Artificial Intelligence and Pathology

Joseph Gaut, MD, PhD

WashU Medicine

Disclosures

- A portion of the presentation represents work supported in part by the MTS Foundation and the NIH. J. Gaut may receive royalty income based on technology developed by Drs. Gaut, Swamidass, and Marsh and licensed by Washington University to PlatformSTL. That technology is evaluated in this research.

- J. Gaut is listed as a co-inventor on a US patent regarding acute kidney injury biomarkers.

Outline



- What is Pathology?
- Pathology Goes Digital
- AI in Pathology
- Diagnosis
- Outcome Prediction
- Foundation Models
- Beyond Human Vision

What is Pathology?

Pathology is the study of disease. It is the bridge between science and medicine.

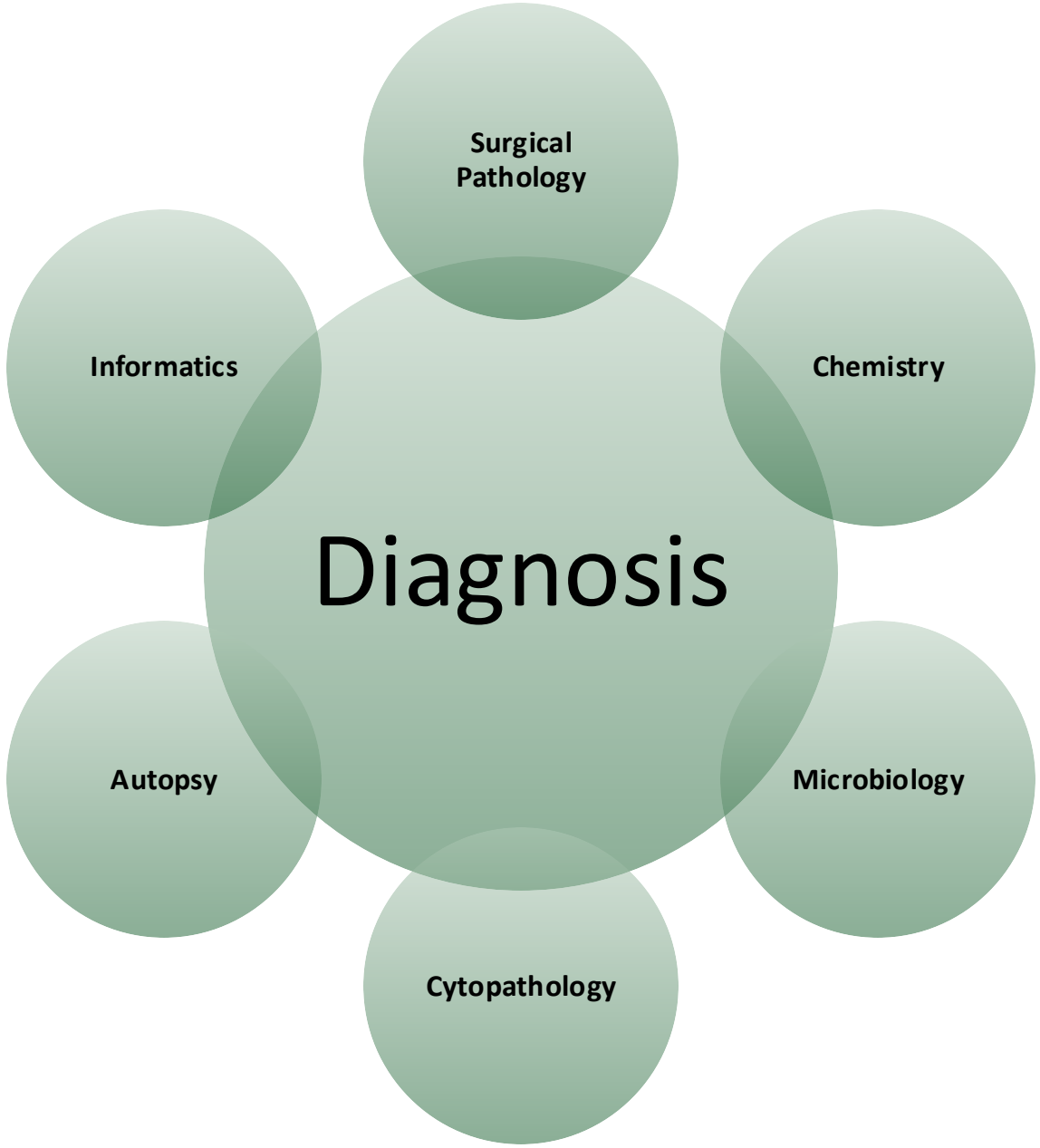
It underpins every aspect of patient care, from diagnostic testing and treatment advice to using cutting-edge genetic technologies and preventing disease.

In the U.S., pathology laboratories conduct approximately 14 billion laboratory tests every year

70 percent of medical decisions depend on laboratory test results



The art and science of predicting biologic behavior in the future by examining a snapshot of the present.


Surgical Pathology
Informatics
Chemistry
Diagnosis
Autopsy
Microbiology
Cytopathology





Pathology Subspecialties

- Anatomic
  - Surgical Pathology
  - Hematopathology
  - Dermatopathology
  - Cytopathology
  - Neuropathology
  - Pediatric Pathology
  - Autopsy Pathology
  - Forensic Pathology
  - Informatics

- Clinical (Laboratory Medicine)
  - Transfusion Medicine
  - Clinical Chemistry
  - Microbiology
  - Genetics & Genomics
  - Informatics
  - Histocompatibility



Surgical Pathology

From Tissue to Slide

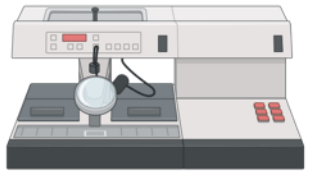
1. Obtain a biopsy sample from the patient.



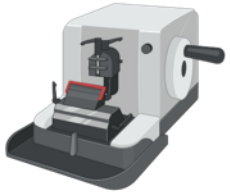
2. Fix the tissue sample in formalin.



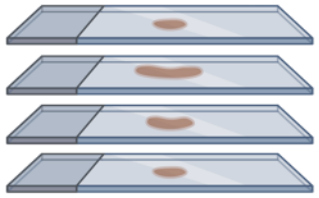
3. Dehydrate and embed the sample in paraffin.



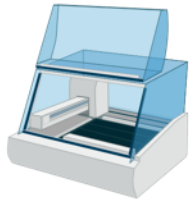
4. Section the tissue using a microtome.



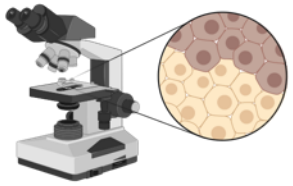
5. Mount the sections on slides.



6. Stain the slides using an auto stainer.



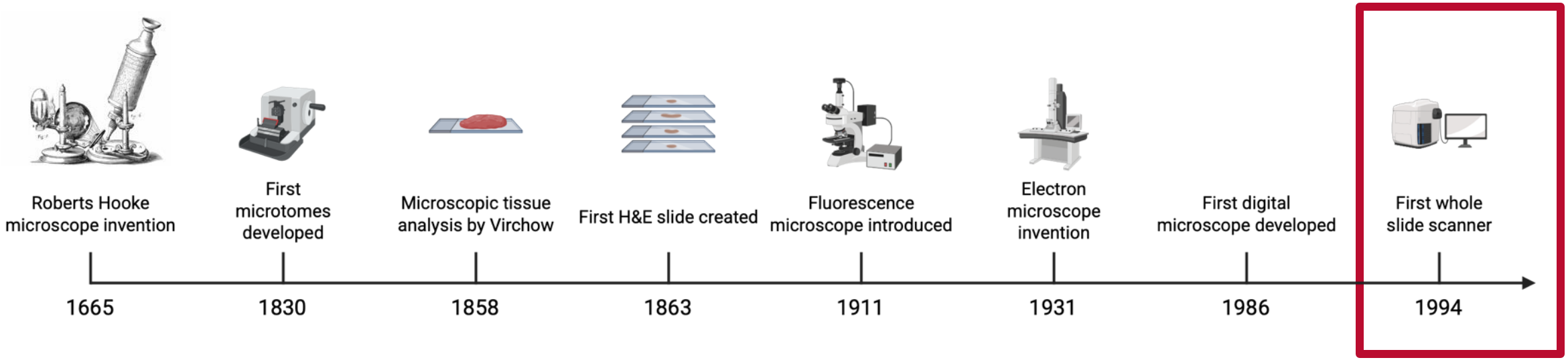
7. Examine the slides using a microscope.



8. Document and report the findings.



Pathology Imaging Evolution



Digital Pathology Workflow

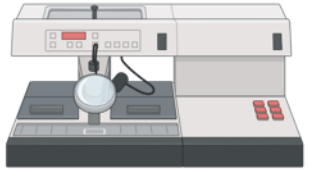
1. Obtain a biopsy sample from the patient.



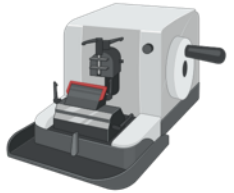
2. Fix the biopsy sample in formalin.



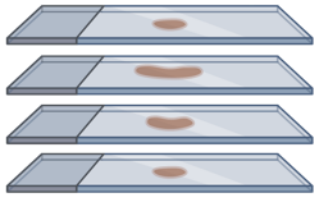
3. Dehydrate and embed the sample in paraffin.



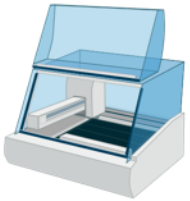
4. Section the sample using a microtome.



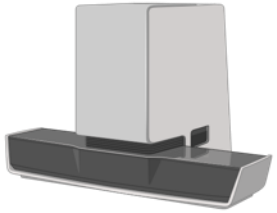
5. Mount the sections on slides.



6. Stain the slides using an auto stainer.



7. Scan the stained slides.



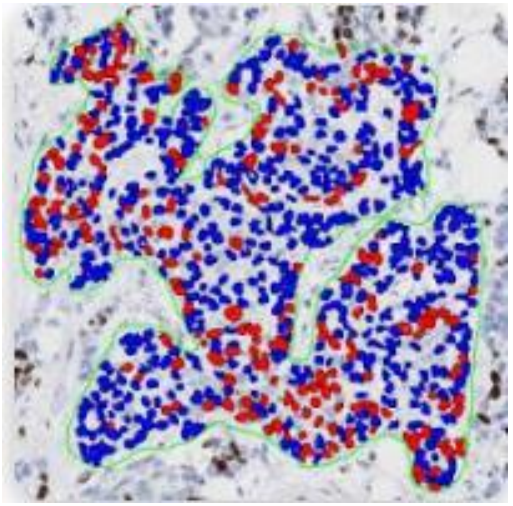
8. Review and sign out the digital slides.

