16th CONGRESS Lung ON CANCER

BARCELONA 27 / 28 NOVEMBER 2025

GECP lung cancer research

AI in digital pathology and multiplex immunohistochemistry (IHC)

Sebastian Michels

University Hospital Cologne

... from a hemato-oncologist point of view!

GeCP
lung cancer research

CONFLICTO DE INTERESES

Grants or contracts from any entity (to institution)

Else-Kröner Forschungskolleg Cologne, German Cancer Aid, Dr. Rolf M. Schwiete Foundation, Amgen, AstraZeneca, Bristol-Myers Squibb, Janssen-Cilag, Novartis

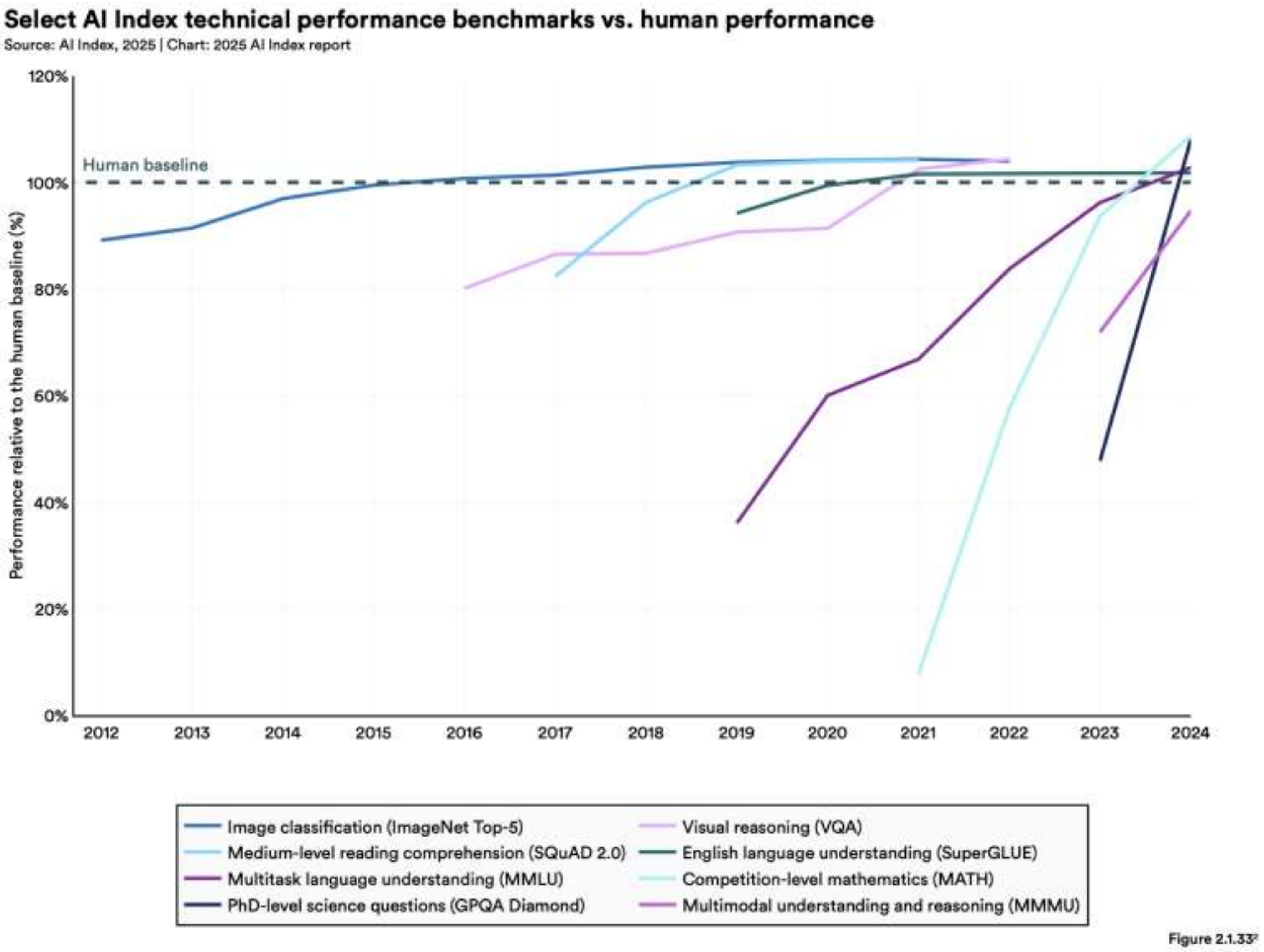
Payment or honoraria (personal)

