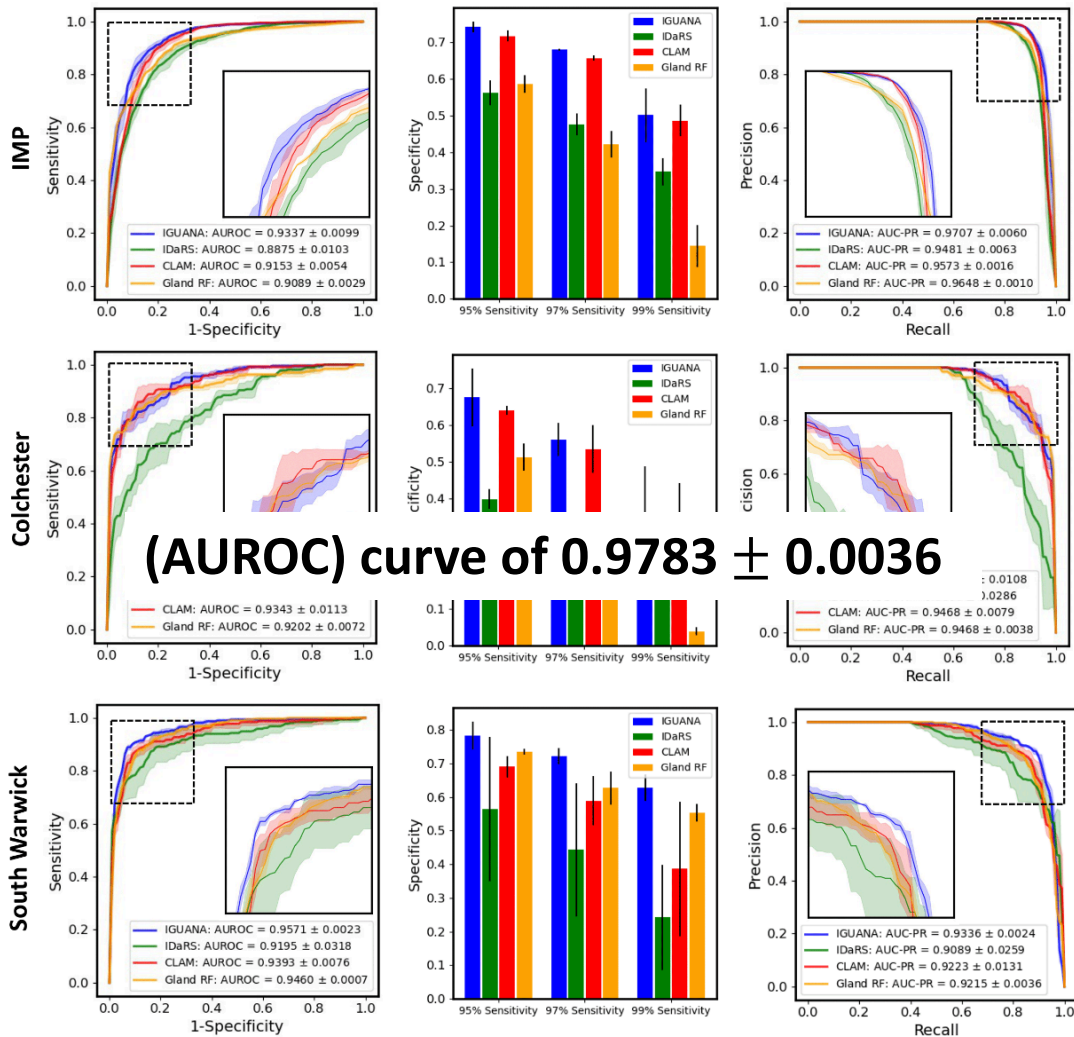


Pathologist 2

Supervised learning approach



Independent Validation



Annotations for cellular pathology

- Regions of interest
- Key diagnostic areas
- Key prognostic information
- Areas of uncertainty
- Counting of cells
- Counting mitoses
- Record observations
- Highlight important features
- MDT review
- Peer review and discussion
- Measure parameters needed
- Case summary

Why standardize?

- Improves communication
- More efficient slide review
- Appropriate tool for the task
- Better designed tools improves pathologist's experience
- Translation between systems
 - Improve interoperability
- Simplifies selection of slides for research

How do we do it?