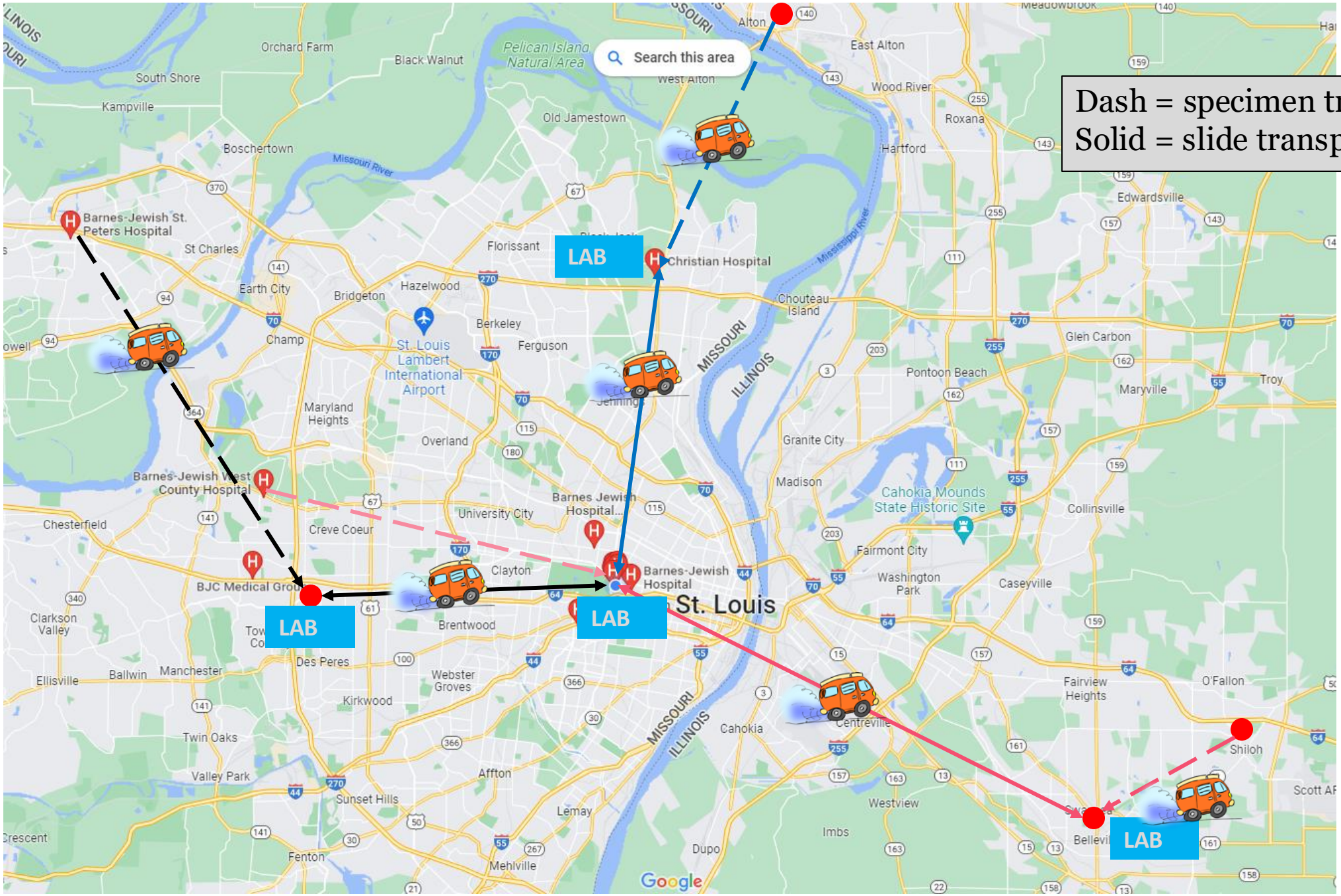


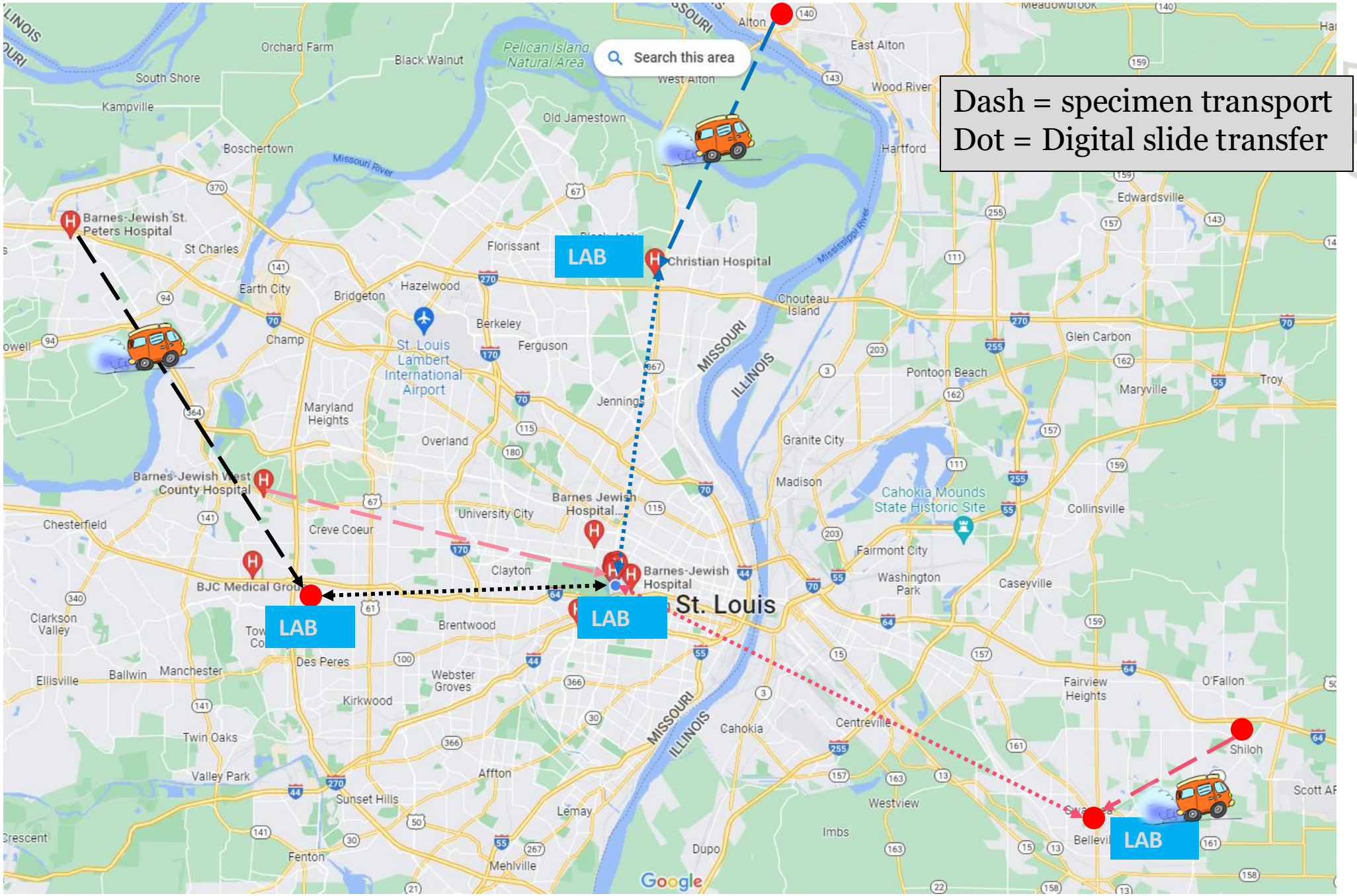
Why Digitization?

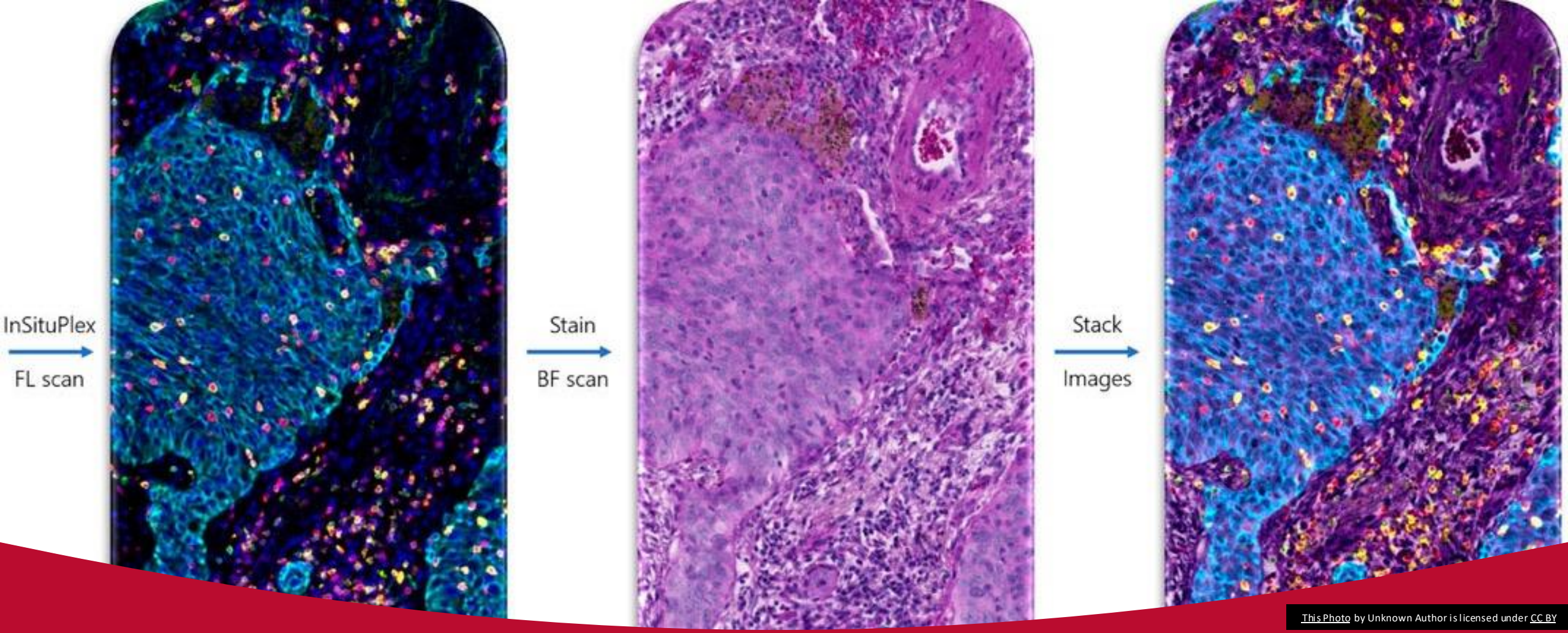
- Improved **efficiency** and **scalability**
- **Minimizes risk** of damage/loss during transport
- Better **patient care**
  - Faster results
  - Easy and fast consultation
- Enables **AI deployment**
- Improved system **integration**
- Develops infrastructure for **growth**
- Essential for **training** future pathologists






Dash = specimen transport
Solid = slide transport
LAB
LAB
LAB
LAB
Search this area
Barnes-Jewish St. Peters Hospital
Christian Hospital
Barnes-Jewish West County Hospital
Barnes-Jewish Hospital
St. Louis Lambert International Airport
St. Louis
Alton
East Alton
West Alton
Wood River
Roxana
Hartford
Edwardsville
Florissant
Hazelwood
Bridgeton
Berkeley
Ferguson
Chouteau Island
Glen Carbon
Pontoon Beach
Maryville
Troy
Granite City
Madison
Cahokia Mounds State Historic Site
Collinsville
Fairmont City
Washington Park
Caseyville
Fairview Heights
O'Fallon
Shiloh
Centreville
Cahokia
Westview
Imbs
Dupo
Lemay
Affton
Sunset Hills
Valley Park
Twin Oaks
Kirkwood
Webster Groves
Brentwood
Clayton
University City
Overland
Maryland Heights
Creve Coeur
Chesterfield
Des Peres
Manchester
Ballwin
Ellisville
Clarkson Valley
Fenton
Mehlville
Scott AF
Missouri River
Mississippi River
MISSOURI
ILLINOIS
Google


Dash = specimen transport
Dot = Digital slide transfer
LAB
LAB
LAB
Barnes-Jewish St. Peters Hospital
Christian Hospital
Barnes-Jewish West County Hospital
Barnes-Jewish Hospital
BJC Medical Group
St. Louis
St. Louis Lambert International Airport
Search this area
Pelican Island Natural Area
Missouri River
Mississippi River
MISSOURI
ILLINOIS
Google



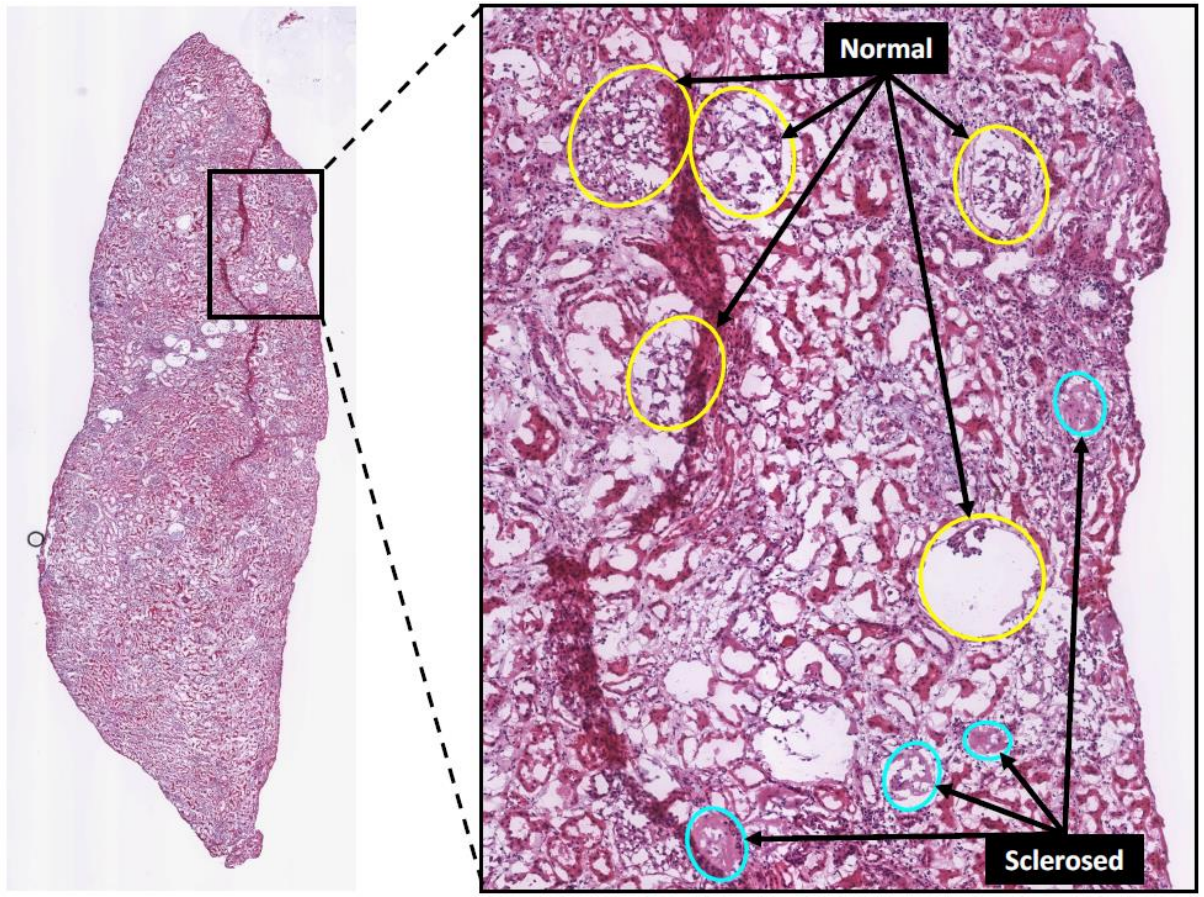
This Photo by Unknown Author is licensed under CC BY

Artificial Intelligence in Anatomic Pathology

Use Cases for AI in Anatomic Pathology

- Perform quantitative tasks
- Diagnosis
- Screening lymph nodes for metastatic cancer
- Outcome prediction
- Identify subvisual features indicative of molecular changes
- Report generation

Preimplantation Kidney Biopsies - Challenges

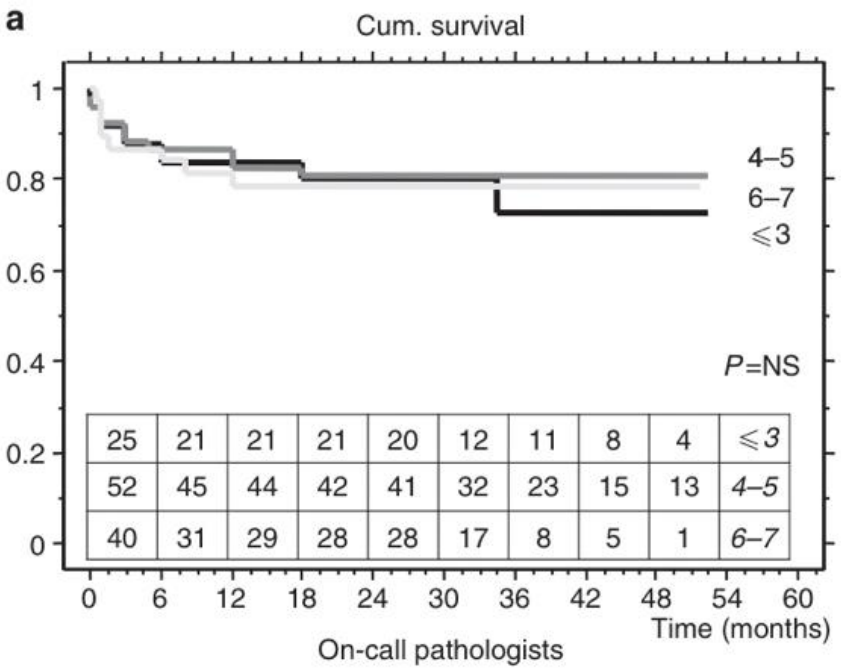


- Interpreted by **on call pathologists**
- Majority **lack expertise** in renal pathology
- Use **frozen sections** to ensure rapid TAT

kidney INTERNATIONAL ISN

Donor kidney biopsies: pathology matters, and so does the pathologist

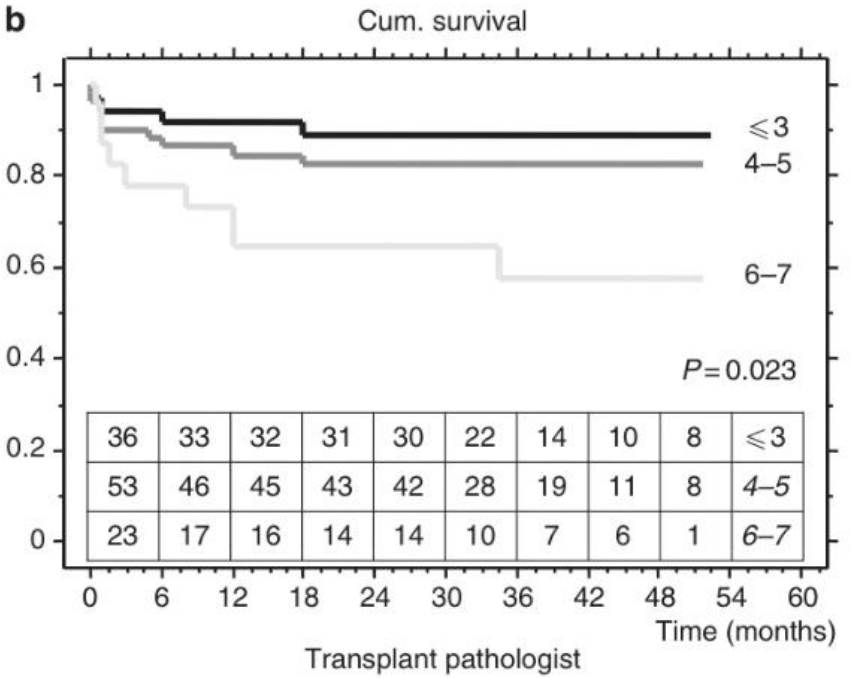
Mark Haas[1]