BeOne Medicines, Bristol-Myers Squibb, Takeda, Janssen-Cilag, Pfizer

Support for attending meetings and/or travel (personal)

Nationales Centrum für Tumorerkrankungen West, BeOne Medicines, Janssen-Cilag, Eli Lilly and Company, Takeda

AI vs human performance

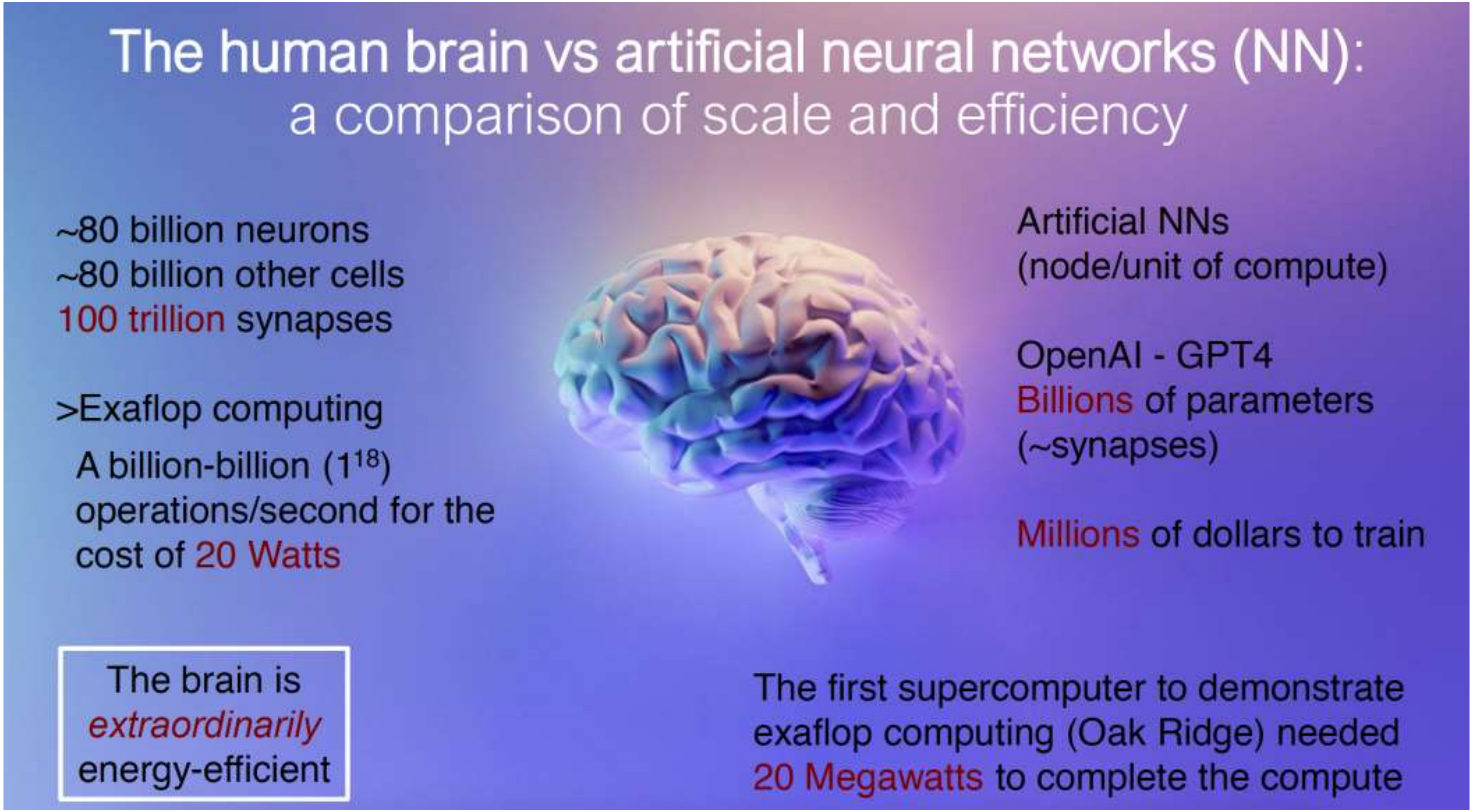


AI is rapidly outperforming humans in diverse tasks

and the rate is increasing...