- Construct a library of annotations
- Define the annotation needed for each task
- Assign motif design and size
- Define additional functionality needed
- Allocate SNOMED CT code
- Allocate colour with Hex code
- Prevent / eliminate duplication
- Publish standard

Cellular pathology annotation requirements

Annotation group	Motif pattern	units
Area measurement	polygon	mm ²
Linear measurement	line	mm
Region of interest	ellipse	N/A
Anatomical boundary breach	ellipse	N/A
Surgical margin	ellipse	N/A
≤ Cell level object	circle	N/A
> Cell level object	circle	N/A

Options

Search Mode: Partial matching search mode

Status: Active components only

Group by concept

Filter results by Language

english 47

Filter results by Semantic Tag

finding 31

disorder 10

observable entity 6

Filter results by Module

SNOMED CT core module (core metadata concept) 47

Type at least 3 characters ✓ Example: *shou fra*

lymph node metastasis

47 matches found in 0.268 seconds.

No regional lymph node metastasis	No regional lymph node metastasis (finding)
Metastasis to cervical lymph nodes	Metastatic malignant neoplasm to lymph nodes of neck (disorder)
Metastasis to multiple lymph nodes	Metastatic malignant neoplasm to multiple lymph nodes (disorder)
Metastasis to upper limb lymph node	Metastatic malignant neoplasm to lymph nodes of upper limb (disorder)
Metastasis to head and neck lymph node	Metastasis to head and neck lymph node (disorder)
Regional lymph node metastasis present	Regional lymph node metastasis present (finding)
No regional lymph node metastasis (finding)	No regional lymph node metastasis (finding)
Left regional lymph node metastasis present	Left regional lymph node metastasis present (finding)
Metastasis to lymph node of unknown primary	Metastasis to lymph node of unknown primary (disorder)

Parents

- SNOMED CT Concept (SNOMED RT+CTV3)
- Clinical finding (finding)
- Finding of lesion (finding)
- Tumor finding (finding)

Stated Inferred

Regional lymph node metastasis present (finding)

SCTID: 399374009

399374009 | Regional lymph node metastasis present (finding) |


- Regional lymph node metastasis present (finding)
- Regional lymph node metastasis present
- Regional lymph node involvement present

