Navigating the Foundation Model Era in Pathology

Heather Couture

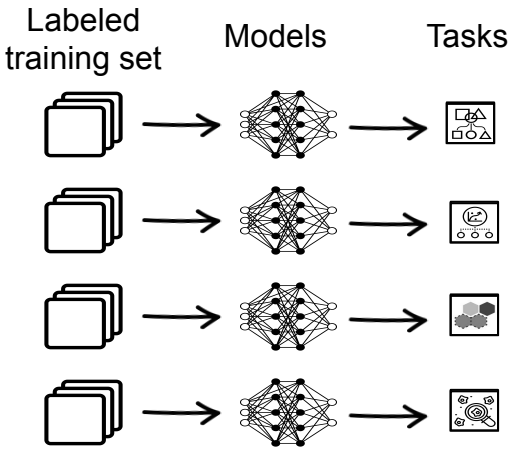
African Congress on Digital Pathology

April 29, 2026

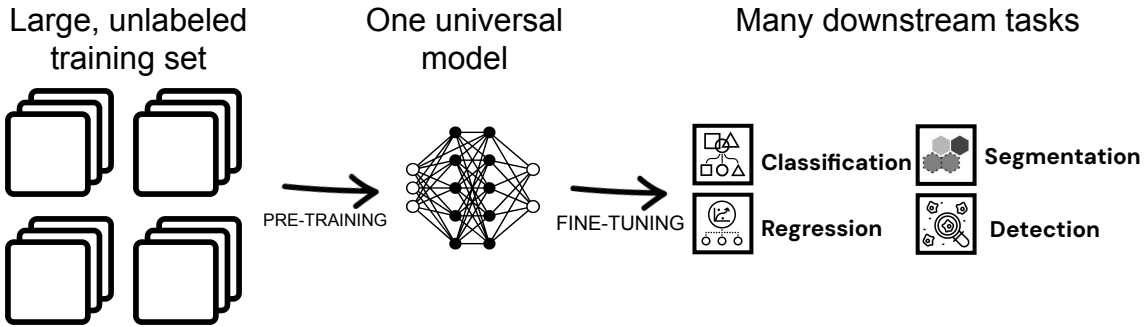
Pixel Scientia LABS

The Paradigm Shift

Traditional ML



Foundation Models



“Foundation model” coined by Stanford Institute for Human-Centered Artificial Intelligence's Center for Research on Foundation Models

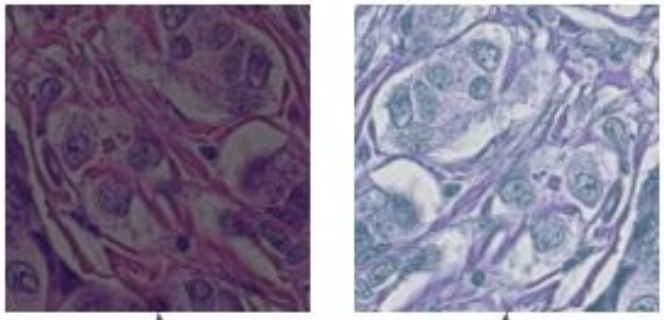
“A foundation model is any model that is trained on **broad data** (generally using **self-supervision** at scale) that can be **adapted** (e.g., fine-tuned) to a wide range of downstream tasks.”

Source: Bommasani, On the Opportunities and Risks of Foundation Models, 2021

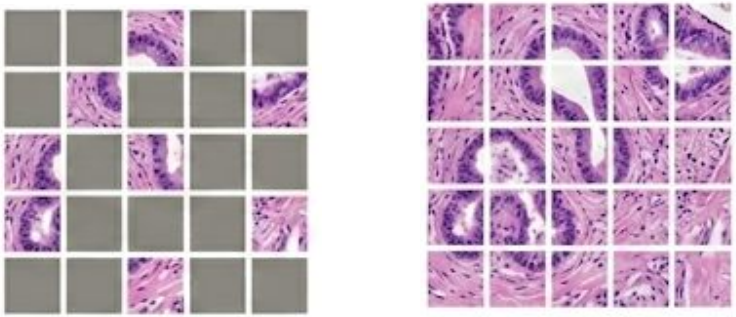
Self-Supervised Pre-Training

By training on millions of unlabeled images, the model solves puzzles to learn the inherent structure of tissue.

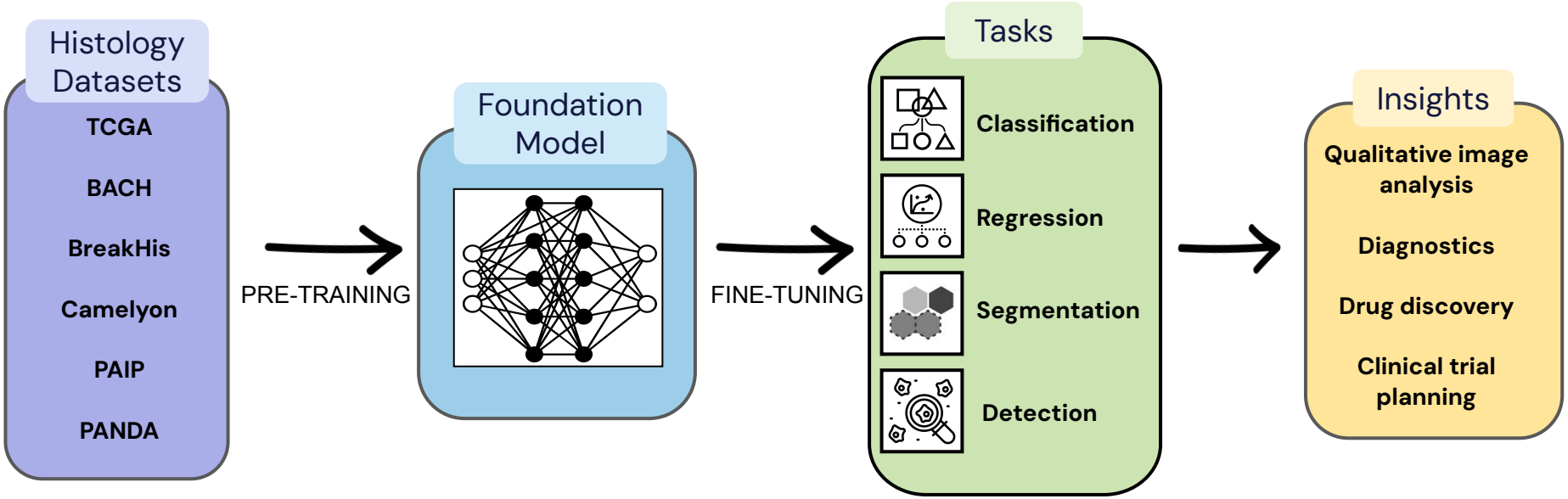
Contrastive:
Same or different?



Reconstruction:
Fill in masked patches

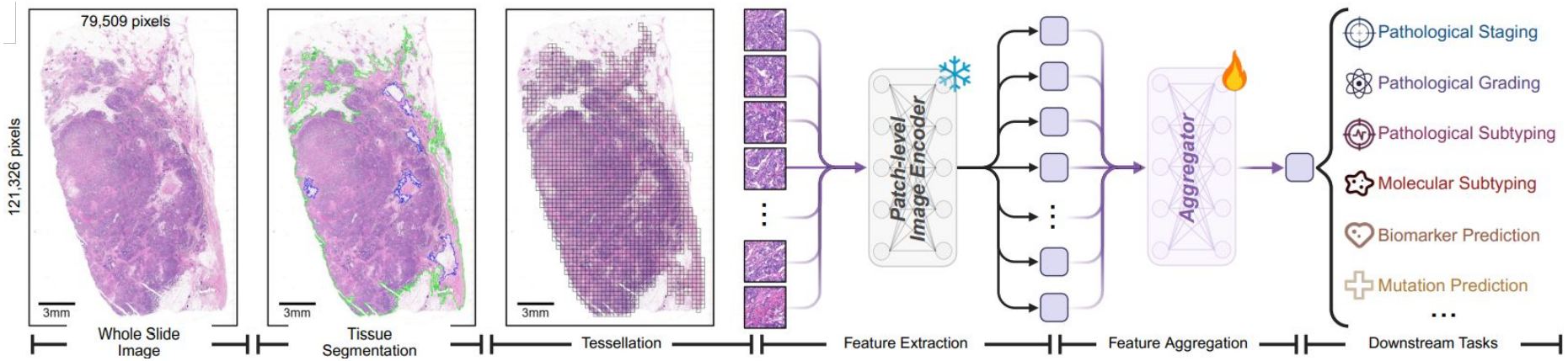


Foundation Models for Histology: Adaptable & Generalizable



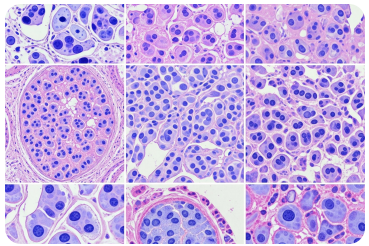
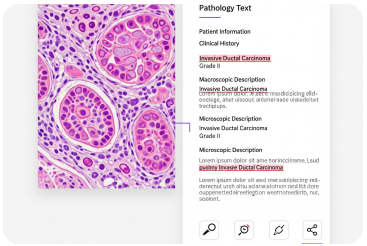
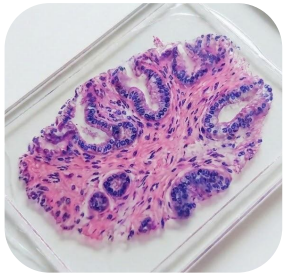
First publicly available model ~2022
“Self-supervised learning”