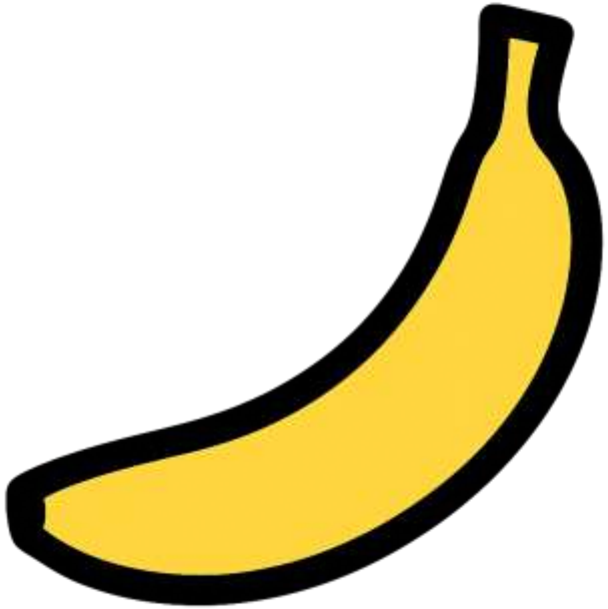
Adapted with permission of Christina Curtis from „AI-based Drug Development“ at ESMO 2025

Efficancy: The Brain’s Enduring Edge



Used with permission of Christina Curtis from „AI-based Drug Development“ at ESMO 2025