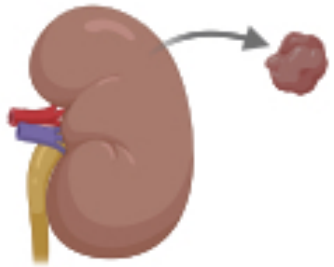
kidney INTERNATIONAL ISN

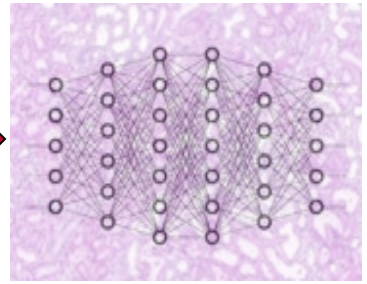
The reproducibility and predictive value on outcome of renal biopsies from expanded criteria donors

M. Antonieta Azancot[1], Francesc Moreso[1], Maite Salcedo[2], Carme Cantarell[1], Manel Perello[1], Irina B. Torres[1], Angeles Montero[2], Enric Trilla[3], Joana Sellarés[1], Joan Morote[3] and Daniel Seron[1]

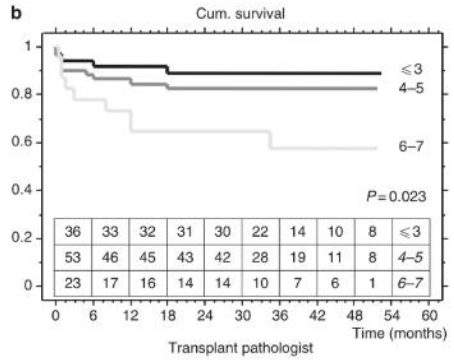


Is Artificial Intelligence Capable of Improving On-Call Pathologist Performance?











S. Joshua Swamidass, MD, PhD

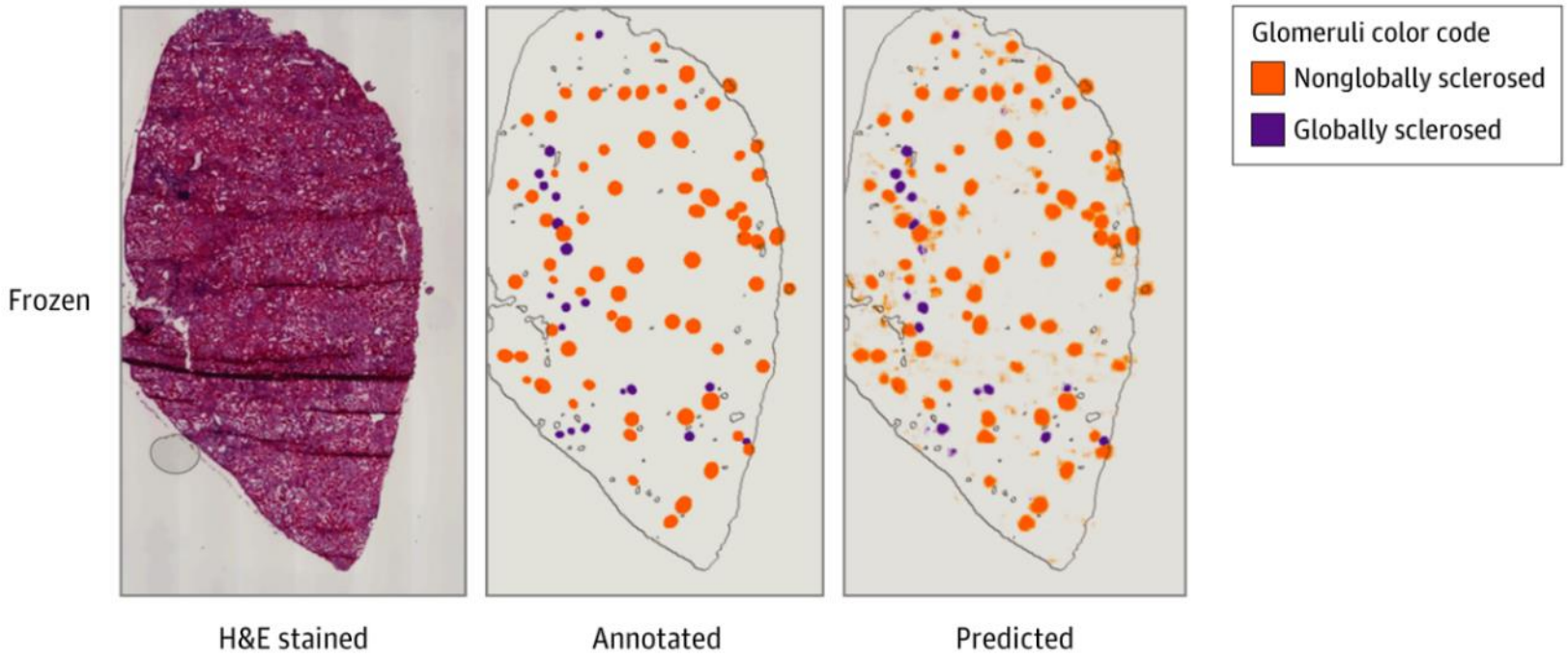


Jon Marsh, PhD



Victoria Swamidass

AI Quantitation of Glomeruli



JAMA Network | Open

Original Investigation | Nephrology

Development and Validation of a Deep Learning Model to Quantify Glomerulosclerosis in Kidney Biopsy Specimen

Jon N. Marsh, PhD; Ta-Chiang Liu, MD, PhD; Parker C. Wilson, MD, PhD; S. Joshua Swamidass, MD, PhD; Joseph P. Gaut, M

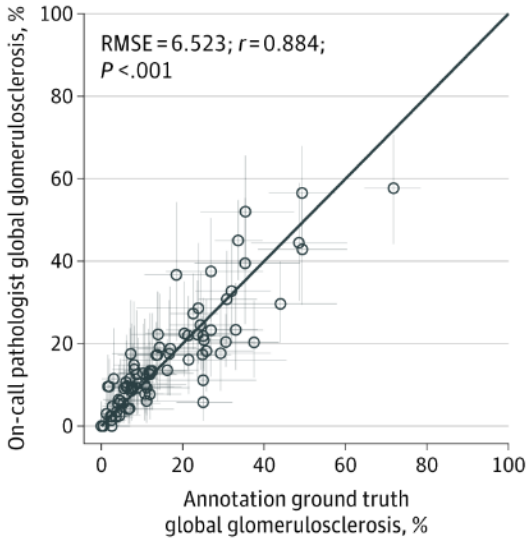
Deep Learning Global Glomerulosclerosis in Transplant Kidney Frozen Sections

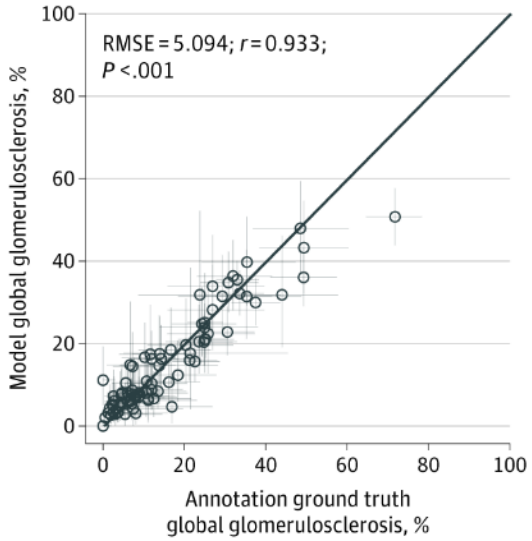
Jon N. Marsh, *Member, IEEE*, Matthew K. Matlock, Satoru Kudose, Ta-Chiang Liu, Thaddeus S. Stappenbeck, Joseph P. Gaut, and S. Joshua Swamidass

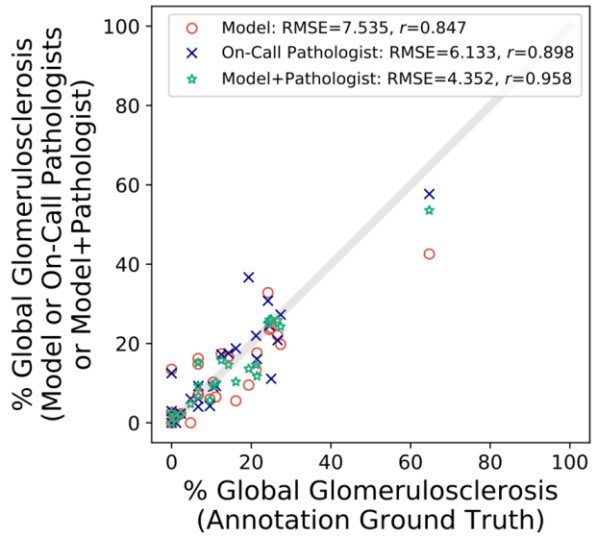
Model Performance

r = **0.884** < **0.933** < **0.958**

On Call Pathologists | **Model** | **Model + Pathologists**



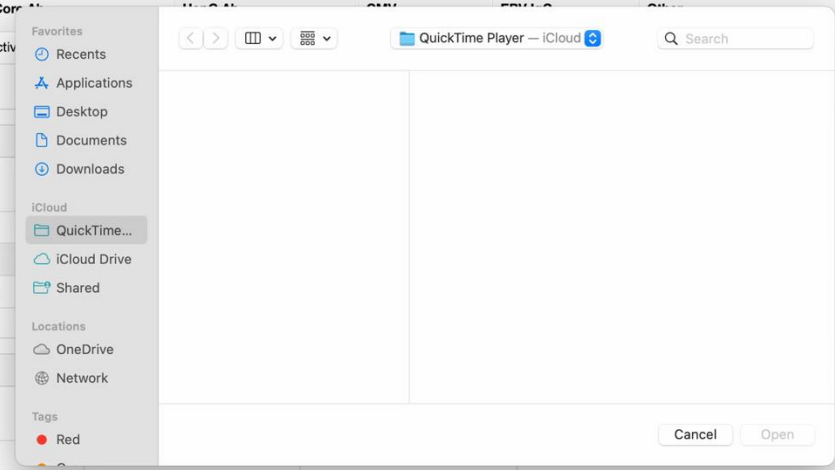




Alcohol? Yes 15 years **Smoking?** No

Obesity? Yes **Comments:**

Serology		
	HepB Surface Ag	**HepB Corr Ab**
Results	not reactive	not reactiv



Kidney Specimen Details

Cysts or Tumors? No

	Peak	**Final**	
CR	2.7	2.7	**BUN**

Gross Description

Number of containers? 2

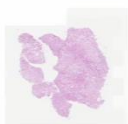
	Identifier (UNOS):	**Part/Sample Type:**			
Specimen :	AMAB338	Kidney, Wedge	1	fresh	1.2 x 0.6 x 0.4 cm
Comments:	left kidney (inked blue)				
Specimen :	AMAB338	Kidney, Wedge	1	fresh	1.1 x 0.6 x 0.3 cm
Comments:	right kidney				

Hot Seat Evaluation

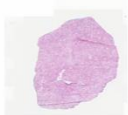
QA Review:

Please enter your findings:

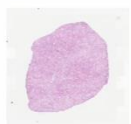
Slides



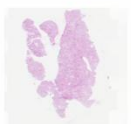
MTS25-10 Right Kid L2.svs
Right L2
Kidney Right - Wedge



MTS25-10 Left Kid L1.svs
Left L1
Kidney Left - Wedge



MTS25-10 Left Kid L2.svs
Left L2
Kidney Left - Wedge



MTS25-10 Right Kid L1.svs
Right L1
Kidney Right - Wedge

Comments:

Open Trusted Kidney

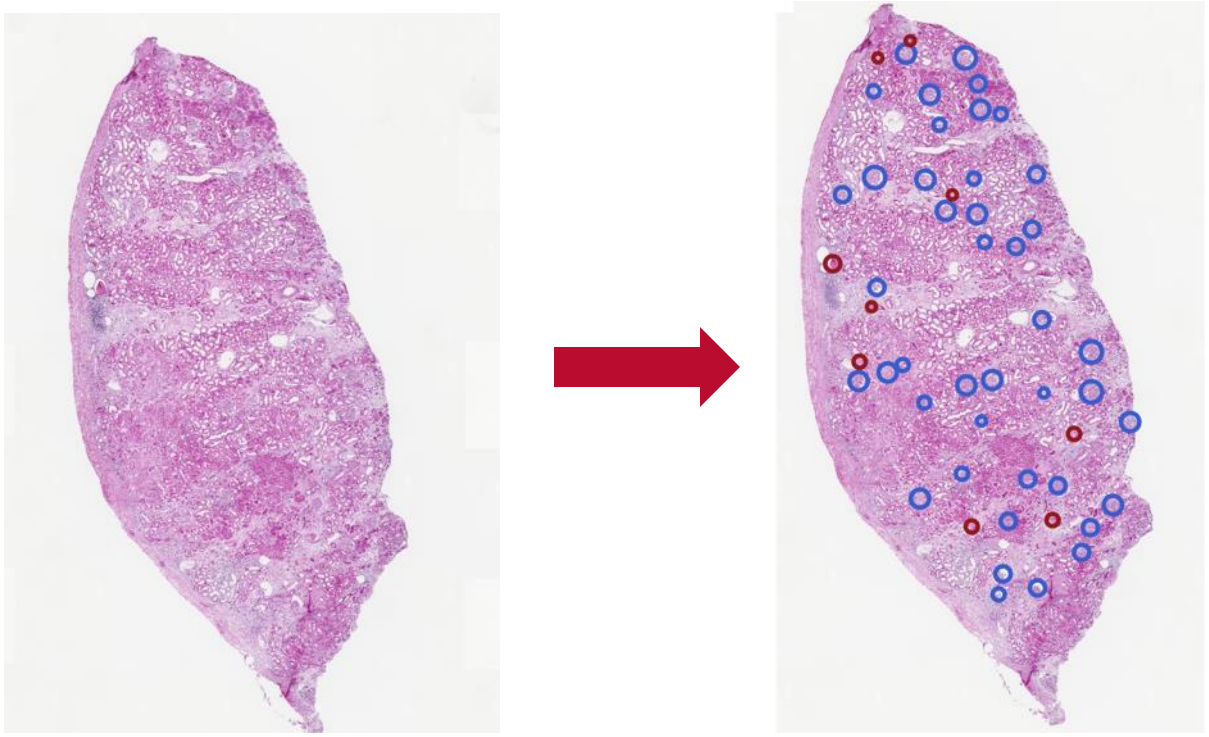
Practice Transformation

Manual

Tedious

Error-prone

Time-consuming

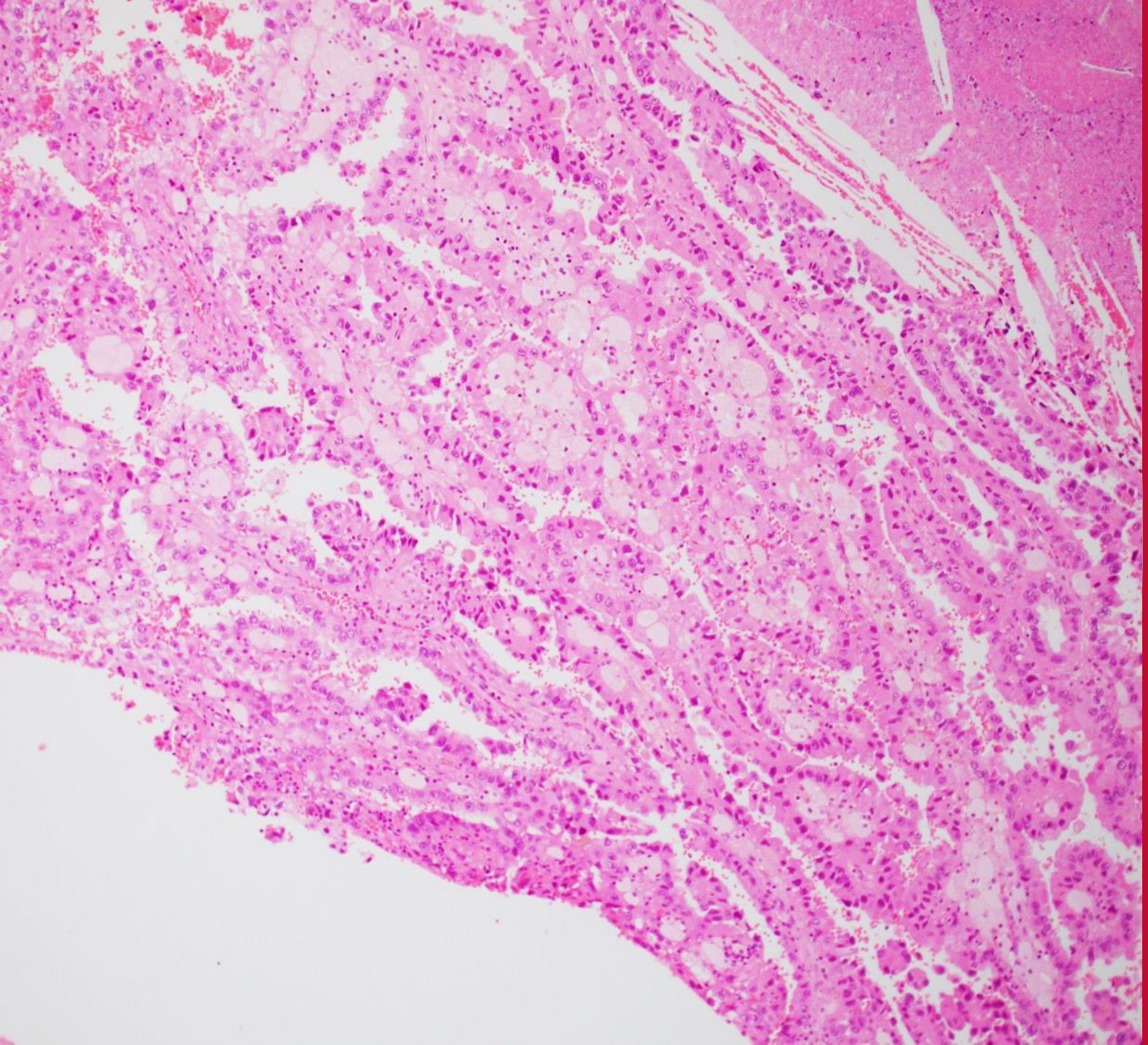


Automated

Liberating

Accurate and precise

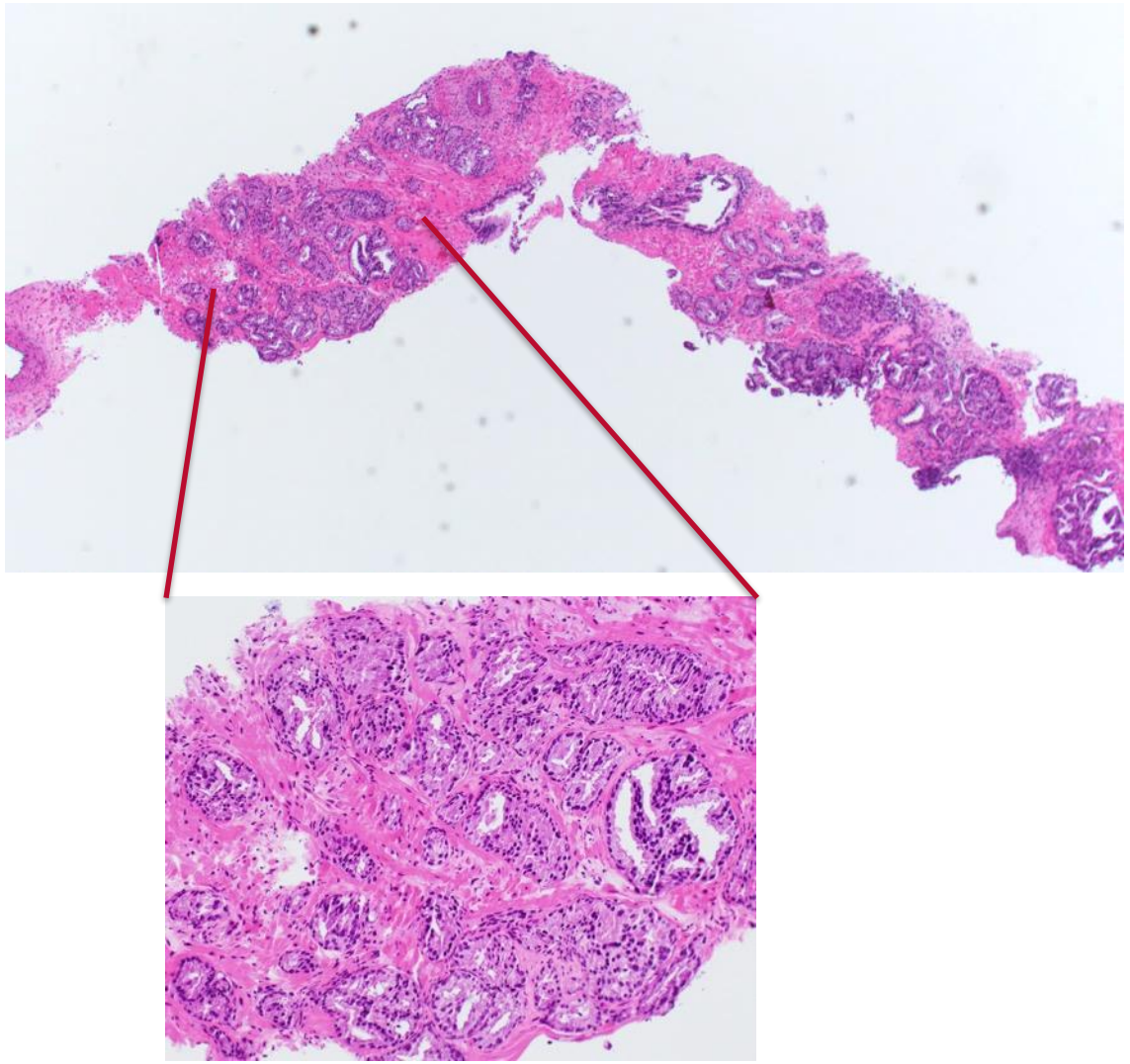
Rapid



Cancer Diagnosis, Typing, Grading

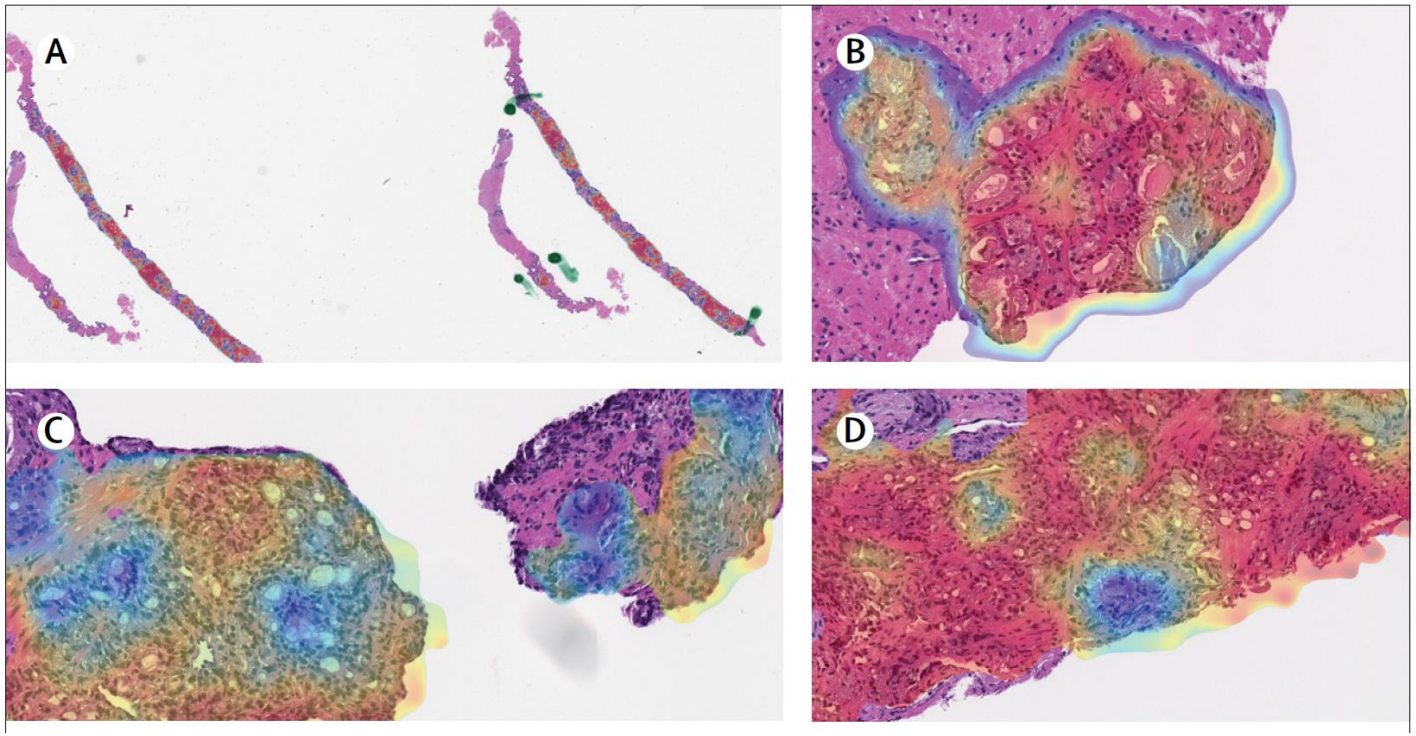
Prostate Cancer

- Biopsies performed in patients with elevated PSA
- Time intensive saturation biopsy procedures
  - 40-50+ slides per patient
- Pathologist identifies, grades, and quantifies cancer on every slide
- Modest reproducibility – less so with higher grade cancer



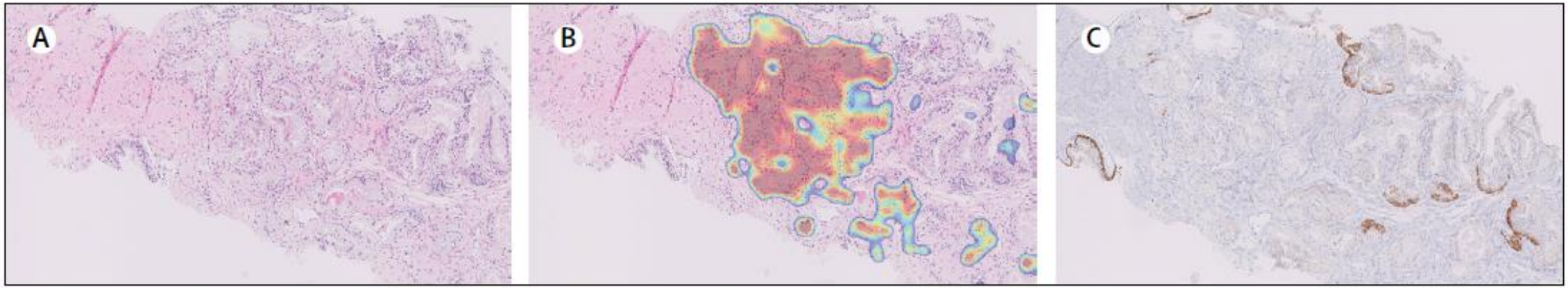
An artificial intelligence algorithm for prostate cancer diagnosis in whole slide images of core needle biopsies: a blinded clinical validation and deployment study

Liron Pantanowitz, Gabriela M Quiroga-Garza, Lilach Bien, Ronen Heled, Daphna Laifenfeld, Chaim Linhart, Judith Sandbank, Anat Albrecht Shach, Varda Shalev, Manuela Vecsler, Pamela Michelow, Scott Hazelhurst, Rajiv Dhir



- Trained on 549 slides, 1,357,480 image patches
- Validated on 1627 external slides (100 cases)