How to Use a Pathology Foundation Model

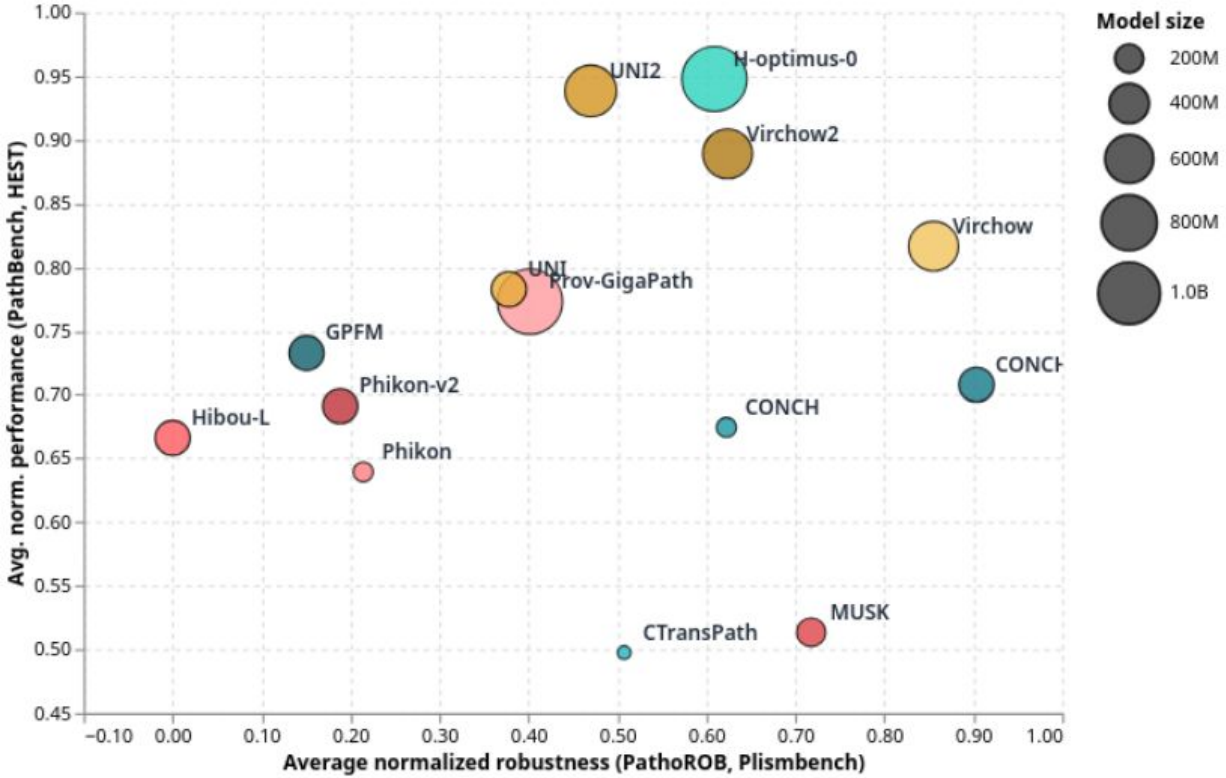


Source: Ma, PathBench: A comprehensive comparison benchmark for pathology foundation models towards precision oncology, 2025

Taxonomy of Pathology Foundation Models

	Single Modality (Images)	Multimodal (Images + Text/Data)
Tile-Level (Patches)	**Tile Features** Visual representations of localized morphology. 	**Tile + Text** Alignment of image patches with text descriptions. 
Slide-Level (WSI Context)	**Slide Reasoning** Global representations capturing spatial context across the entire tissue section. 	**Integrated Insights** Combining WSIs with radiology, genomics, transcriptomics, and/or electronic health records. 

A Diverse Set of Tile-Based Foundation Models

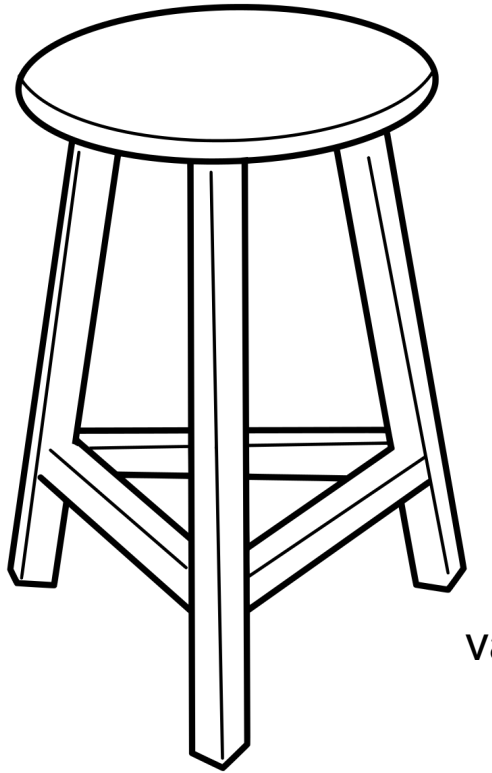


Histoboard currently aggregates results from 10 published benchmarks covering:

- 411 evaluation tasks spanning classification, survival prediction, biomarker detection, and more
- 46 foundation models from academic and industry labs worldwide
- 20 organs including Bladder, Brain, Breast, Cervix, Colorectal, and others
- Robustness evaluation across domain shifts, scanners, and staining variations

Source: https://wearewaiv.github.io/histoboard

From Promise to Practice: Three Pillars

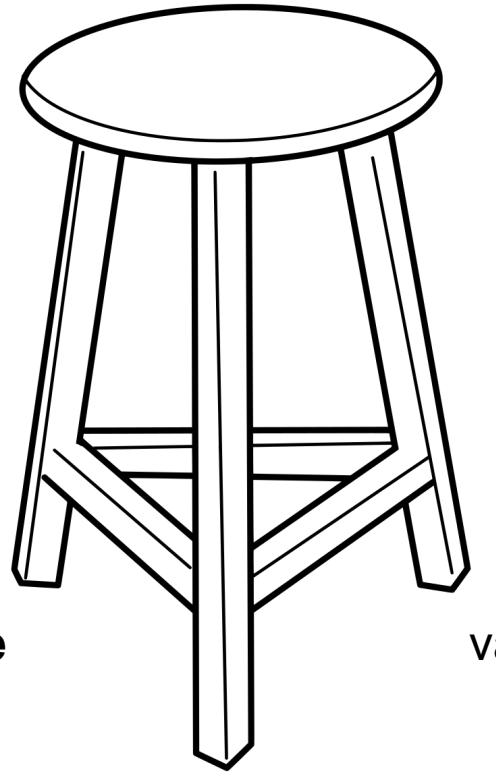


Trustworthy:
robust & reliable

Actionable:
valuable & sustainable

Responsible:
fair & transparent

From Promise to Practice: Three Pillars



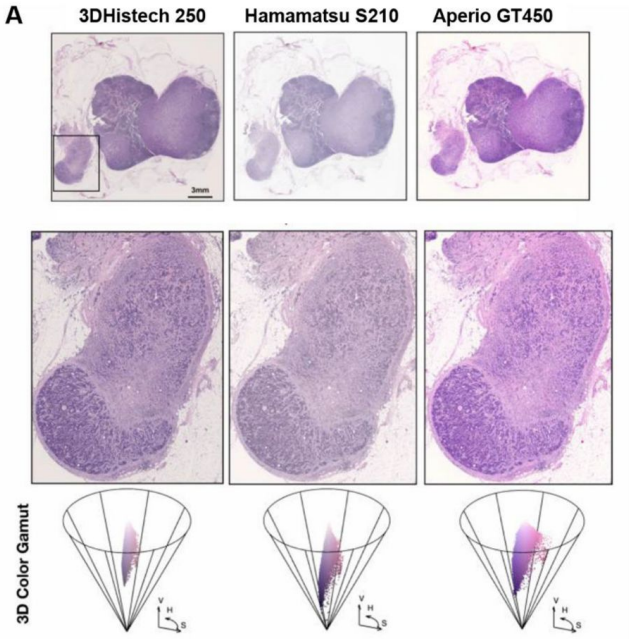
Trustworthy:
robust & reliable

Actionable:
valuable & sustainable

Responsible:
fair & transparent

Examples of Distribution Shift & Batch Effects in Pathology

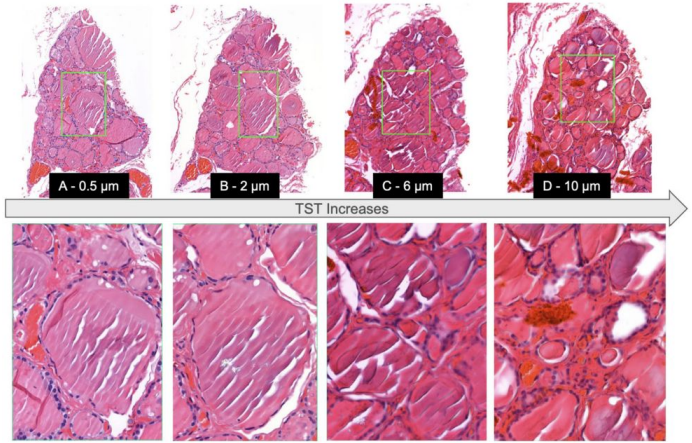
Scanner Variability



Source: Chen, Algorithm fairness in artificial intelligence for medicine and healthcare, 2023