Efficancy: The Brain’s Enduring Edge



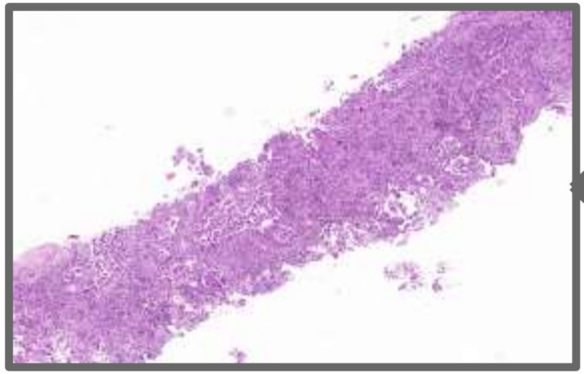
VS

1,000,000 bananas or ∼120 T

AI in digital pathology - ideally

Conventional pathology vs AI-enhanced digital pathology

Conventional



Tumor tissue
and conventional HE slide

Microscopy

IHC

Molecular analyses

Others

Final diagnosis

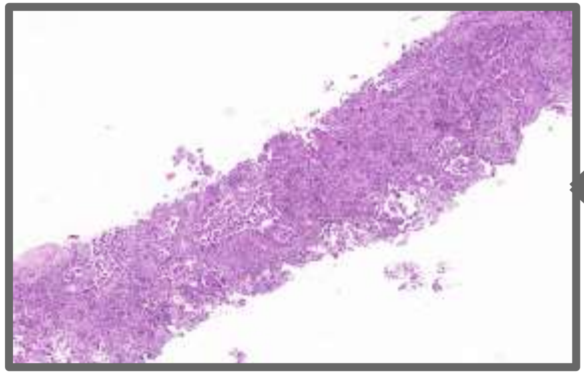
Final report

Digitalized slides provided by Yuri Tolkach

AI in digital pathology - ideally

Conventional pathology vs AI-enhanced digital pathology

Conventional

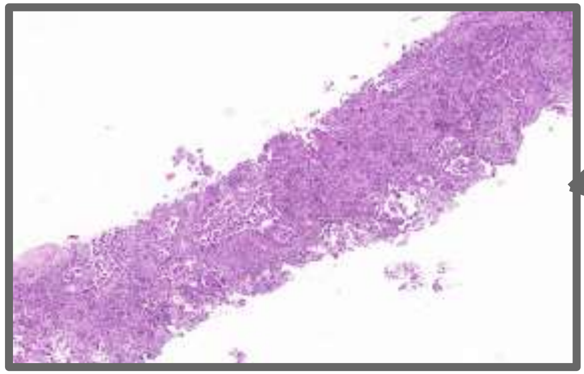


Tumor tissue
and conventional HE slide

- Microscopy
- IHC
- Molecular analyses
- Others

- Final diagnosis
- Final report

AI-enhanced



Digitalized HE slide
➔ Whole slide image (WSI)

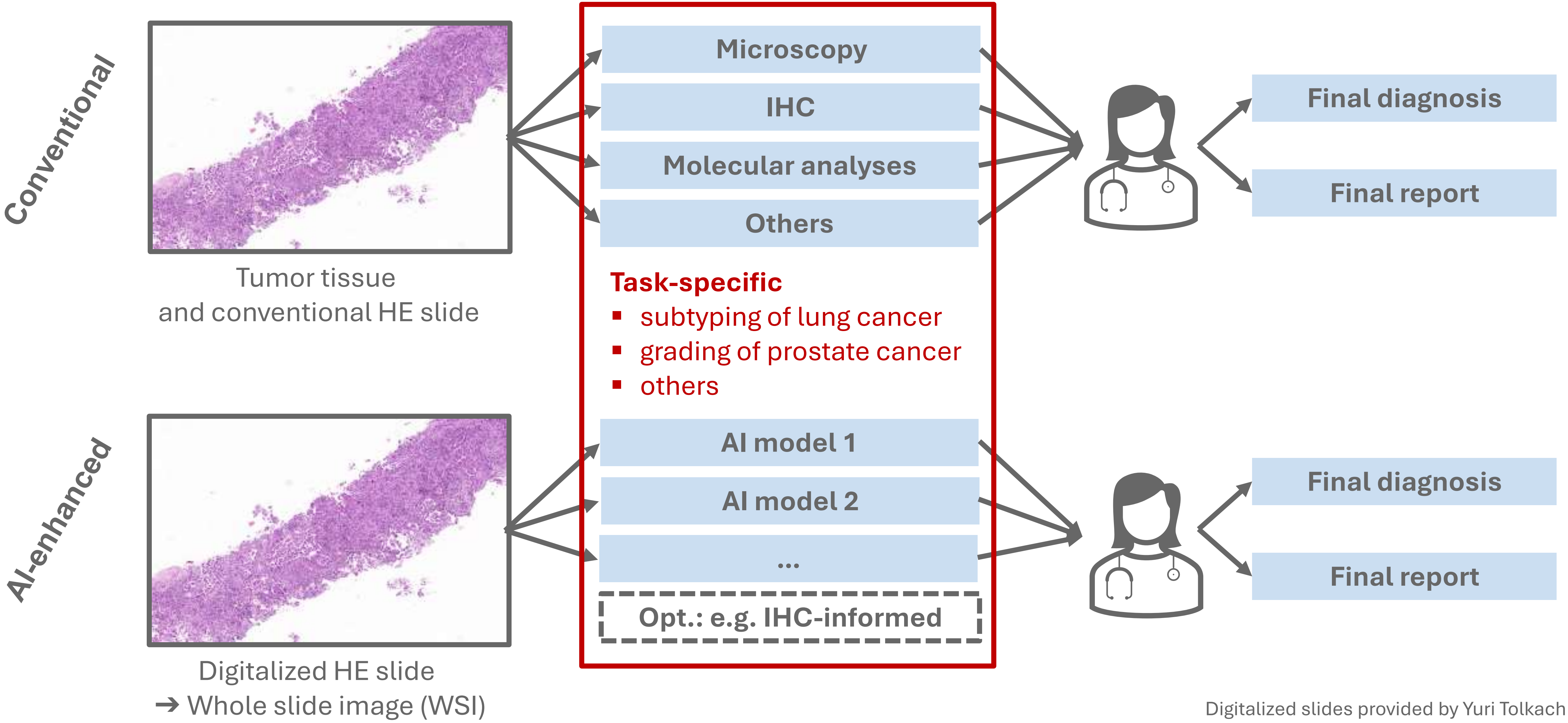
- AI model 1
- AI model 2
- ...
- Opt.: e.g. IHC-informed