Prostate Cancer Detection



- Above example was missed by the pathologist
- Overall AUC = 0.991 for cancer detection on external validation set
- Now FDA approved for clinical use

Breast Cancer Detection

Human Pathology 162 (2025) 105819

Contents lists available at ScienceDirect

Human Pathology

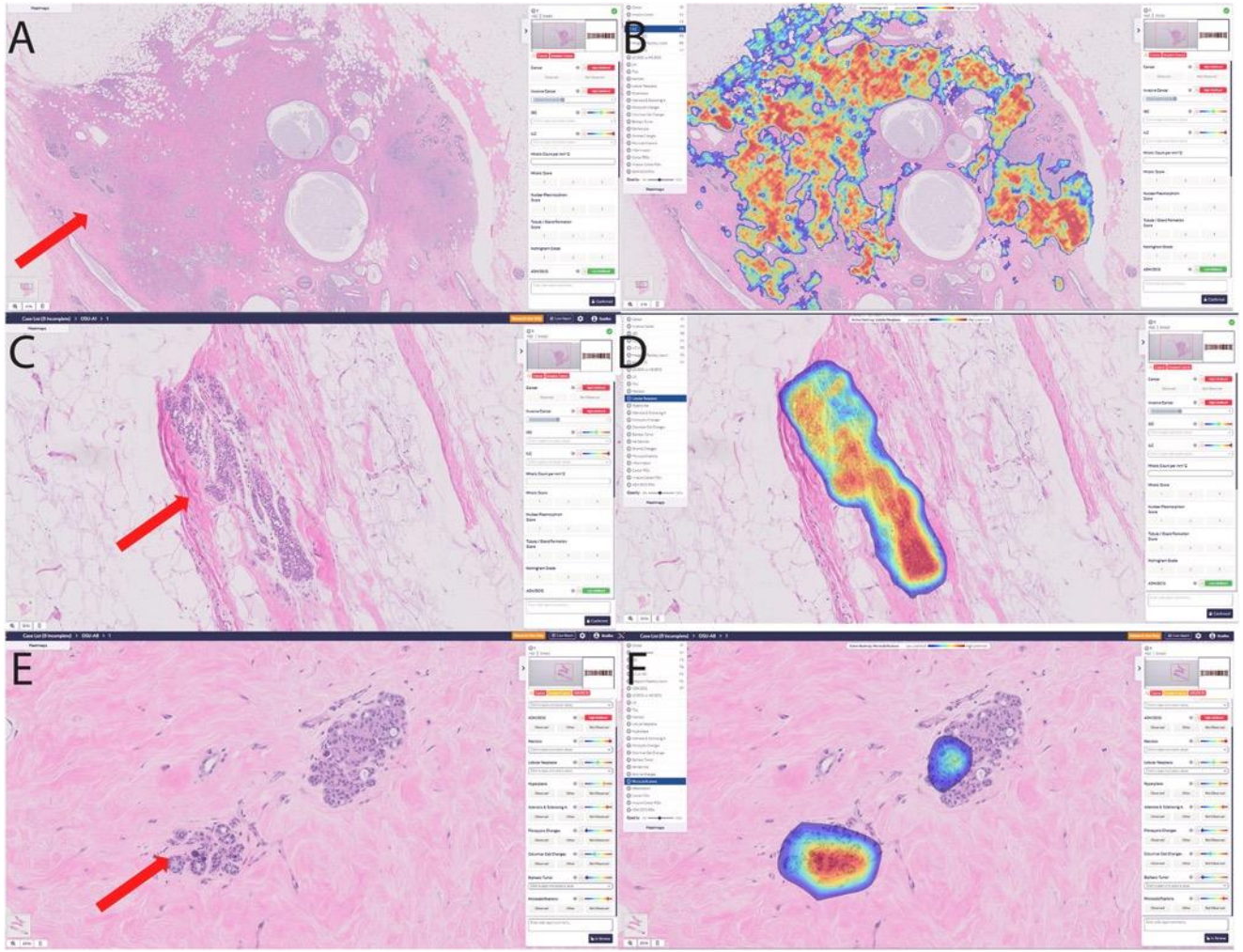
journal homepage: www.elsevier.com/locate/humpath

Artificial intelligence in breast pathology: Overview and recent updates

Sneha Datwani, Hikmat Khan, Muhammad Khalid Khan Niazi, Anil V. Parwani, Zaibo Li *

Department of Pathology, The Ohio State University Wexner Medical Center, Columbus, OH, 43210, USA

- Assist pathologists with diagnosis and classification of breast lesions
- AUC 0.976 for detecting breast cancer
- Pathologist accuracy improved from 97.1% to 100% with AI support.
- Reduced review time by 16.5%
- Decreased IHC usage by 33%

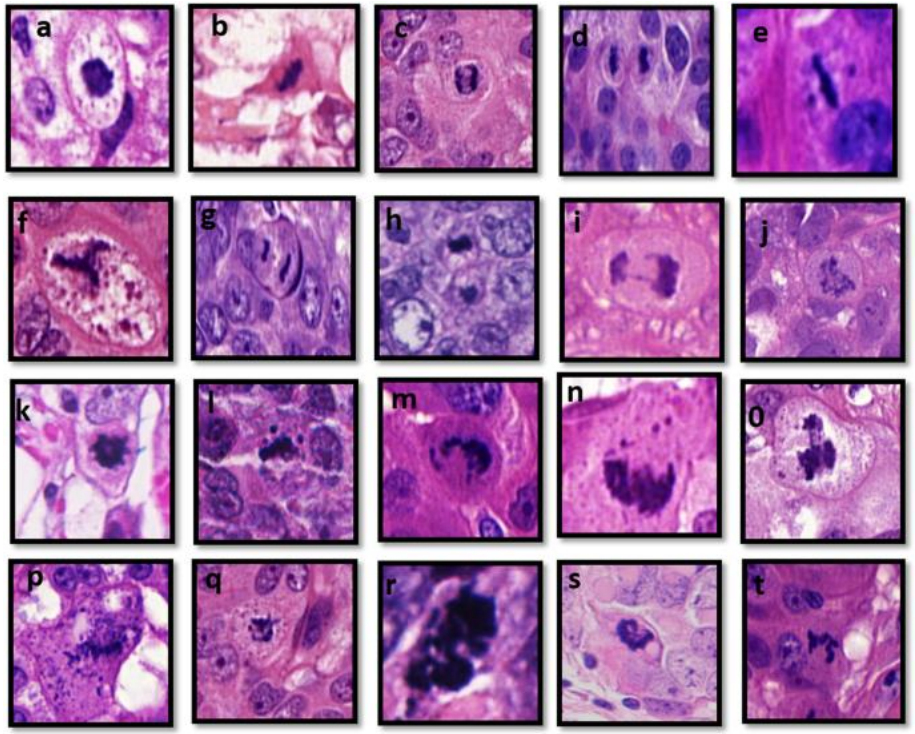


Breast Cancer Grading

Mitotic activity

- Important for pathologic grading of many types of cancer
- Difficult and tedious
- Reproducibility is modest at best
- Results help drive therapy and provide prognostic information

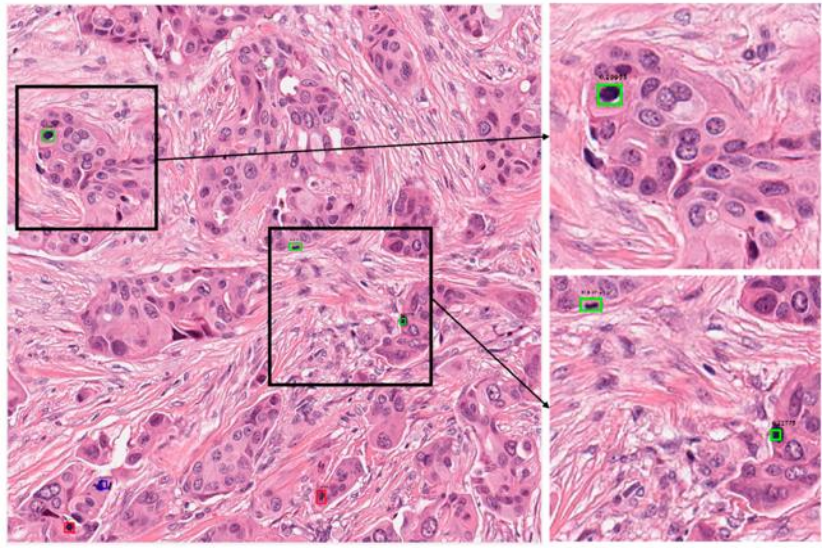
Examples of mitoses in breast cancer



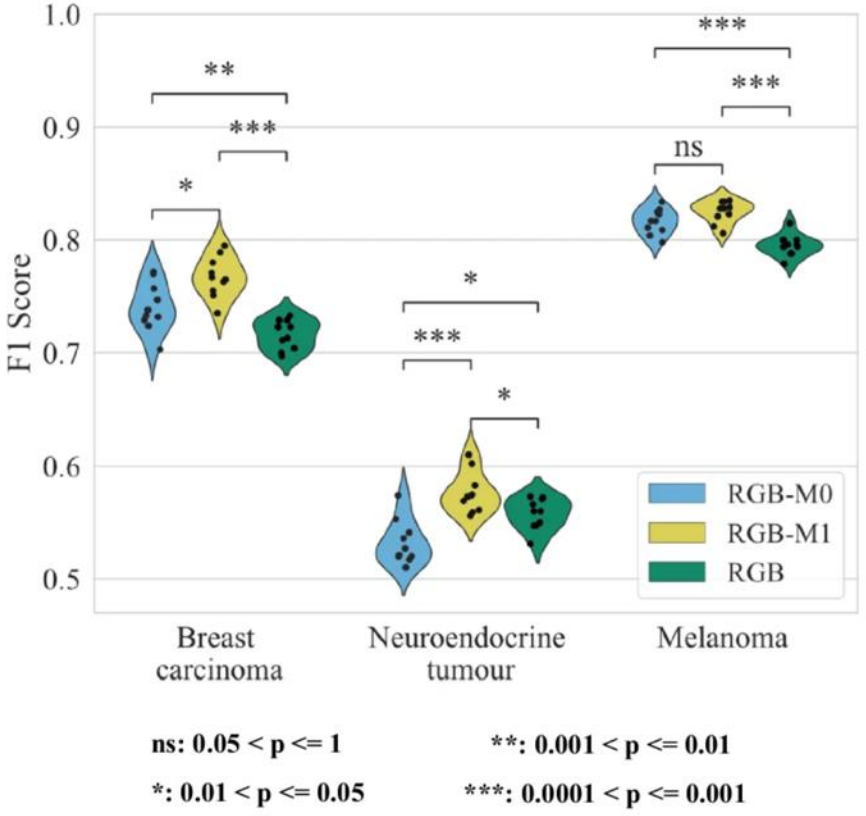
Lashen A, et al. Mod Pathol. 2022:35:1341-1348.

Automated Mitotic Counting

- Mitotic counting across tumor types
- F1 score = 0.84 for pan-cancer mitoses
- Improved reproducibility – kappa 0.457 to 0.726 with AI



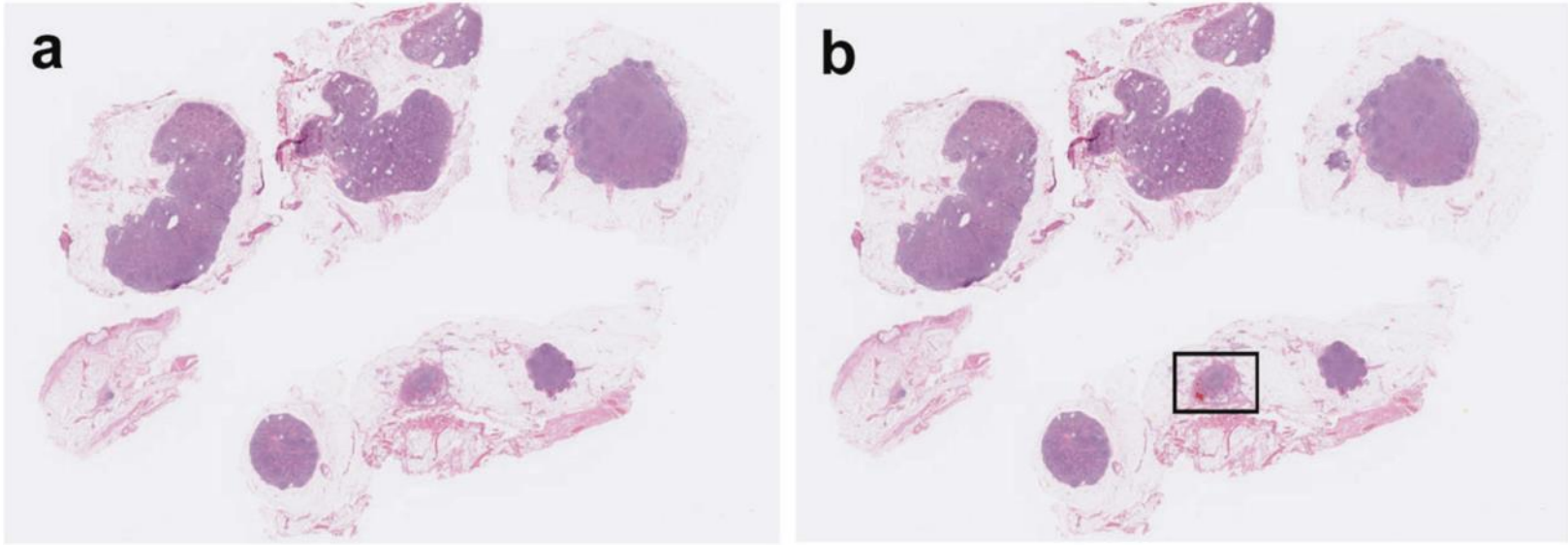
Mahmoud T, et al. J Clin Med. 2020:9(3):749.
Simmat C, et al. Diagnostics (Basel). 2025:15(9):1127.



Shen Z, et al. Comm Biol. 2024:7:1674.

Lymph Node Screening

- Essential component of staging many types of cancer
- Manual review of lymph node slides to find small foci of cancer
- Cancer metastatic to lymph nodes alters therapy



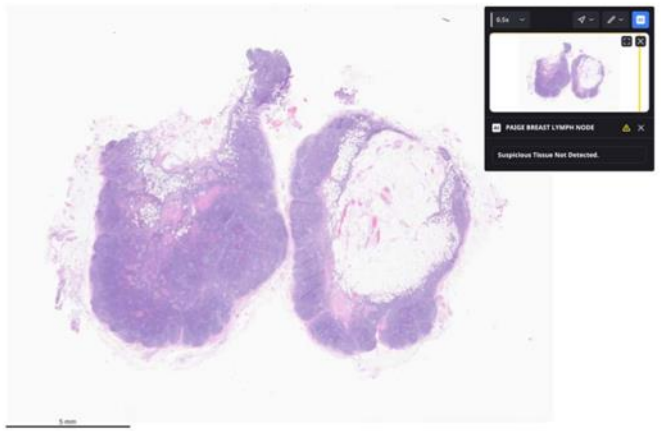
Chuang, W-Y, et al. Mod Pathol. 2021:34:1901-1911.