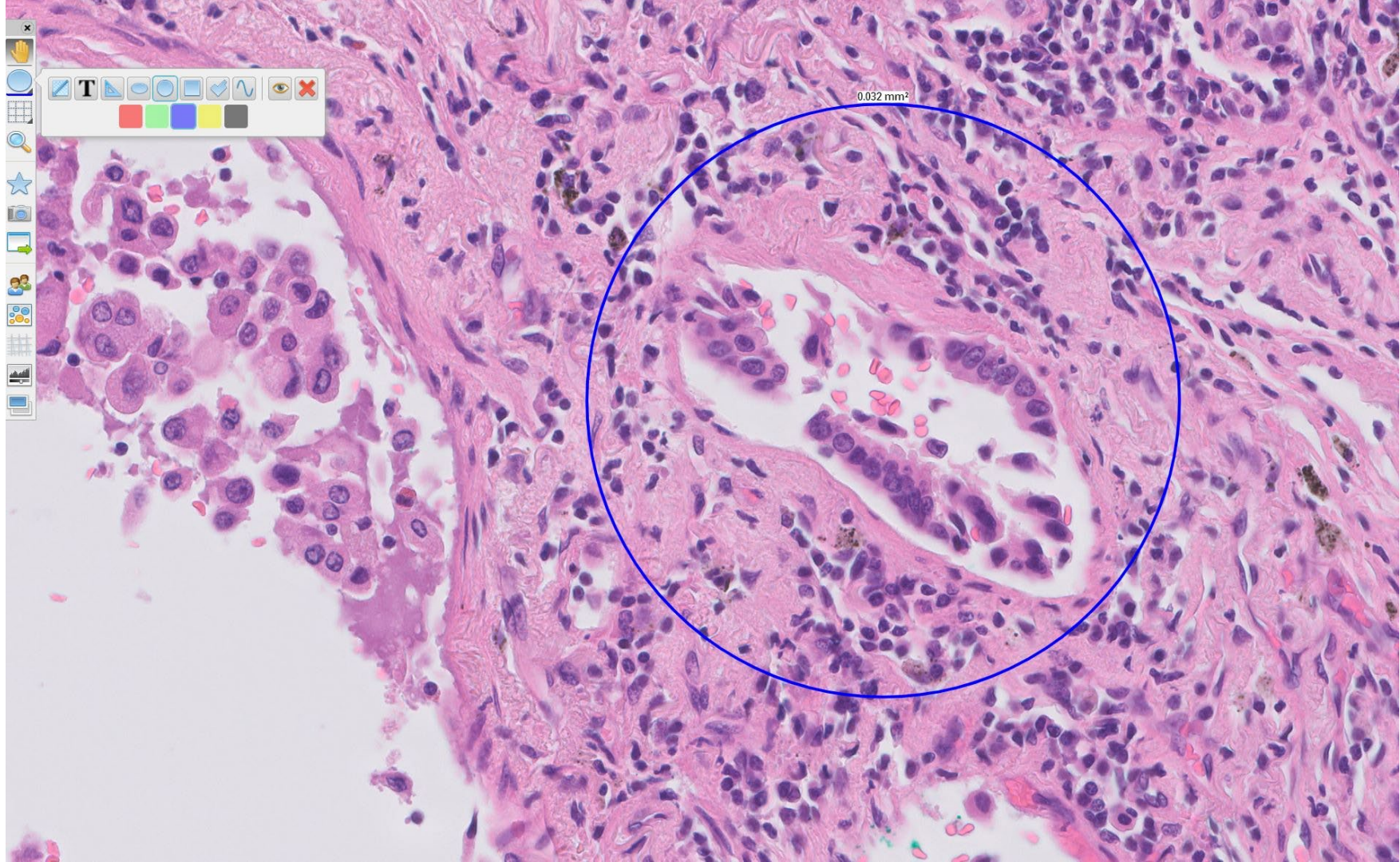
Associated morphology → Lesion

Interprets → Lesion observable

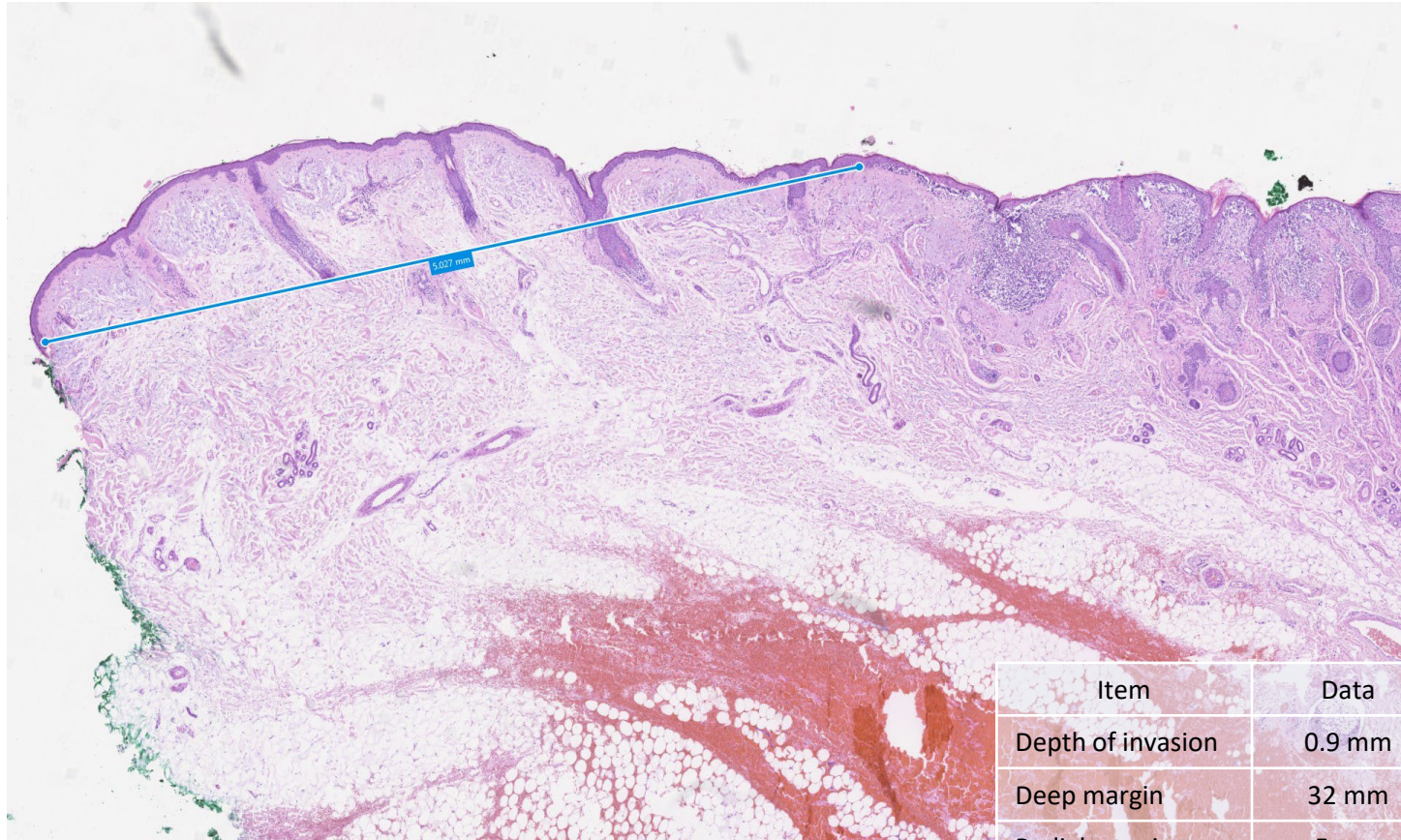
Children (2)