- Final diagnosis
- Final report

Digitalized slides provided by Yuri Tolkach

AI in digital pathology - ideally

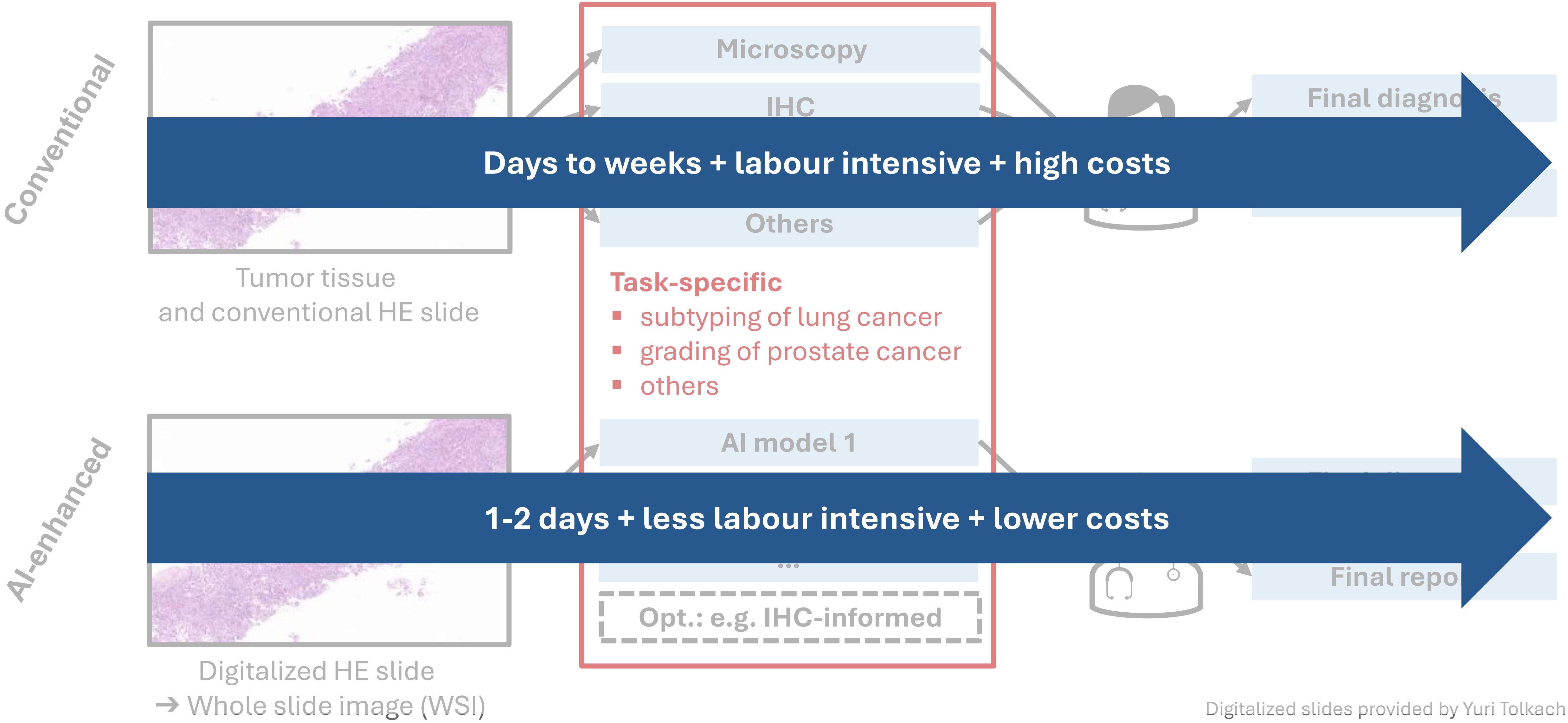
Conventional pathology vs AI-enhanced digital pathology