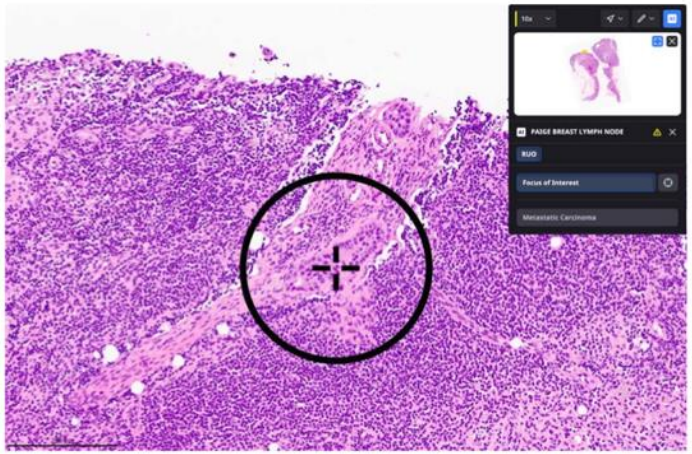
ORIGINAL ARTICLE

OPEN

Artificial Intelligence Helps Pathologists Increase Diagnostic Accuracy and Efficiency in the Detection of Breast Cancer Lymph Node Metastases

Juan Antonio Retamero, MD, MSc, Emre Gulturk, MSc,* Alican Bozkurt, MSc,* Sandy Liu, MD,† Maria Gorgan, MD,† Luis Moral, MD,† Margaret Horton, PhD,* Andrea Parke, PhD,* Kasper Malfroid, MSc,* Jill Sue, MS,* Brandon Rothrock, PhD,* Gerard Oakley, MD,* George DeMuth, MS,‡ Ewan Millar, BSc, FRCPath,*§ Thomas J. Fuchs, DSc,*‖¶ and David S. Klimstra, MD**





- Trained using >32,000 breast sentinel lymph node whole slide images
- >8000 patients
- Pathologists examined a subset of lymph nodes with and without AI assistance
- Speed improved by 55%
- Sensitivity improved for 2 of the 3 pathologists by 20%

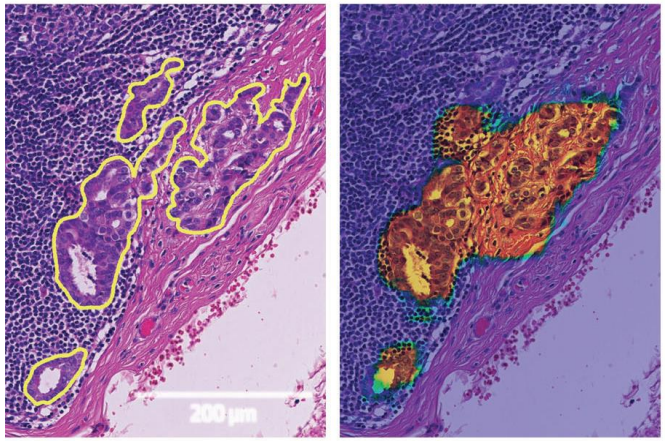
Screening Lymph Nodes

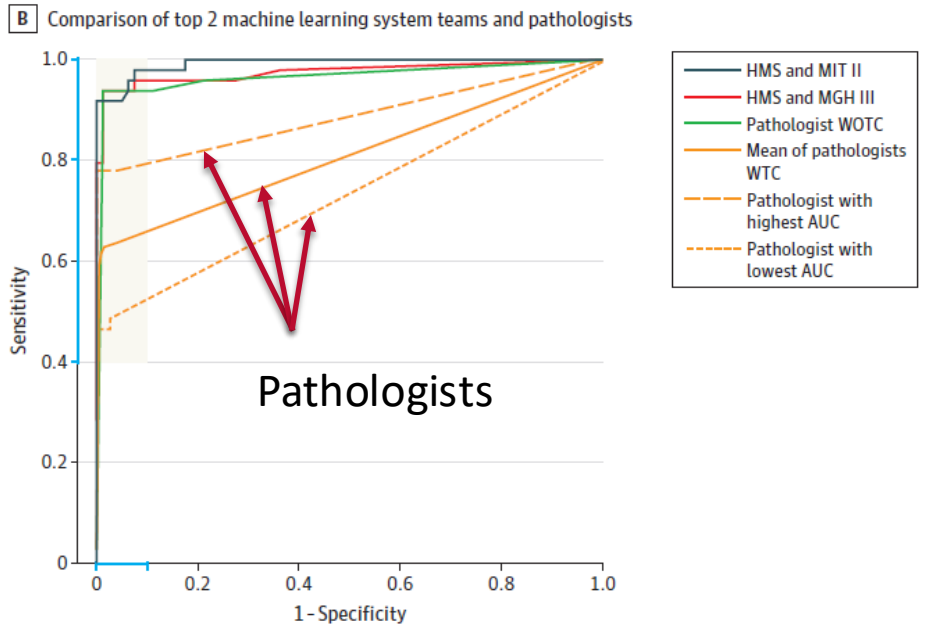
JAMA | Original Investigation

Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer

Babak Ehteshami Bejnordi, MS; Mitko Veta, PhD; Paul Johannes van Diest, MD, PhD; Bram van Ginneken, PhD; Nico Karssemeijer, PhD; Geert Litjens, PhD; Jeroen A. W. M. van der Laak, PhD; and the CAMELYON16 Consortium

- Challenge competition (CAMELYON16)
- WSI with (n=110) and without (n=160) lymph node metastasis
- Validated in independent test set n=129
- Best algorithm AUROC = 0.994
- Pathologists AUROC = 0.966







Predicting Outcomes

Cancer Recurrence

- Trained on 1,039 patients
- Model 1 – recognize tumor
- Model 2 – predict 21 gene recurrence score from histopathology
- Combined model developed
- AUROC = 0.83 for predicting recurrence

npj | breast cancer

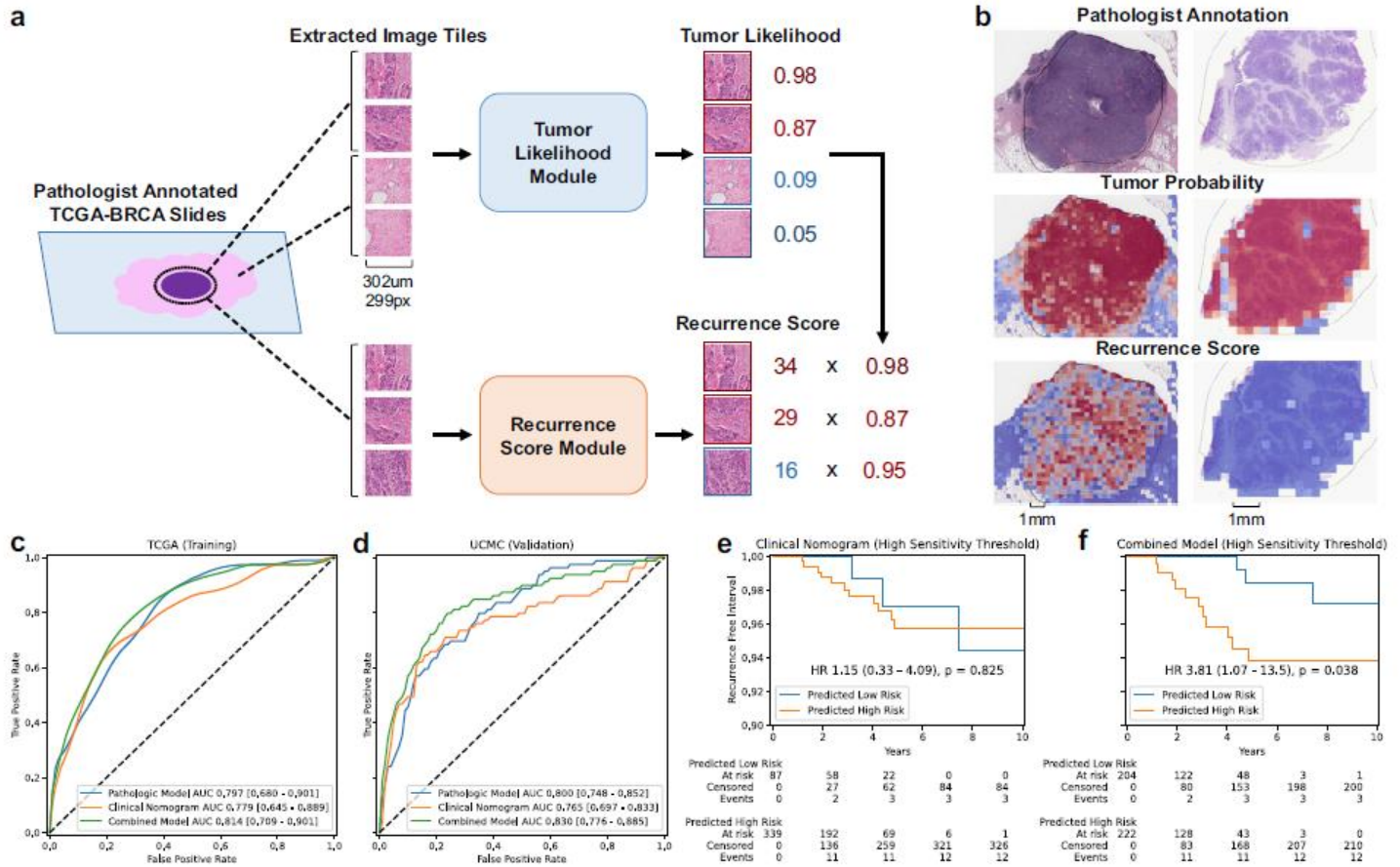
www.nature.com/npjbcancer

BRIEF COMMUNICATION OPEN

Check for updates

Integration of clinical features and deep learning on pathology for the prediction of breast cancer recurrence assays and risk of recurrence

Frederick M. Howard[1], James Dolezal[1], Sara Kochanny[1], Galina Khramtsova[1], Jasmine Vickery[2], Andrew Srisuwananukorn[3], Anna Woodard[1,4], Nan Chen[1], Rita Nanda[1], Charles M. Perou[5], Olufunmilayo I. Olopade[1], Dezheng Huo[6] and Alexander T. Pearson[1]✉





This Photo by Unknown Author is licensed under CC BY-SA

Foundation Models

Foundation Models

nature medicine

Article https://doi.org/10.1038/s41591-024-02857-3

Towards a general-purpose foundation model for computational pathology

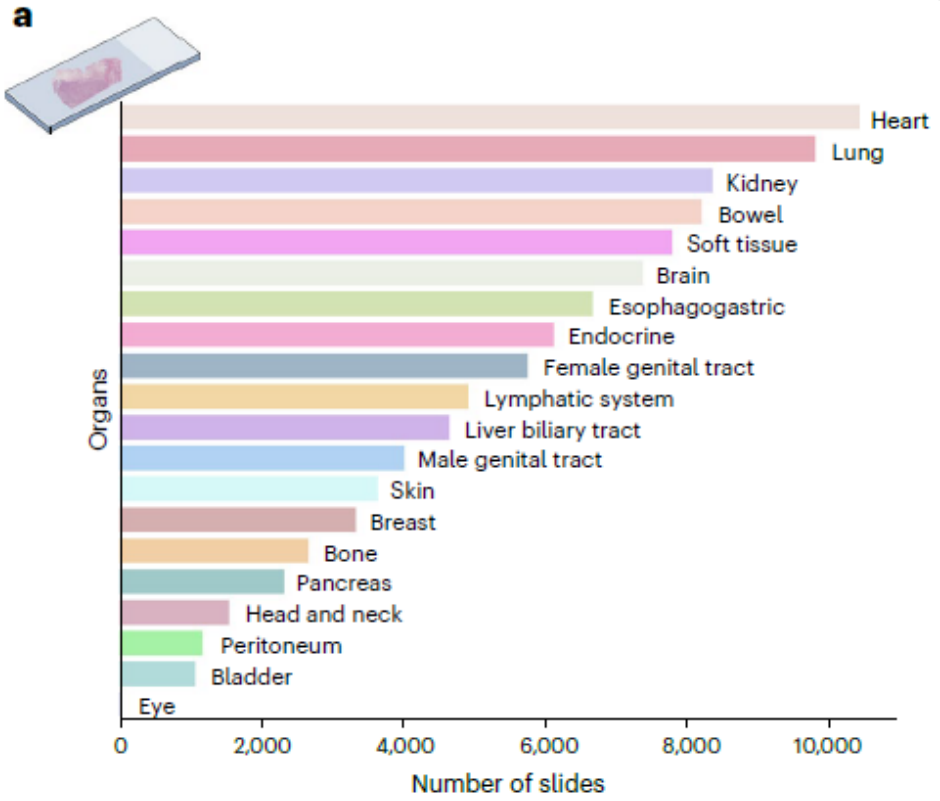
Received: 28 August 2023

Accepted: 5 February 2024

Published online: 19 March 2024

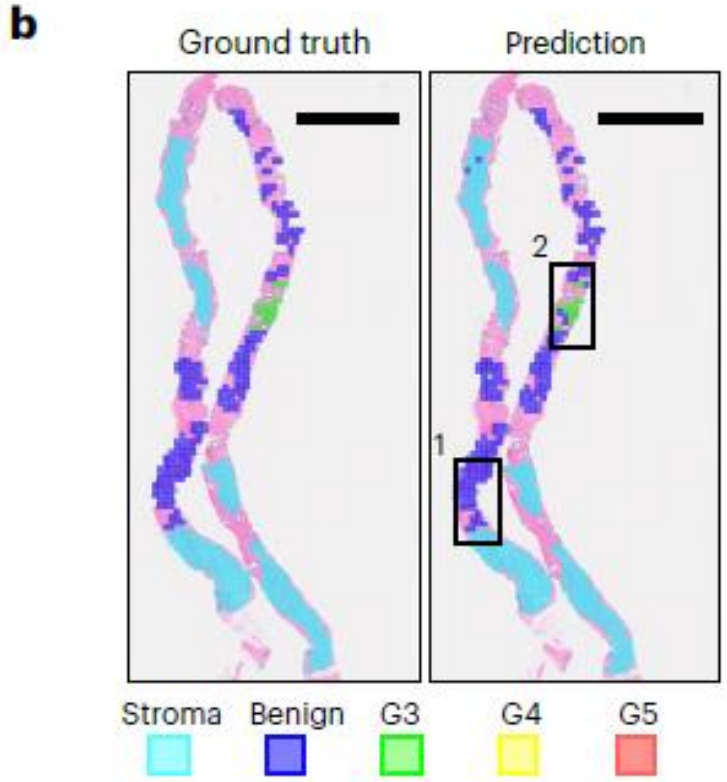
Check for updates

Richard J. Chen[1,2,3,4,5,11], Tong Ding[1,6,11], Ming Y. Lu[1,2,3,4,7,11], Drew F. K. Williamson[1,2,3,11], Guillaume Jaume[1,2,3,4], Andrew H. Song[1,2,3,4], Bowen Chen[1,2], Andrew Zhang[1,2,3,4,8], Daniel Shao[1,2,3,4,8], Muhammad Shaban[1,2,3,4], Mane Williams[1,2,3,4,5], Lukas Oldenburg[1], Luca L. Weishaupt[1,2,3,4,8], Judy J. Wang[1], Anurag Vaidya[1,2,3,4,8], Long Phi Le[2,8], Georg Gerber[1], Sharifa Sahai[1,2,3,4,9], Walt Williams[1,6] & Faisal Mahmood[1,2,3,4,10]