- Left regional lymph node metastasis present (finding)
- Right regional lymph node metastasis present (finding)

1	Annotation Dictionary	Column1	Column2	Column3	Column4	Column5	Column6	Column7
2	Universal tumour related objects	SCTID (SNOMED CT CODE)	Home tab	Annotation sha	Colour	Colour name	Hex code #RRG	Decimal Code R,G,B
3	Tumour length	372299002	tumour	line		Golden Poppy	#FCC200	252,194,0
4	Tumour diameter / depth of invasion	396237006	tumour	line		Cyan Process	#00B7EB	0,183,235
5	Tumour area / volume	258261001	tumour	polygon		GO Green	#00AB66	0,171,102
6	Invasive carcinoma	722688002	tumour	polygon		Red	#FF0000	255,0,0
7	Adenocarcinoma in situ	51642000	Respiratory	Polygon		Classic Rose	#FBCCE7	251,204,231
8	Vascular invasion	372287009	tumour	ellipse		medium blue	#0000CD	0,0,205
9	perineural invasion	369731000	tumour	ellipse		rebeccapurple	#663399	102,51,153
10	Mitosis	396447006	tumour	7 micron circle		dark yellow 1	#CCCC00	204,204,0
11	atypical mitosis	27159004	tumour	7 micron circle		Psychedelic Purple	#DF00FF	223,0,255
12	Serosal involvement	370062000	tumour	ellipse		Pumpkin	#FF7518	255,117,24
13	Lymph node metastasis	399374009	tumour	ellipse		Wild watermelon	#fc6c85	252, 108, 133
14	non necrotising granulomas	45647009	inflammatio	ellipse		Bright Turquoise	#08E8DE	8,232,222
15	 psis	112674009	vascular	ellipse		True Blue	#0073CF	0,115,207
16	viral inclusions	53012008	infection	Stamp		sea green	#2E8B57	46,139,87
17	Giant cells	60401004	infection	Stamp		cadet blue	#5F9EA0	95,158,160
18	Bacteria	409822003	infection	Stamp		pink	#FFC0CB	255,192,203
19	Deep margin clearance	85270005	tumour	line		Neon Green	#39ff14	57,255,20
20	Radial margin clearance	384822004	tumour	line		Maya Blue	#73c2fb	115,194,251