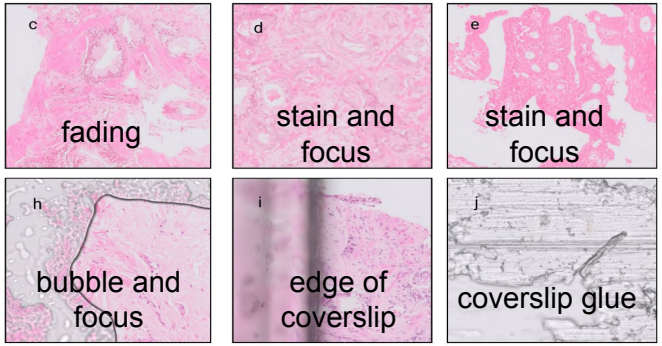
Tissue Thickness Variations

Source: Shah, Impact of Tissue Thickness on Computational Quantification of Features in Whole Slide Images for Diagnostic Pathology, 2025

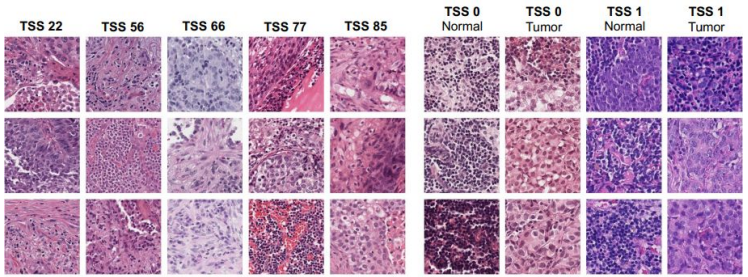


Slide Preparation and Image Acquisition Artifacts

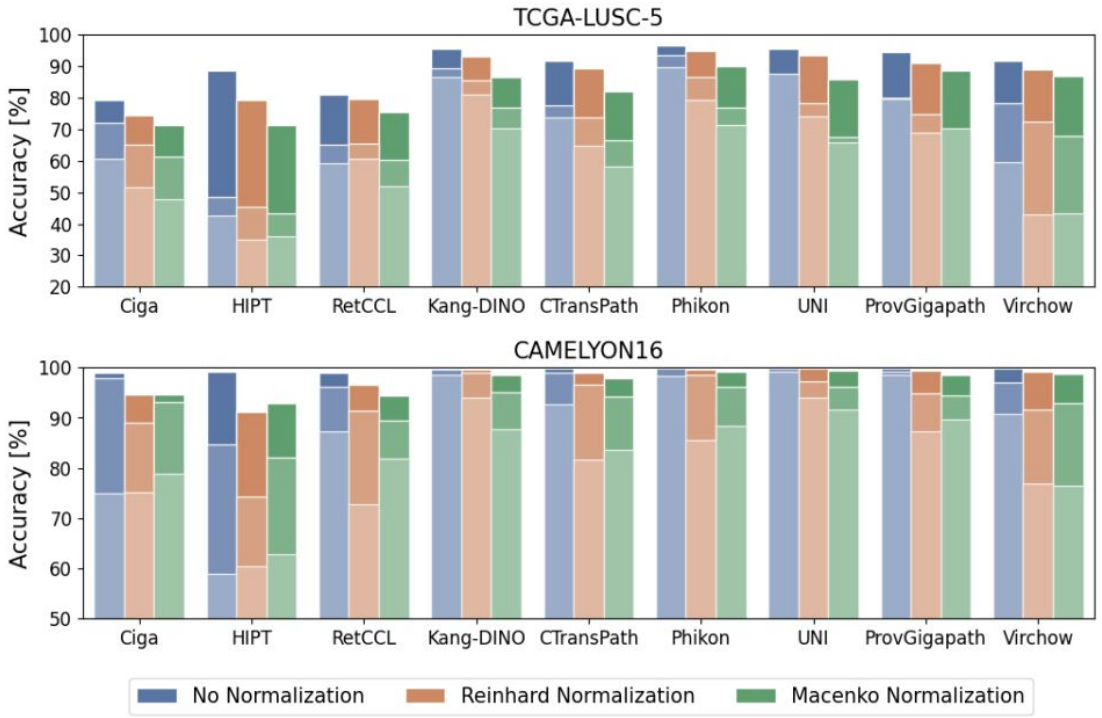
Source: Haghighat, PathProfiler: Automated Quality Assessment of Retrospective Histopathology Whole-Slide Image Cohorts by Artificial Intelligence – A Case Study for Prostate Cancer Research, 2021



Foundation Models Encode Batch Effects



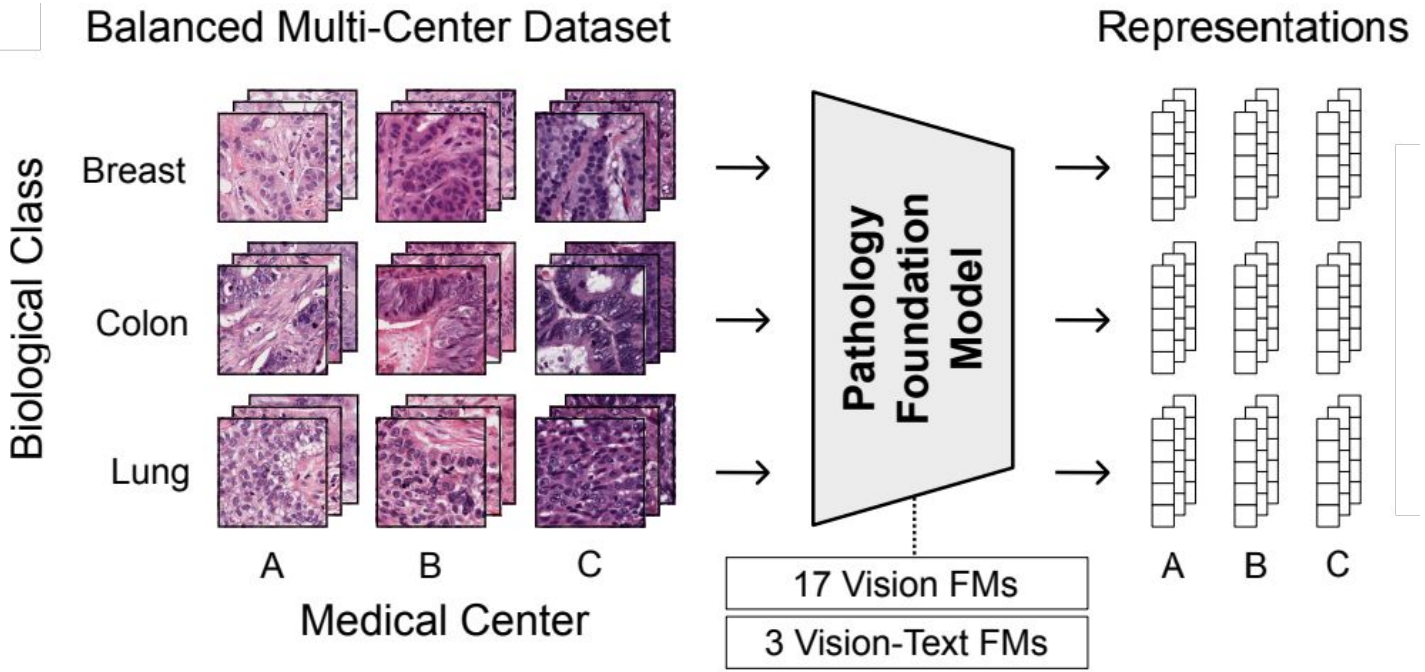
Tissue source site prediction



Source: Kömen, Do Histopathological Foundation Models Eliminate Batch Effects? A Comparative Study, 2024