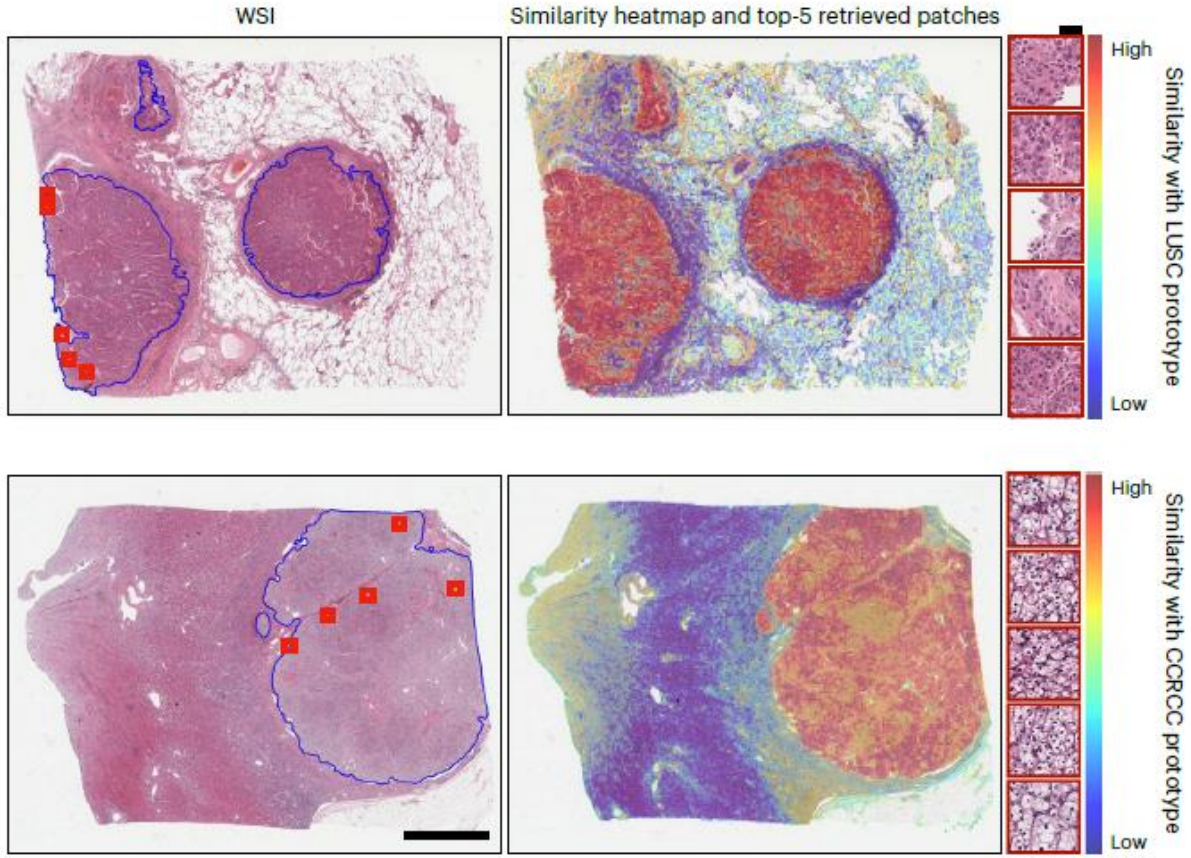
- 100 million images from 100,000 H&E slides
- Slides primarily from 2 institutions
- 20 tissue types
- Self-supervised learning
- Assessed performance on 34 clinical tasks such as classification of cancer types

Selected Results



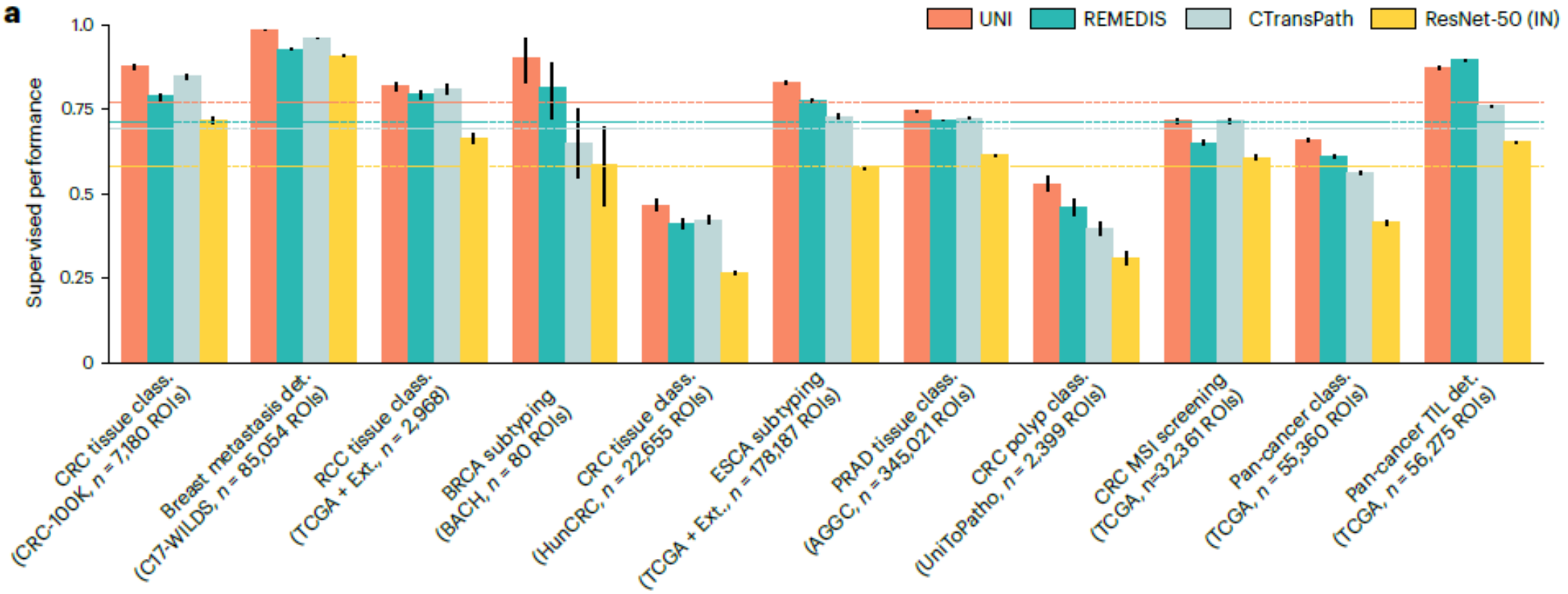
Prostate cancer detection and grading

Lung cancer detection and subtyping



Kidney cancer detection and subtyping

Foundation Model Performance



Foundation Model Comparison

nature communications

Article

https://doi.org/10.1038/s41467-025-58796-1

A clinical benchmark of public self-supervised pathology foundation models

Received: 8 July 2024

Accepted: 2 April 2025

Published online: 17 April 2025

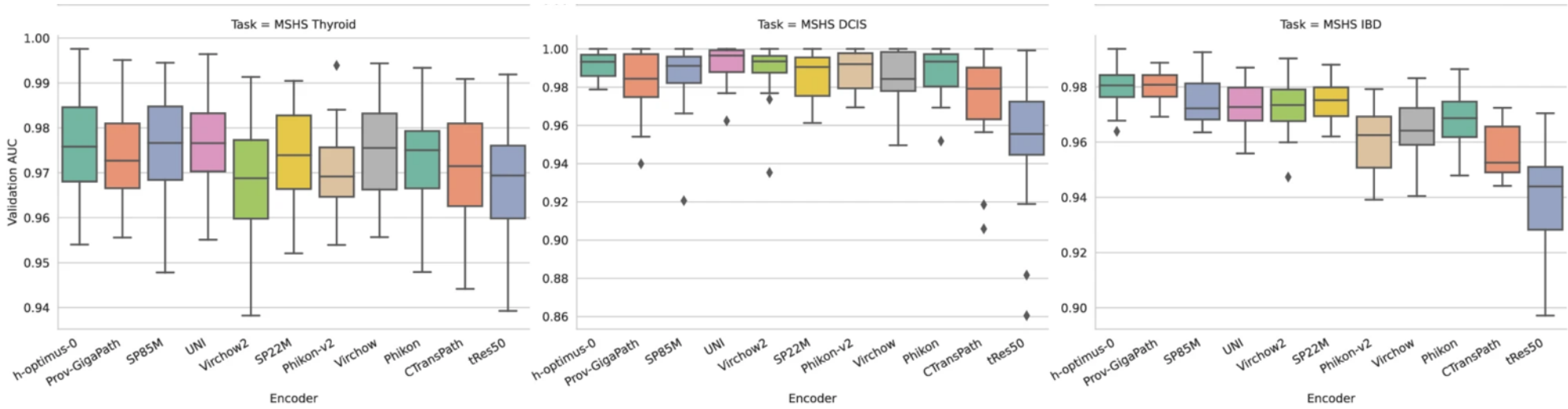
Check for updates

Gabriele Campanella[1,2], Shengjia Chen[1,2], Manbir Singh[1,2], Ruchika Verma[1,2], Silke Muehlstedt[1,2], Jennifer Zeng[3], Aryeh Stock[3], Matt Croken[3], Brandon Veremis[3], Abdulkadir Elmas[4], Ivan Shujski[5,6], Noora Neittaanmäki[5,6], Kuan-lin Huang[4], Ricky Kwan[3], Jane Houldsworth[3], Adam J. Schoenfeld[7] & Chad Vanderbilt[8]

Model	Param. (M)	Algorithm	Training Data Source	Tiles (M)	Slides (K)	Training Resolution
CTransPath[18]	28	SRCL	TCGA, PAIP	16	32	20x
Phikon[21]	86	iBOT	TCGA	43	6	20x
UNI[23]	303	DINOv2	MGB	100	100	20x
Virchow[25]	631	DINOv2	MSKCC	2000	1488	20x
Ref. 26	22	DINO	MSHS	1600	423	20x
Ref. 26	303	MAE	MSHS	3200	423	20x
Rudolf-V[29]	304	DINOv2	Multicenter	1200	134	20x
Prov-GigaPath[30]	1135	DINOv2	PHS	1300	171	20x
Virchow2[33]	631	DINOv2	MSKCC	1700	3100	5,10,20,40 ×
H-optimus-0[34]	1135	DINOv2	Proprietary	>100	>500	20x
Phikon-v2[35]	307	DINOv2	Multicenter	456	58	20x

MGB Mass General Brigham, *MSKCC* Memorial Sloan Kettering Cancer Center, *MSHS* Mount Sinai Health System, *PHS* Providence Health and Services.

Performance Comparison of Foundation Models - Detection



- Model performance for detection tasks
- Thyroid cancer, breast ductal carcinoma in situ, inflammatory bowel disease, and others
- All models show validation AUC >0.9 (best above 0.96)