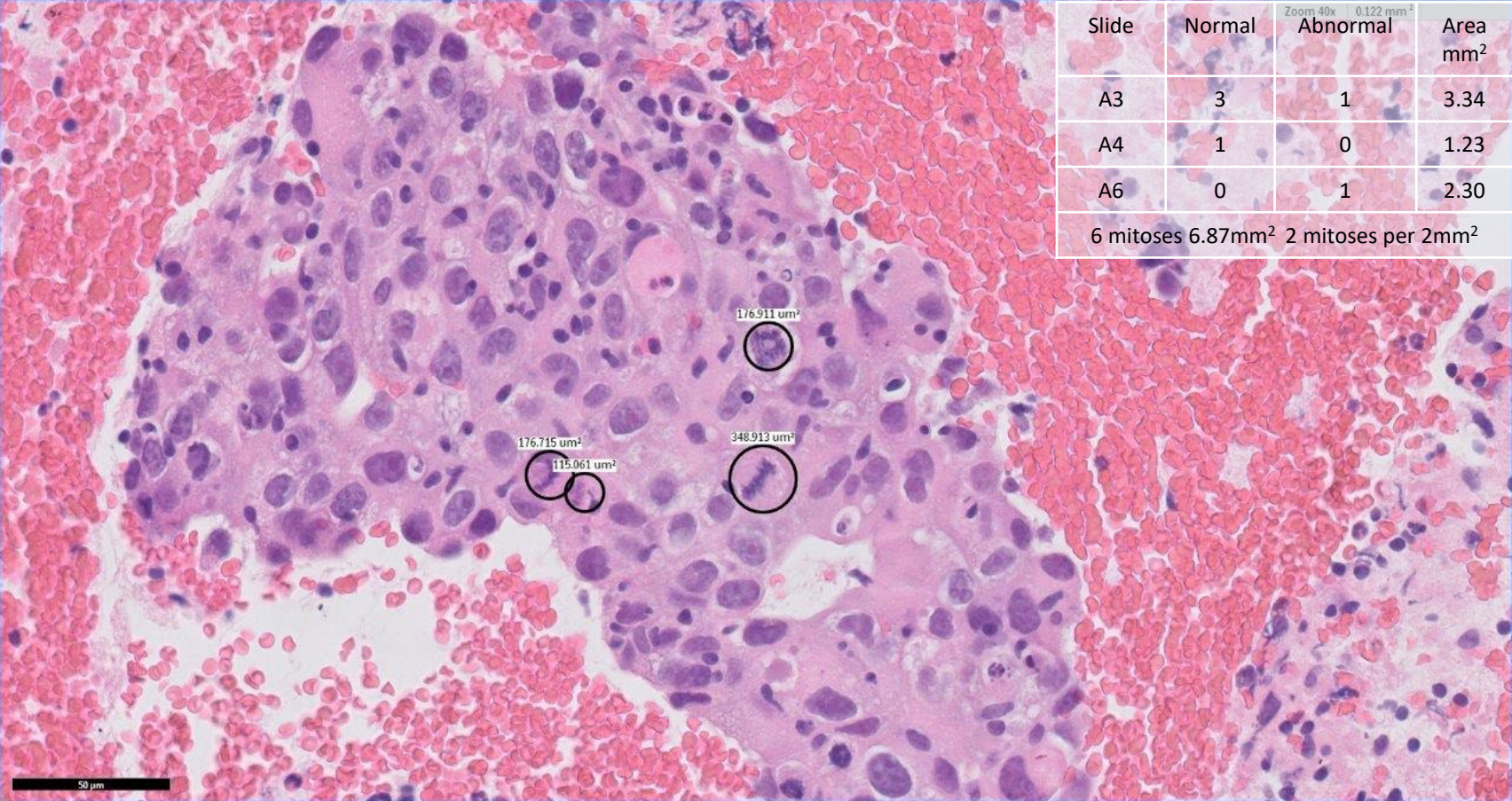


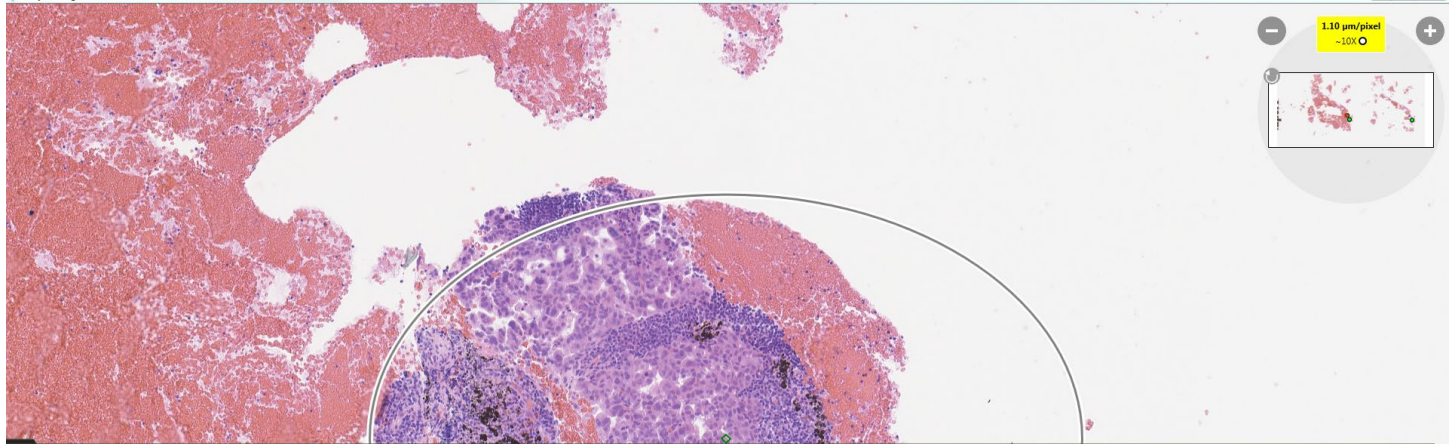
0.032 mm²



Item	Data	Slide
Depth of invasion	0.9 mm	A3
Deep margin	32 mm	A3
Radial margin	5 mm	A1
Mitotic count	6 per mm ²	A3, A4, A6

Slide	Normal	Zoom 40x 0.122 mm ² Abnormal	Area mm ²
A3	3	1	3.34
A4	1	0	1.23
A6	0	1	2.30
6 mitoses			6.87mm ² 2 mitoses per 2mm ²





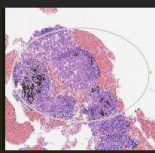
1.10 $\mu\text{m}/\text{pixel}$
~10X

Share Image

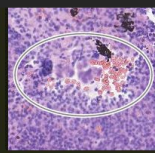
Slide Tray Annotations (3) Snapshots (0)

Part B: EBUS SPECIMEN: Cytology sample station 4L

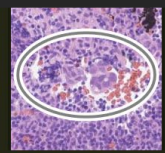
Part C: EBUS SPECIMEN station 10R



Drawn at ~10X



Drawn at ~20X



Drawn at ~10X

Station	Nodes	Involved	Slide
7	2	0	A1, A2
4L	1	0	B1
4R	1	1	C1

Lymph node stage pN2
(Lung TNM8)

Summary

- Annotations are important
- Pathologists apply them routinely at reporting (Radiologists don't)
- Standardising annotation “motifs” improves their utility