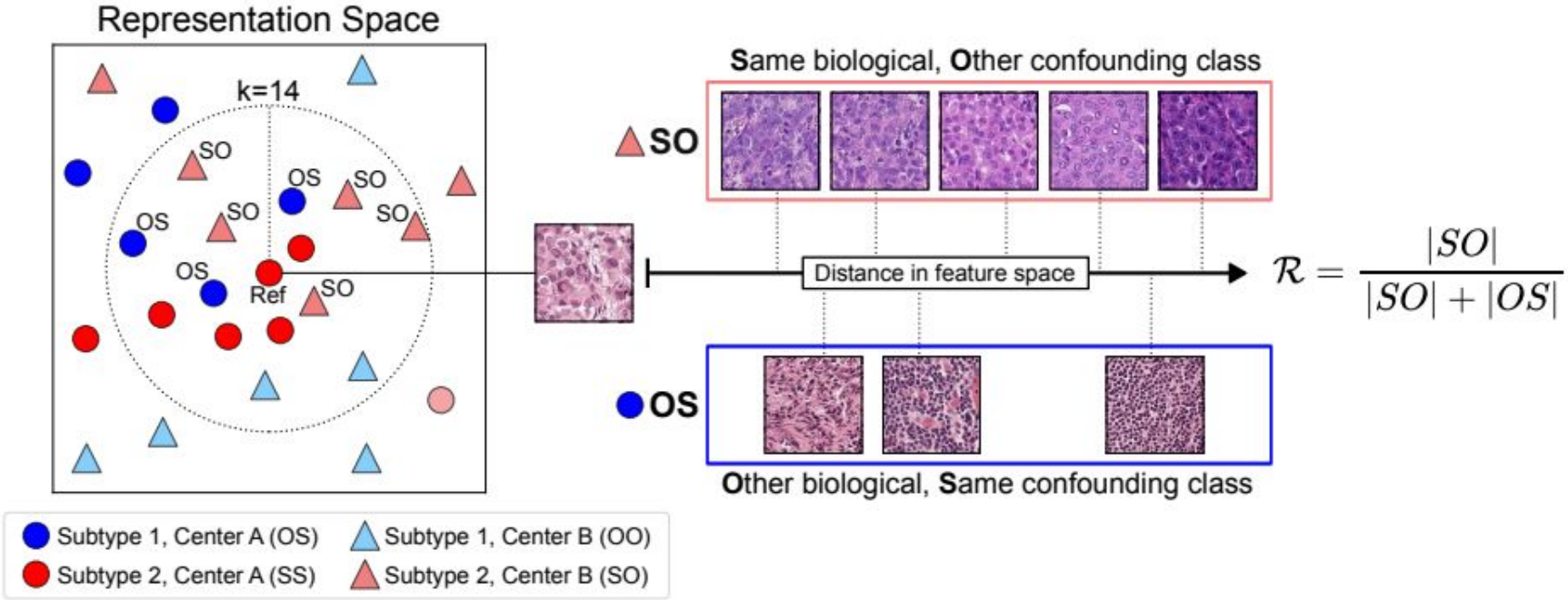
Benchmarking Robustness

PathoROB Benchmark



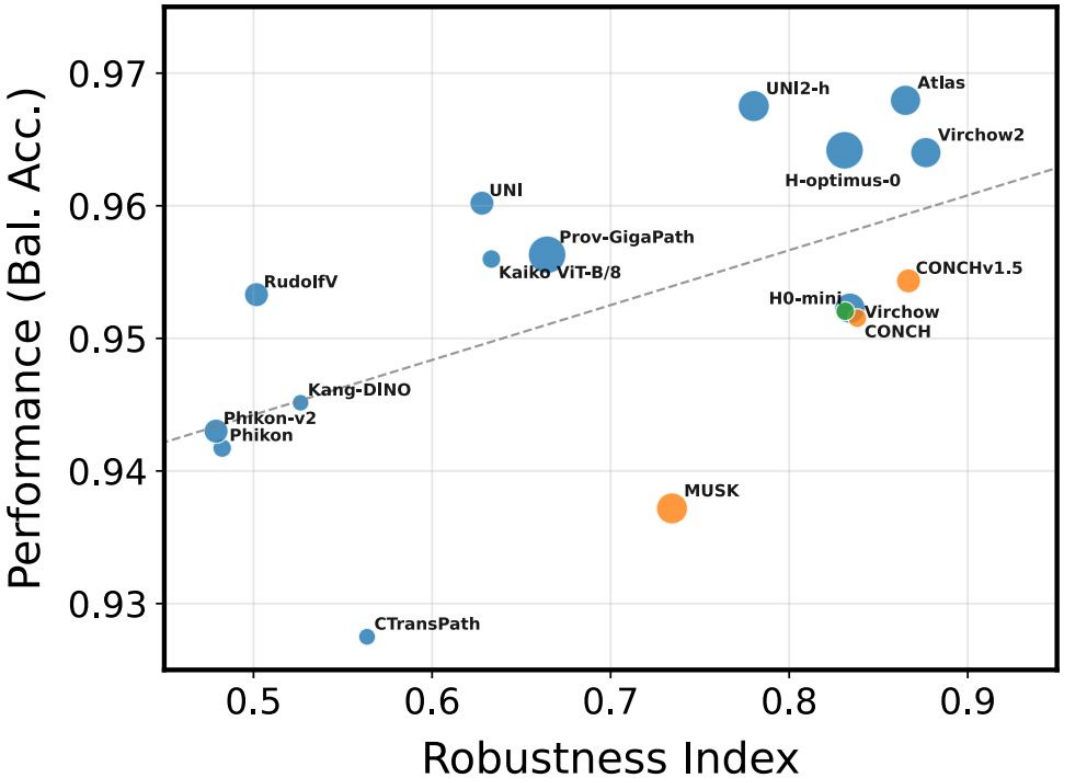
Source: Komen, Towards Robust Foundation Models for Digital Pathology, 2025

Measuring Model Robustness



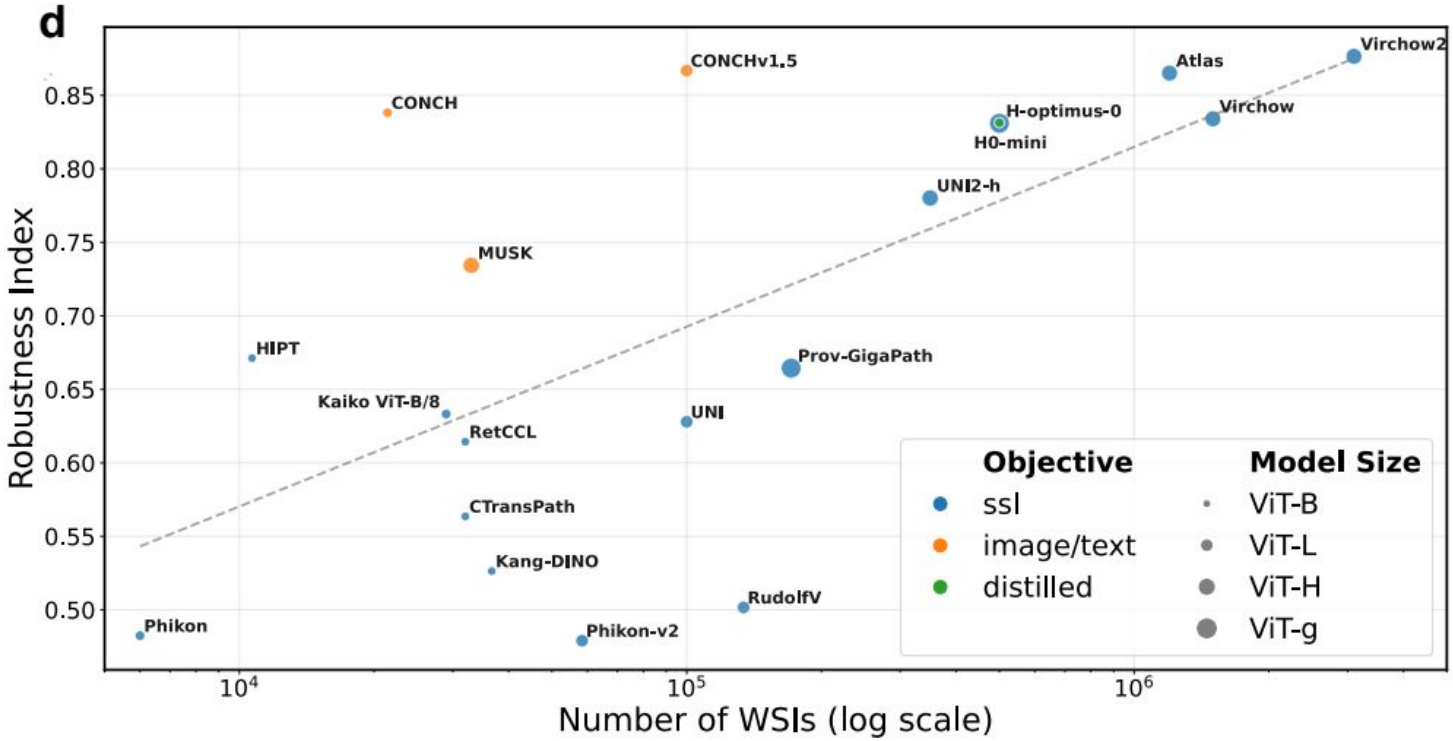
Source: Komen, Towards Robust Foundation Models for Digital Pathology, 2025

Some Models Are More Robust Than Others



Source: Komen, Towards Robust Foundation Models for Digital Pathology, 2025

What Makes a Model More Robust?



Larger training set?

More diverse training set?

Multimodal objective?

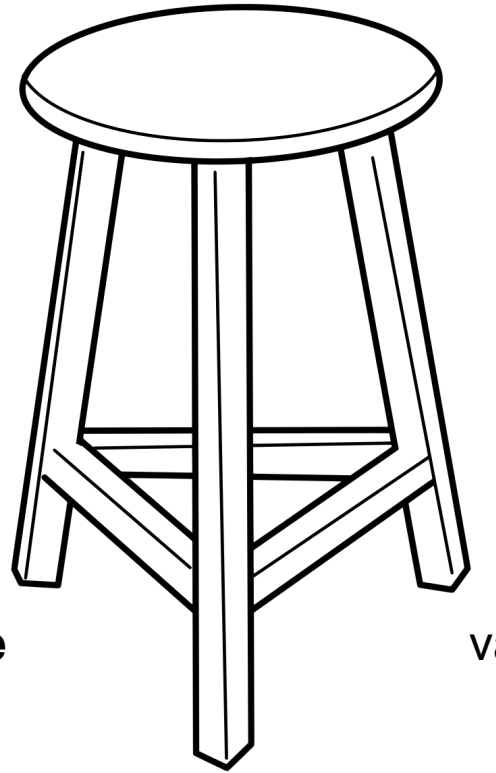
Source: Komen, Towards Robust Foundation Models for Digital Pathology, 2025

From Promise to Practice: Three Pillars

Beware of technical variations.

Some FMs are more robust than others.

All are susceptible to batch effects.