Beyond Human Vision

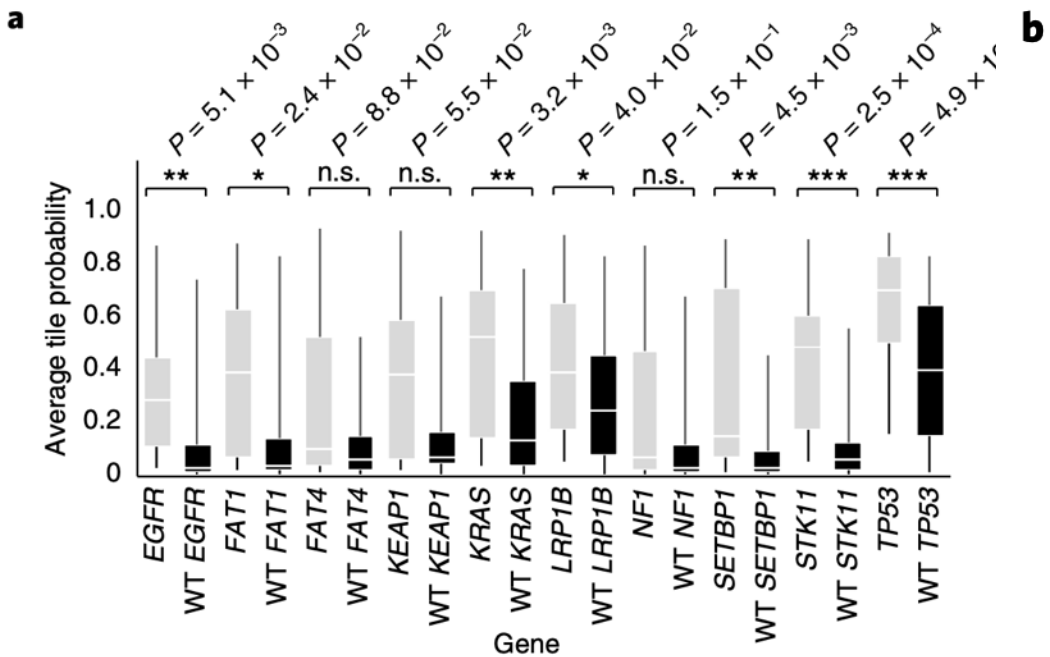
nature medicine

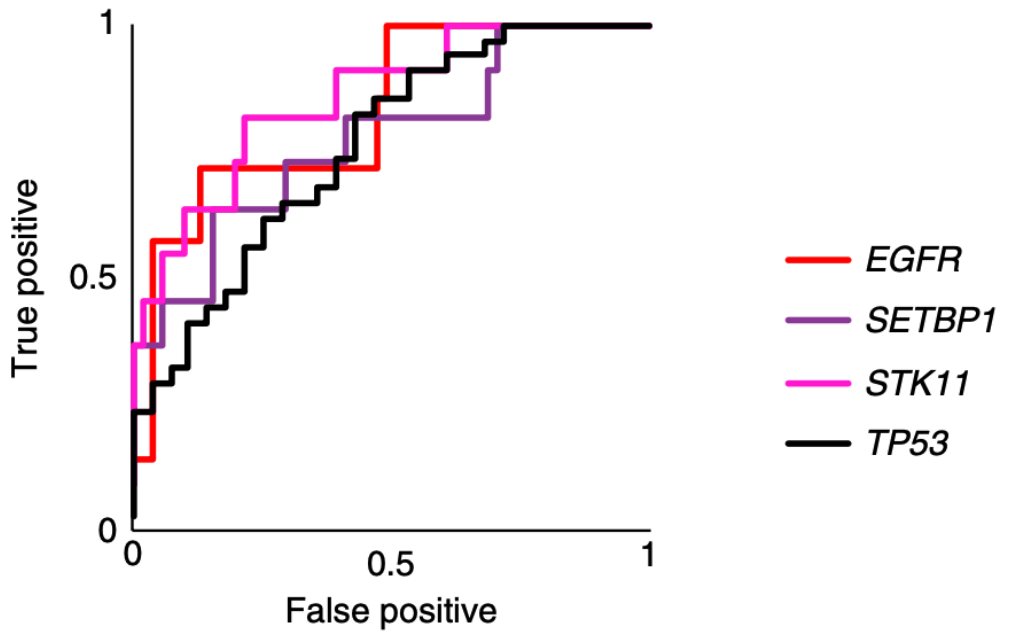
ARTICLES

https://doi.org/10.1038/s41591-018-0177-5

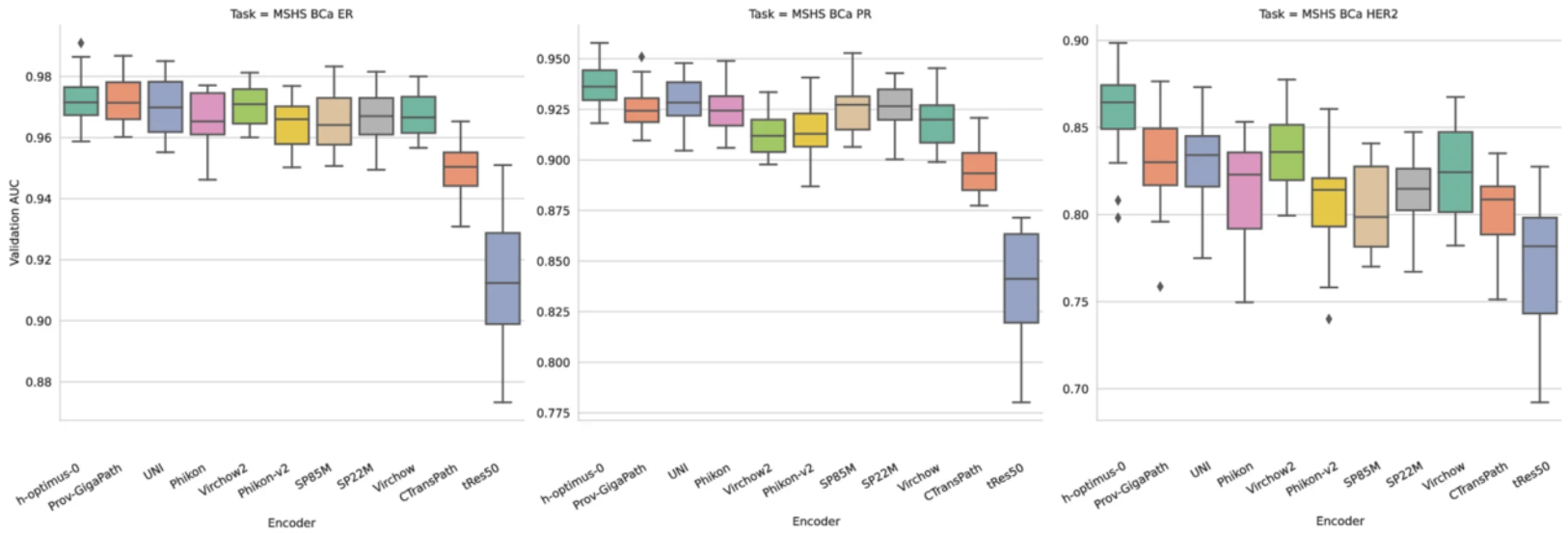
Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning

Nicolas Coudray[1,2,9], Paolo Santiago Ocampo[3,9], Theodore Sakellaropoulos[4], Navneet Narula[3], Matija Snuderl[3], David Fenyö[5,6], Andre L. Moreira[3,7], Narges Razavian[8*] and Aristotelis Tsirigos[1,3*]





Performance – Biomarker Prediction



- Prediction of breast cancer biomarkers
- Estrogen receptor, progesterone receptor, Her2
- Majority of models show validation AUC 0.96-0.98

nature communications

Article

https://doi.org/10.1038/s41467-025-58796-1

A clinical benchmark of public self-supervised pathology foundation models

Received: 8 July 2024

Accepted: 2 April 2025

Published online: 17 April 2025

Check for updates

Gabriele Campanella[1,2], Shengjia Chen[1,2], Manbir Singh[1,2], Ruchika Verma[1,2], Silke Muehlstedt[1,2], Jennifer Zeng[3], Aryeh Stock[3], Matt Croken[3], Brandon Veremis[3], Abdulkadir Elmas[4], Ivan Shujski[5,6], Noora Neittaanmäki[5,6], Kuan-lin Huang[4], Ricky Kwan[3], Jane Houldsworth[3], Adam J. Schoenfeld[7] & Chad Vanderbilt[8]

Future Digital Pathology

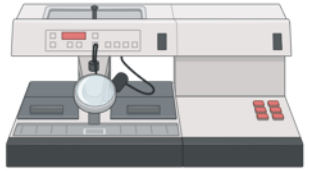
1. Obtain a biopsy sample from the patient.



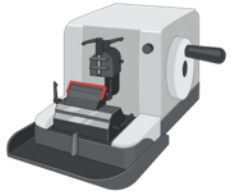
2. Fix the biopsy sample in formalin.



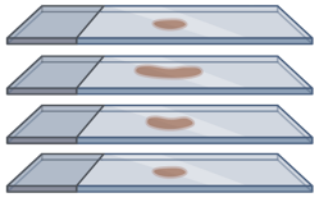
3. Dehydrate and embed the sample in paraffin.



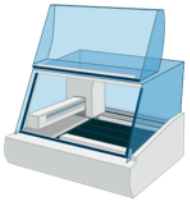
4. Section the sample using a microtome.



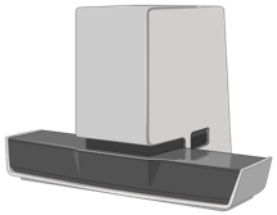
5. Mount the sections on slides.



6. Stain the slides using an auto stainer.



7. Scan the stained slides.



8. Review and sign out the digital slides.



Virtual Slide Staining

1. Obtain a biopsy sample from the patient.



2. Fix the sample in formalin.



3. Embed the sample in paraffin.



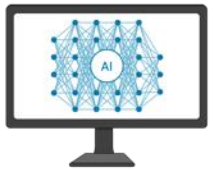
4. Section the sample and create slides.



5. Image the unstained slides.

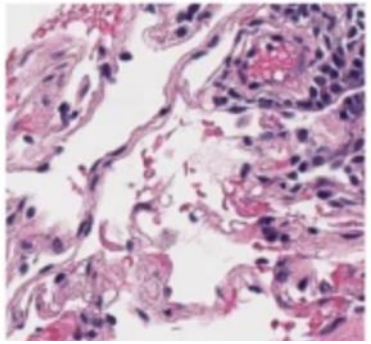


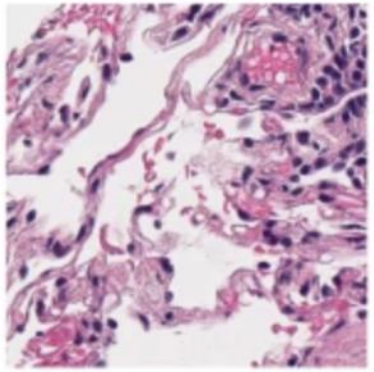
6. Use AI to generate stained slide images.



7. Review digitally and sign out.







nature communications

Article

https://doi.org/10.1038/s41467-025-60387-z

Pixel super-resolved virtual staining of label-free tissue using diffusion models

Received: 21 October 2024

Accepted: 16 May 2025

Yijie Zhang[1,2,3,5], Luzhe Huang[1,2,3,5], Nir Pillar[1,2,3], Yuzhu Li[1,2,3], Hanlong Chen[1,2,3] & Aydogan Ozcan[1,2,3,4]



Implementation Challenges

Model Performance in the Field

scientific reports

Check for updates

OPEN

Validation of histopathology foundation models through whole slide image retrieval

Saghir Alfasly[1], Ghazal Alabtah[1], Sobhan Hemati[1], Krishna Rani Kalari[2], Joaquin J. Garcia[3] & H. R. Tizhoosh[1]✉

Foundation Model Performance

	Top-1	Majority @ Top-3	Majority @ Top-5
Yottixel	28% ± 13% [23% 33%]	28% ± 13% [17% 32%]	27% ± 13% [22% 32%]
Yottixel-UNI	44% ± 15% [38% 50%]	44% ± 14% [38% 50%]	42% ± 14% [36% 48%]
Yottixel-Virchow	41% ± 15% [35% 47%]	41% ± 14% [35% 47%]	40% ± 13% [36% 45%]
Yottixel-GigaPath	43% ± 15% [37% 49%]	43% ± 15% [37% 49%]	41% ± 13% [35% 47%]
GigaPath WSI	42% ± 15% [36% 48%]	41% ± 15% [35% 47%]	40% ± 14% [35% 45%]

The Harmonic Mean of Precision and Recaall

$$F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

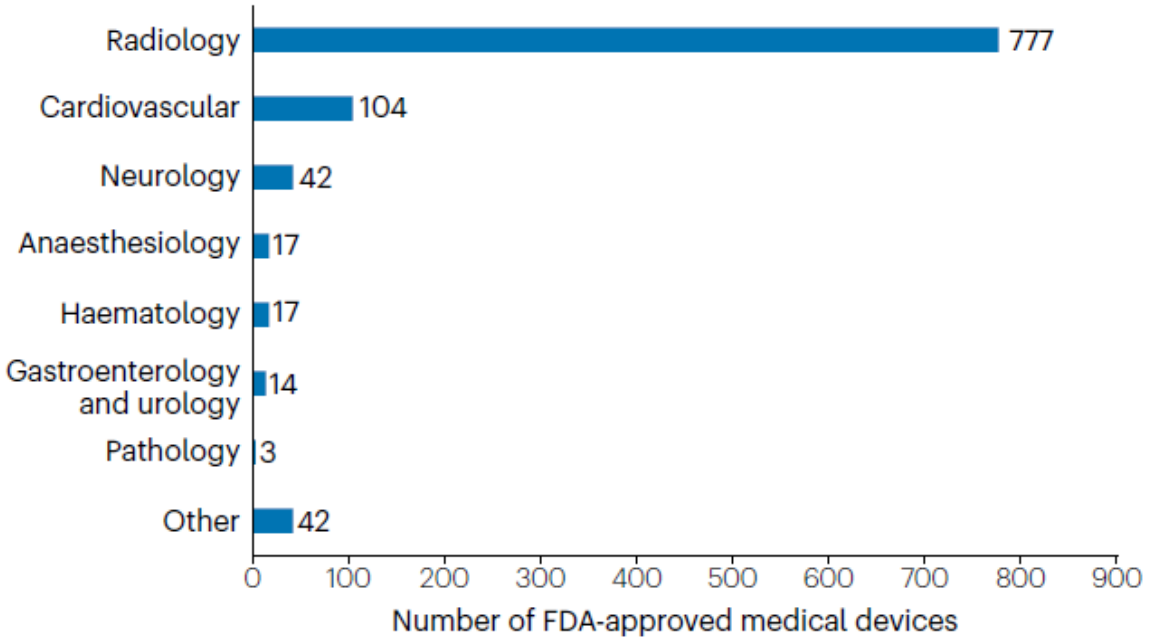
$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

- Evaluated AI driven image retrieval
- 11,444 WSI from 9,339 patients
- 23 organs, 117 cancer types
- Average F1 scores and 95% confidence intervals for whole slide image obtained from The Cancer Genome Atlas – previously not seen by the models under evaluation

Still Early Days for Pathology

- Challenges to clinical use
  - Clinical validation
    - Randomized controlled trials
  - Heterogeneous data in real world
  - Black box nature
  - Reimbursement
  - Workflow integration
  - Digitization required
  - Major hardware investment to run large foundation models



Aggarwal A. et al. Nat Rev Clin Oncology. 2025:22:283-291.

Looking to the Future

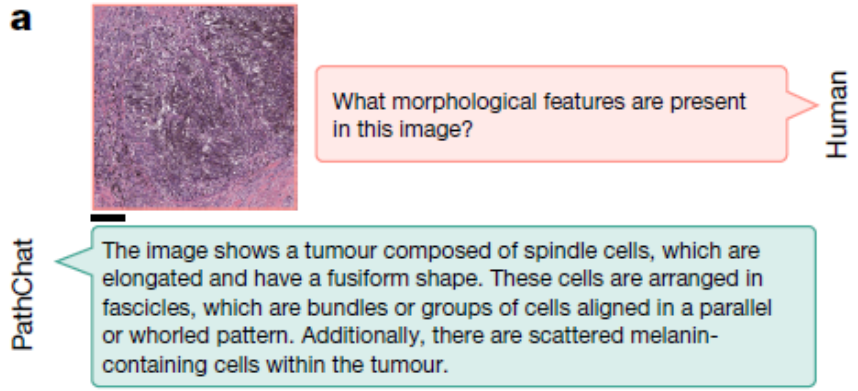
- Additional single task models FDA approved beyond prostate and breast cancer
- Further evolution of foundation models with larger and more representative datasets
- Multimodal models using generative AI technology to assist pathologists
  - PathChat

Article

A multimodal generative AI copilot for human pathology

https://doi.org/10.1038/s41586-024-07618-3
Received: 11 December 2023
Accepted: 28 May 2024

Ming Y. Lu[1,2,3,4,11], Bowen Chen[1,2,11], Drew F. K. Williamson[1,2,3,11], Richard J. Chen[1,2,3], Melissa Zhao[1,2], Aaron K. Chow[5], Kenji Ikemura[1,2], Ahrong Kim[1,6], Dimitra Pouli[1,2], Ankush Patel[7], Amr Soliman[5], Chengkuan Chen[1], Tong Ding[1,8], Judy J. Wang[1], Georg Gerber[1], Ivy Liang[1,9], Long Phi Le[2], Anil V. Parwani[5], Luca L. Weishaupt[1,9] & Faisal Mahmood[1,2,3,10]




WashU Medicine