- Coding these annotations makes them retrievable
- Better design of annotation tools is attractive to pathologists
- Work with pathologists to understand their needs

Acknowledgements

- Nasir Rajpoot
- Clare Verrill
- Maneul Salto-Tellez
- Emad Rakha
- Shan Raza
- Fayyaz Minhas
- Abhir Bhalerao
- Mohsin Bilal
- Noorul Wahab
- Young Park
- Wenqi Lu
- John Euinton
- Simon Graham
- Deborah Griggs
- Rachel Flowers
- Mark Cracknell
- Andrew White
- Sean James
- Matthew Burke
- Kellie Burns
- Jennifer Samboy
- Hugh Cormican

- Steve Roberts
- Simon Jones
- Amar Kansara
- Yee Wah Tsang
- Shatrughan Sah
- Kishore Goparlakrishnan
- Hesham El Daly
- Andrew Robinson
- Sarah Read Jones
- Ayesha Azam
- Katherine Hewitt
- **Harriet Evans**
- **Emily Hero**
- Harvir Sahota
- Katherine Dodd
- Mohammed Nimir
- Liam Hems
- Abi Takyi
- Peter Whitney
- Emma Elliot
- Sasha Gill
- Ceri Jones



Innovate
UK



INDUSTRIAL
STRATEGY
Delivering growth, prosperity, opportunity

NIHR | National Institute
for Health Research