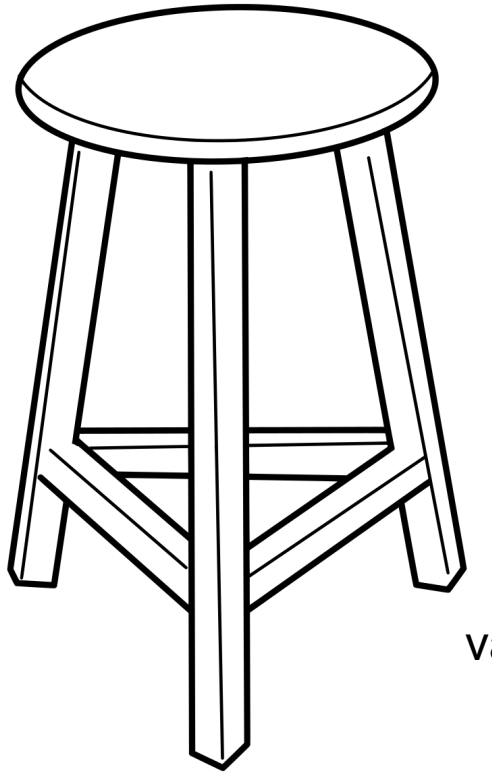


Trustworthy:
robust & reliable

Actionable:
valuable & sustainable

Responsible:
fair & transparent

From Promise to Practice: Three Pillars



Trustworthy:
robust & reliable

Actionable:
valuable & sustainable

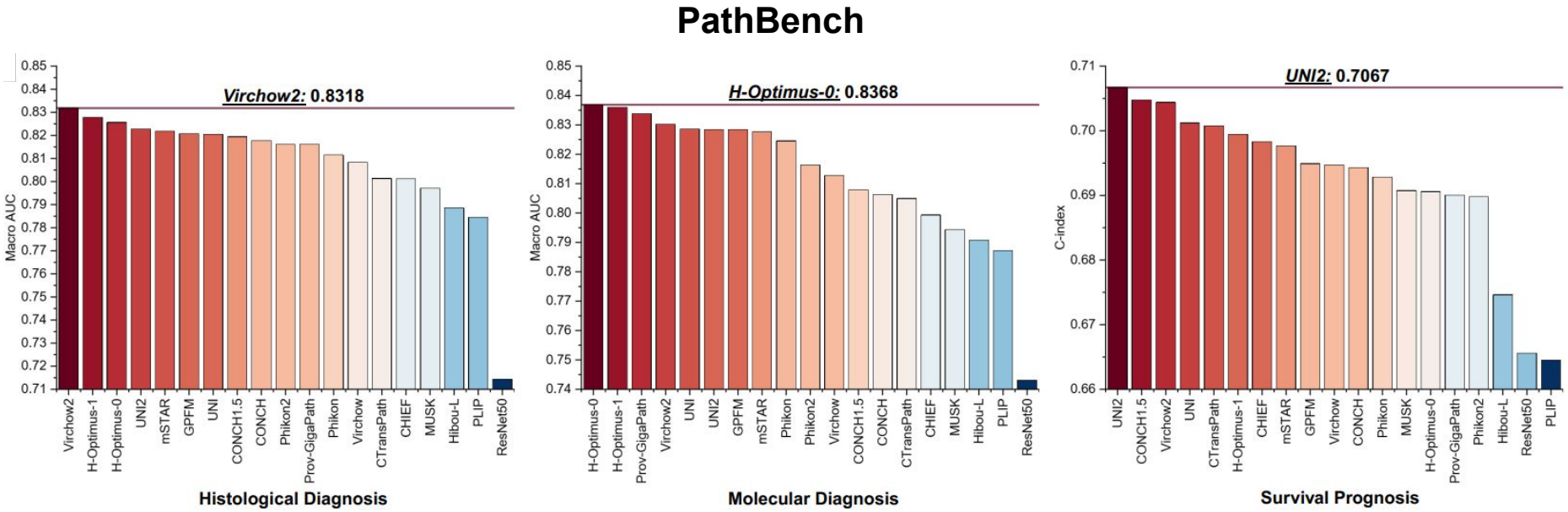
Responsible:
fair & transparent

The Limitations of Benchmarks

Model	EVA	PathBench	Stanford	HEST	Patho-Bench	Sinai	STAMP	THUNDER	PathoROB
Virchow2 NC ViT-H/14 (632M)	1/5	1/8	1/6	2/9	1/2	2/4	1/4	2/7	1/7
Prov-GigaPath Open ViT-g/14 (1.1B)	4/5	4/8	2/8	5/9	-	1/4	2/4	4/7	5/7
Phikon-v2 NC ViT-L/14 (307M)	5/5	5/8	7/8	8/9	-	3/4	-	7/7	7/7
UNI2 NC ViT-H/14 (682M)	2/5	2/8	3/8	1/9	2/2	-	-	1/7	4/7
Hibou-L Open ViT-L/14 (307M)	3/5	8/8	6/8	4/9	-	-	4/4	5/7	-
CTransPath Open Swin-T/4 (28M)	-	7/8	8/8	9/9	-	4/4	3/4	-	6/7
H0-mini NC ViT-B/14 (86M)	-	-	4/8	3/9	-	-	-	3/7	3/7
CONCH 1.5 NC ViT-L/16 (307M)	-	6/8	-	7/9	-	-	-	6/7	2/7
GPFM NC ViT-L/14 (303M)	-	3/8	5/8	6/9	-	-	-	-	-

Source: https://wearewaiv.github.io/histoboard

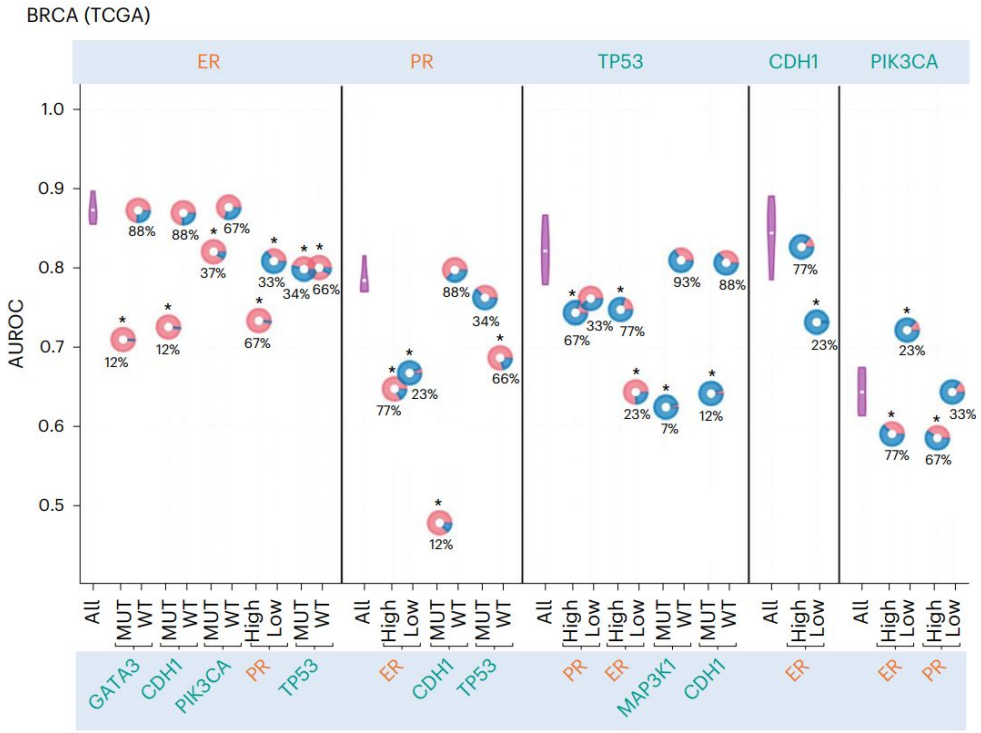
Varied Performance Across Types of Tasks

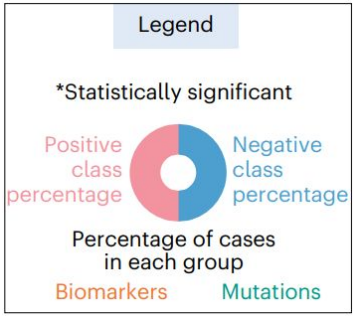


Source: Ma, PathBench: A comprehensive comparison benchmark for pathology foundation models towards precision oncology, 2025

Beyond Global Metrics

Stratified performance of WSI-based biomarker predictors

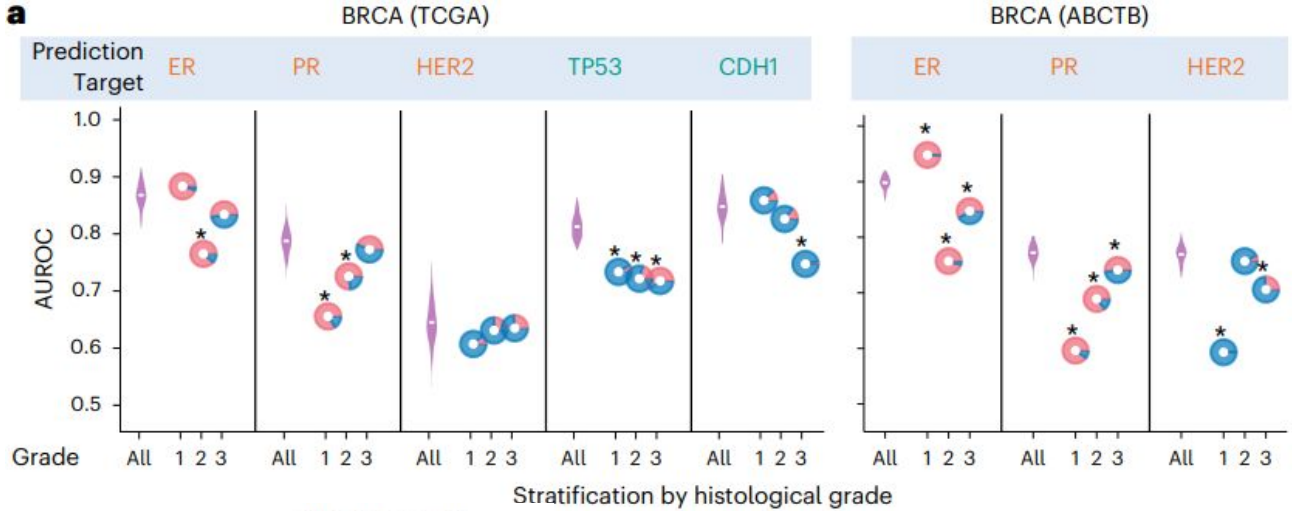


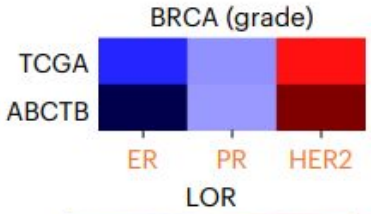


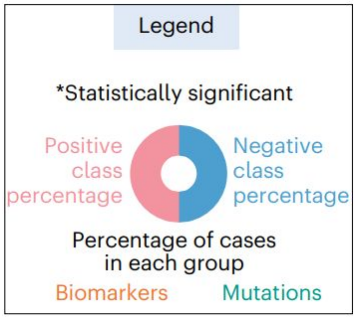
Source: Dawood, Confounding factors and biases abound when predicting molecular biomarkers from histological images, 2026

Checking for Bias

Biased performance of WSI-based biomarker predictors across patients with different histological grades







Source: Dawood, Confounding factors and biases abound when predicting molecular biomarkers from histological images, 2026

Is the model capturing the relevant biological signal or relying on confounders?

Where Does Pretraining Data Come From?

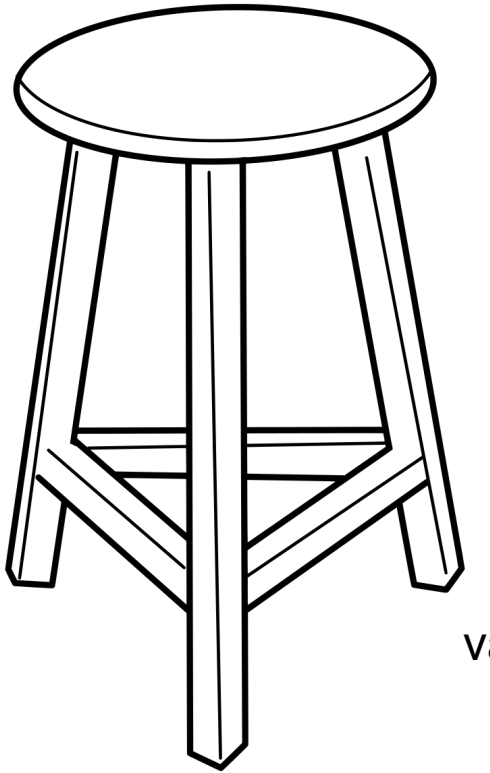
Model	Organiz-ation	Release date	# Cases	# WSIs	# Tiles	# Scan-ners	# Sites	Stains	Mag	# Tissue types
mSTAR	HKUST	2024-03	26k	11k	116m	NA	>30	H&E	20x	32
Hibou-L	Hist.ai	2024-06	306k	1.1m	1.2b	NA	NA	H&E, non-H&E, cytology	NA	Human + vet
Prov-GigaPath	Microsoft	2024-06	30k+	171k	1.4b	NA	28	H&E, IHC	20x	31 tissue types
TITAN	Harvard	2024-07	NA	336k	NA	16	3	H&E, IHC, special	20	20 organ types
Virchow2	Paige	2024-08	225k	3.1m	1.9b	NA	diver se	H&E, IHC	5/10/20/40	200 tissue types
Phikon-v2	Owkin	2024-09	NA	58k	456m	>50	>100	NA	20x	>30 sites
CHIEF	Harvard	2024-09	NA	61k	15m	NA	NA	H&E	10x	19
CanvOI	Imagene	2024-09	NA	633k	70m	NA	100	Mostly H&E	20x	40
UNI2	Harvard	2025-01	NA	>350k	200m	NA	1	H&E, IHC	20x	>20 organ types
H-optimus-1	Bioptimus	2025-03	>800k	>1m	billions	3	4000	H&E	NA	50 organs
Midnight-12k	kaiko.ai	2025-04	NA	12k	NA	NA	NA	NA	5/10/20/40	32 cancer types
Atlas 2	Aignostics	2026-01	NA	5.5m	NA	NA	3	NA	5/10/20/40	NA
GenBio-PathFM	GenBio AI	2026-03	NA	177k	400m	NA	NA	H&E, IHC	5/10/20/40	7

From Promise to Practice: Three Pillars

Benchmarking helps, but what really matters is performance on your data.

Validate thoroughly with stratified metrics and check for confounders.

Ask questions about FM training data.

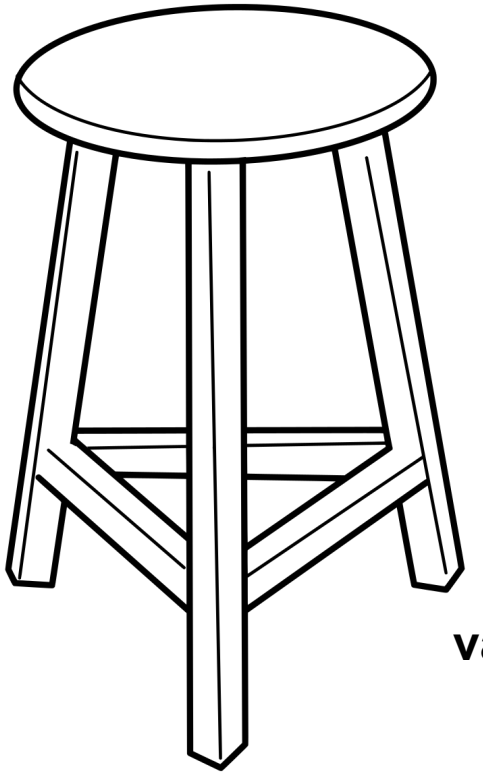


Trustworthy:
robust & reliable

Actionable:
valuable & sustainable

Responsible:
fair & transparent

From Promise to Practice: Three Pillars

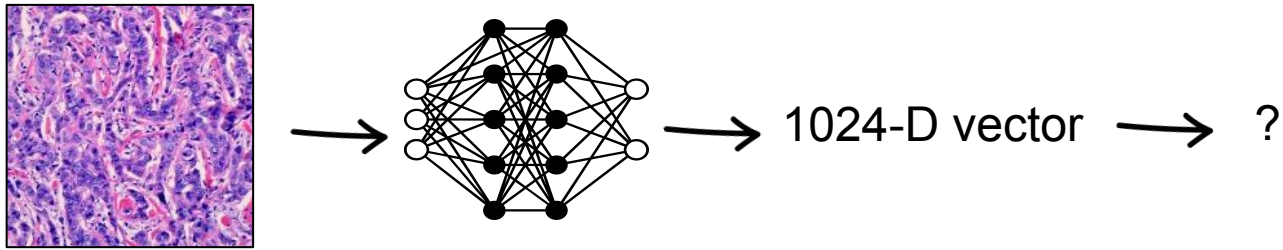


Trustworthy:
robust & reliable

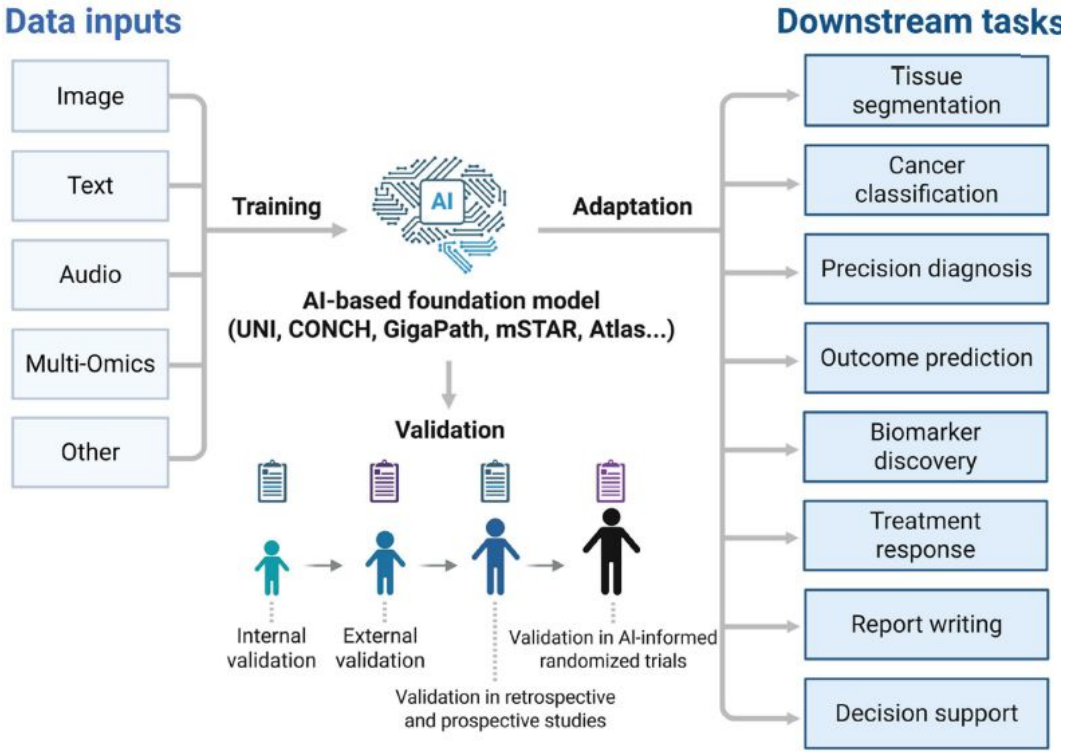
Actionable:
valuable & sustainable

Responsible:
fair & transparent

Embeddings Aren’t Diagnoses

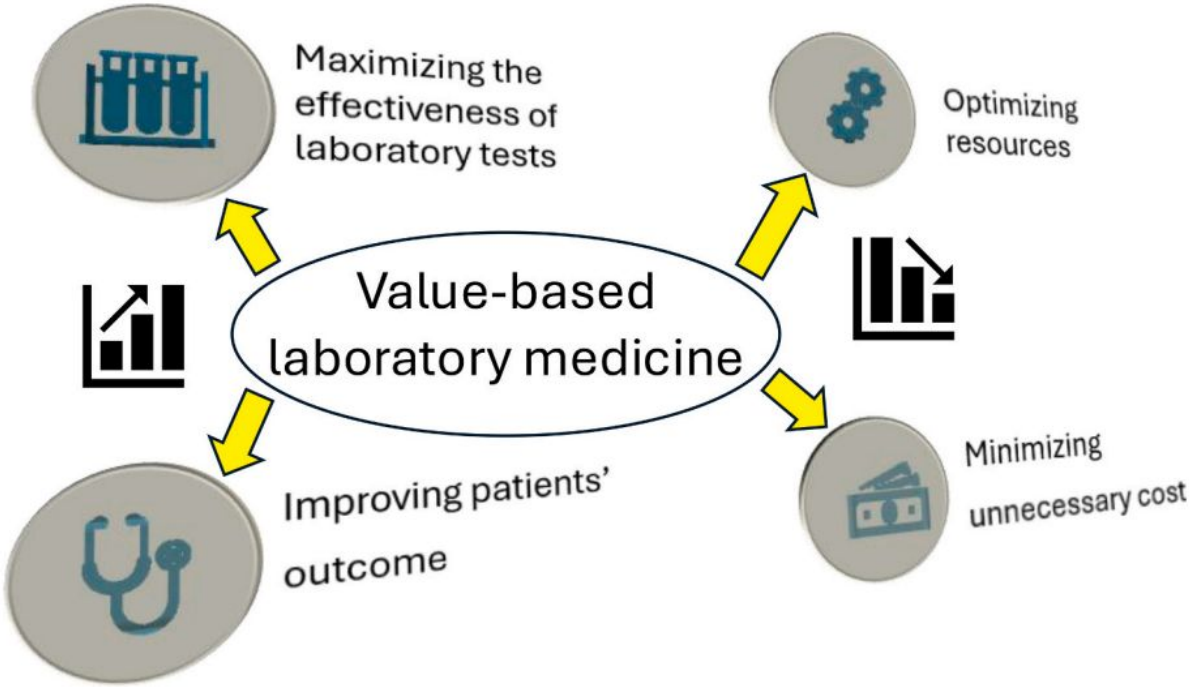


How Do We Use Foundation Models?



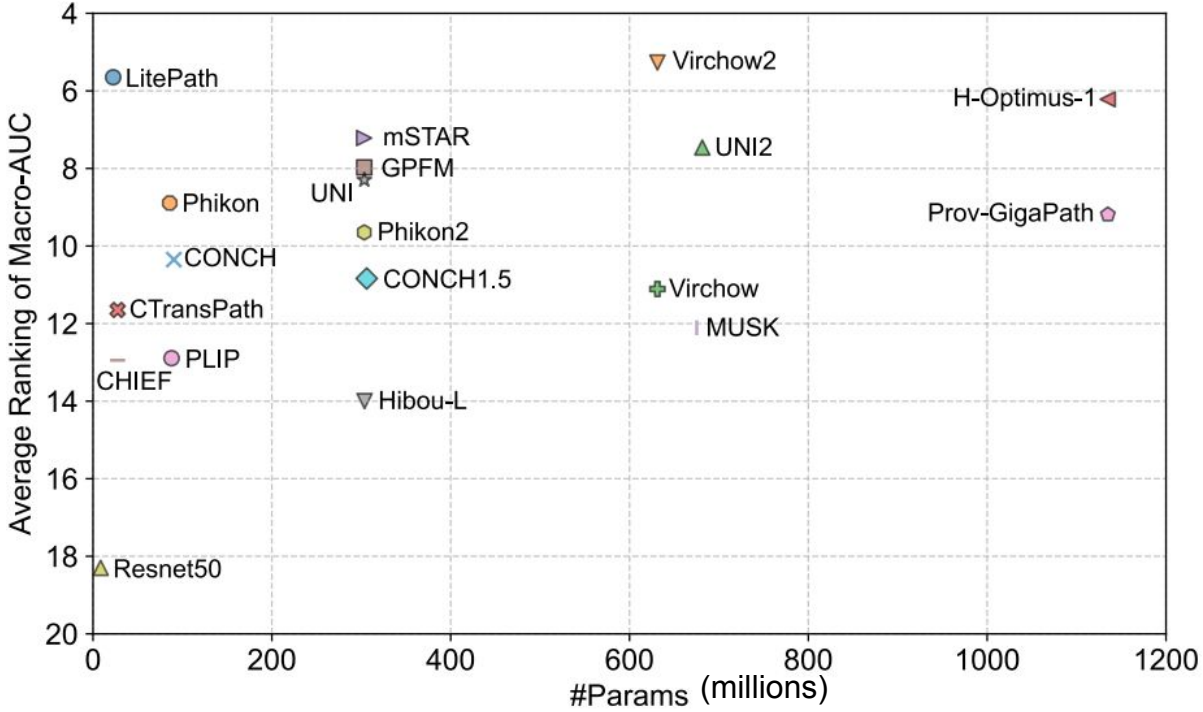
Source: Cheng, The role of artificial intelligence-based foundation models and “copilots” in cancer pathology: potential and challenges, 2026

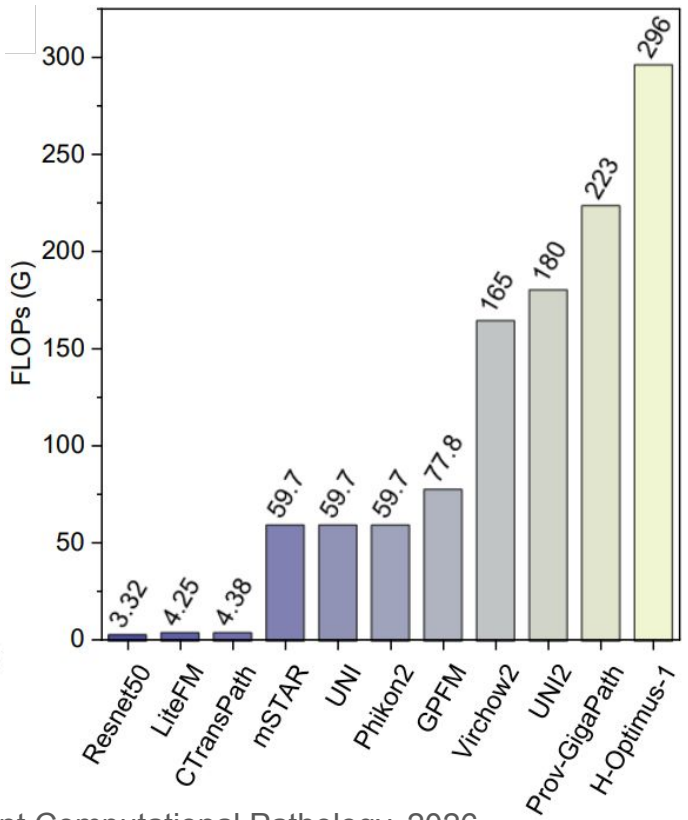
The Value Proposition



Source: Plebani, Promoting value-based laboratory medicine: Moving towards an innovative model of clinical laboratory, 2025

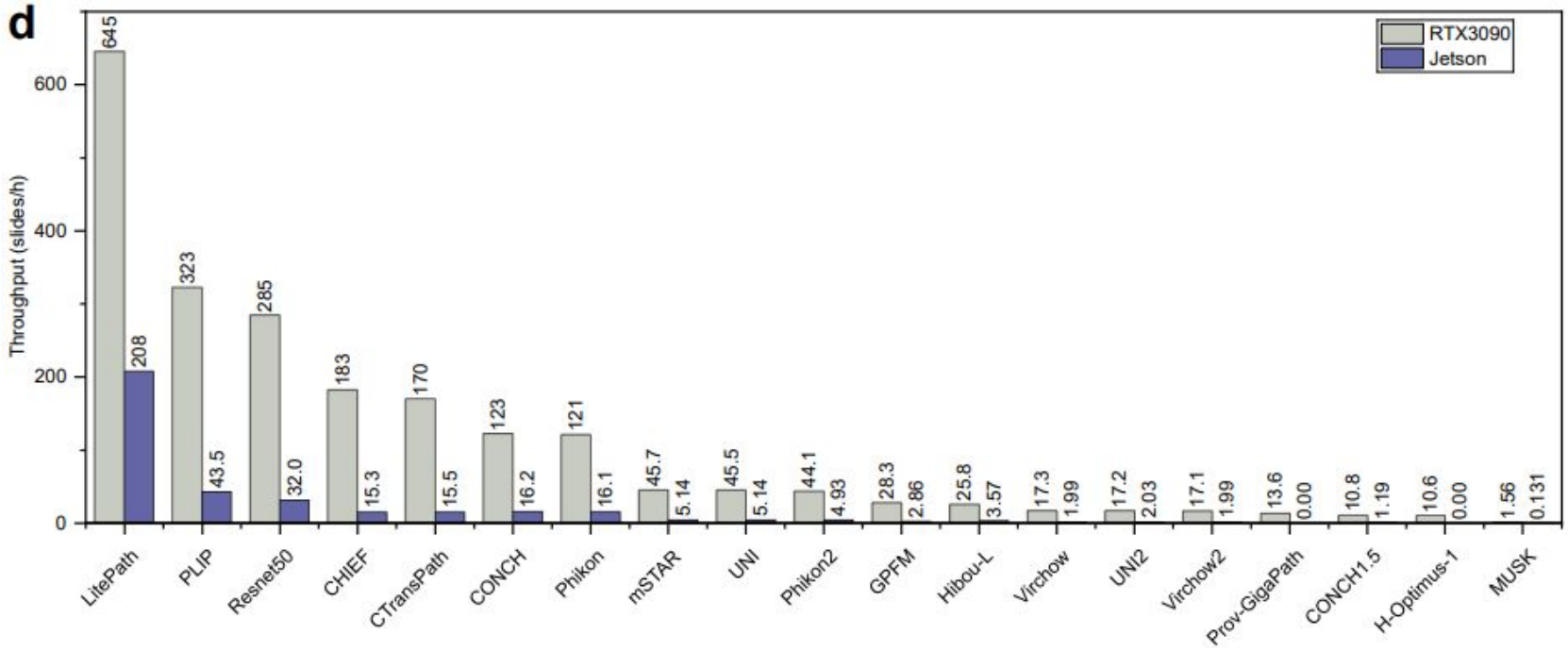
Does Model Size Matter?





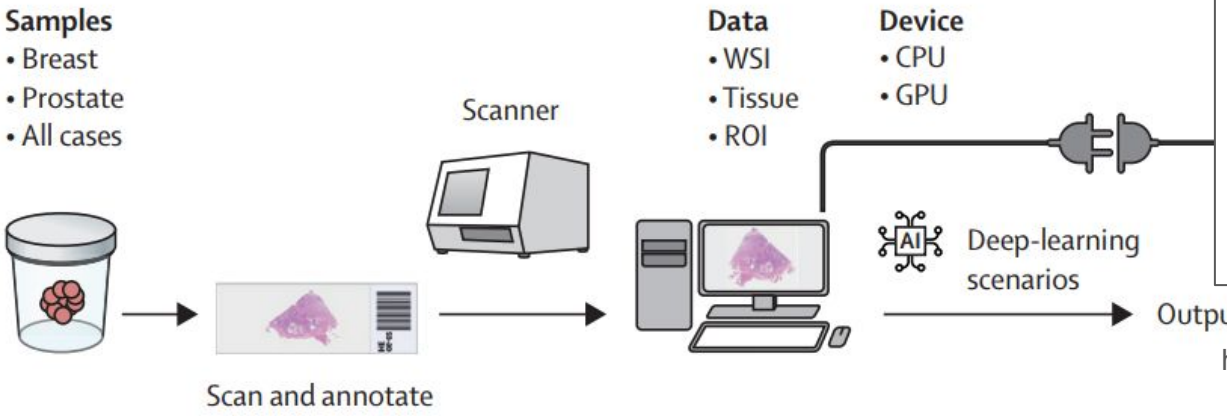
Source: Cai, A Deployment-Friendly Foundational Framework for Efficient Computational Pathology, 2026

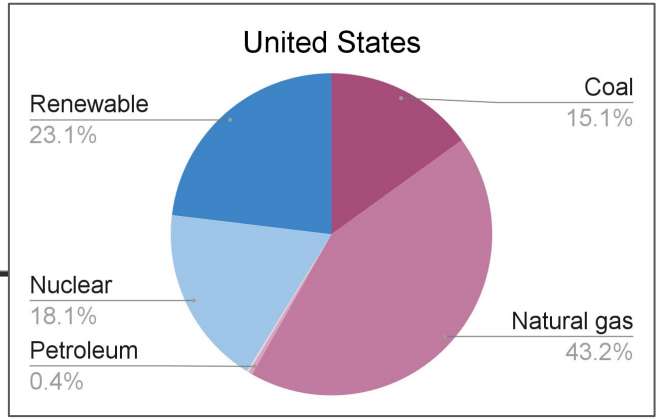
A Comparison of Slide Throughput



Source: Cai, A Deployment-Friendly Foundational Framework for Efficient Computational Pathology, 2026

Processing of Every WSI Uses Electricity & Produces Emissions

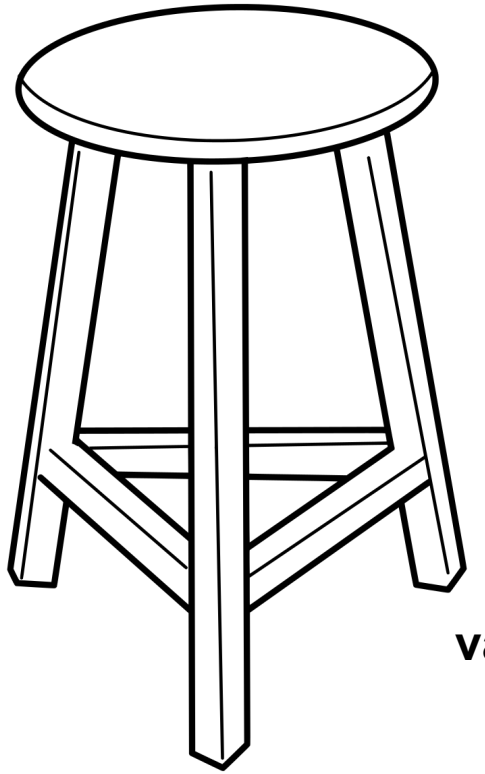




Source: https://www.eia.gov/energyexplained/us-energy-facts/data-and-statistics.php

Source: Sadr, Operational greenhouse-gas emissions of deep learning in digital pathology: a modelling study, 2024

From Promise to Practice: Three Pillars



Trustworthy:
robust & reliable

Actionable:
valuable & sustainable

Responsible:
fair & transparent

Integrate into clinical workflow.

Have a clear value proposition.

Be mindful of compute costs.

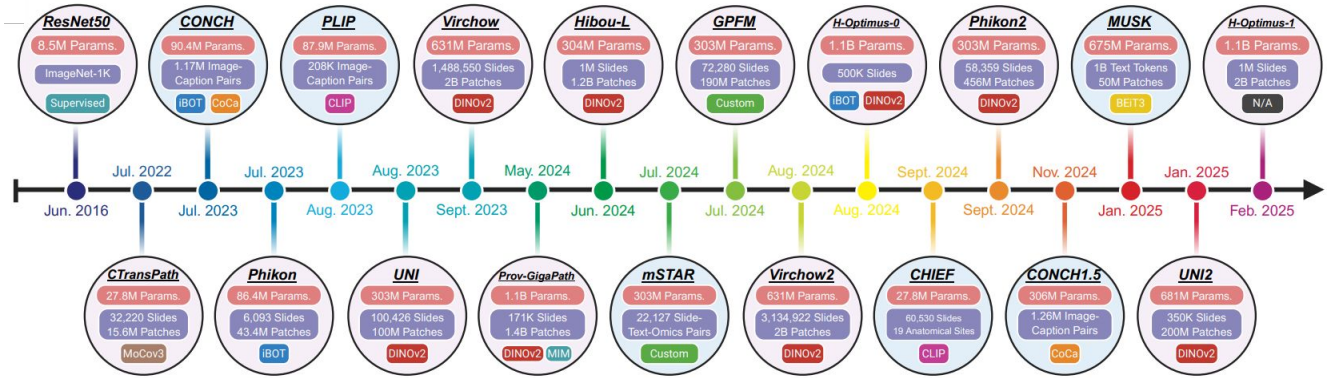
Closing Thoughts

Foundation models are the key components of future digital pathology tools.

They must be built on the pillars of trust, responsibility, and actionability.

Using more powerful and robust FMs doesn't change the need for thorough validation.

Whether you are developing FMs, building products with them, or simply an end user…
Ask questions about the models, how they were trained, and how they were validated.



Source: Ma, PathBench: A comprehensive comparison benchmark for pathology foundation models towards precision oncology, 2025

Resources

Computer Vision Insights Newsletter

Practical briefings to help you close the gap between algorithm and impact — for people, planet, and science
https://pixelscientia.com/newsletter/

Contact

Heather Couture
heather@pixelscientia.com

https://www.linkedin.com/in/hdcouture/

Pixel Scientia LABS