Digitalized slides provided by Yuri Tolkach

AI in digital pathology - ideally

Conventional pathology vs AI-enhanced digital pathology



Digitalized slides provided by Yuri Tolkach

AI applications in pathology and oncology

Basic to advanced = unimodal to multimodal

Types of AI application

Basic

- Tissue segmentation or tumor detection
- Grading
- Subtyping
- Marker quantification

Usually unimodal

Advanced

- Prediction of prognosis
- Prediction of therapy response
- Prediction of biomarkers

Usually multimodal

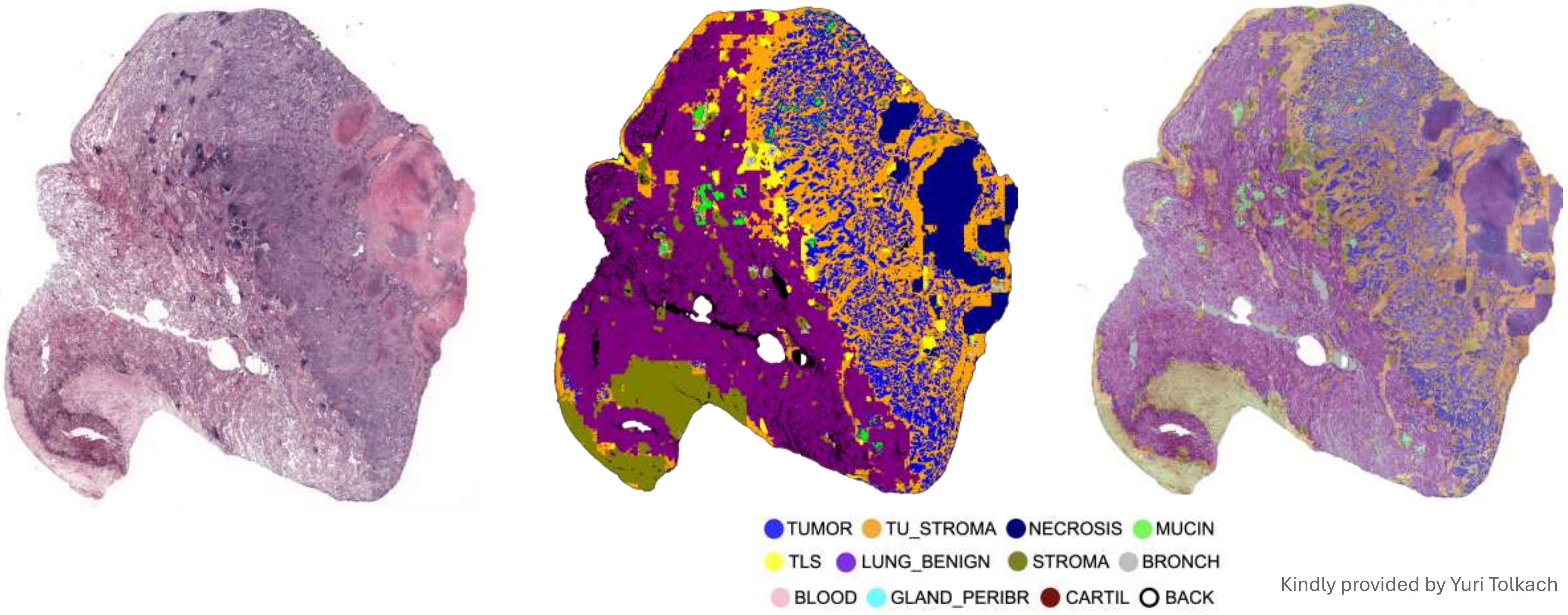
Downstream applications

Basic: Tissue segmentation

A multi-institutional model by the University Hospital Cologne

A supervised (explainable) model to segment tissue of digitalized HE-stained whole slide images (WSI)

➔ **Trained in 345 annotated WSI** and **validated in 4,097 WSI**

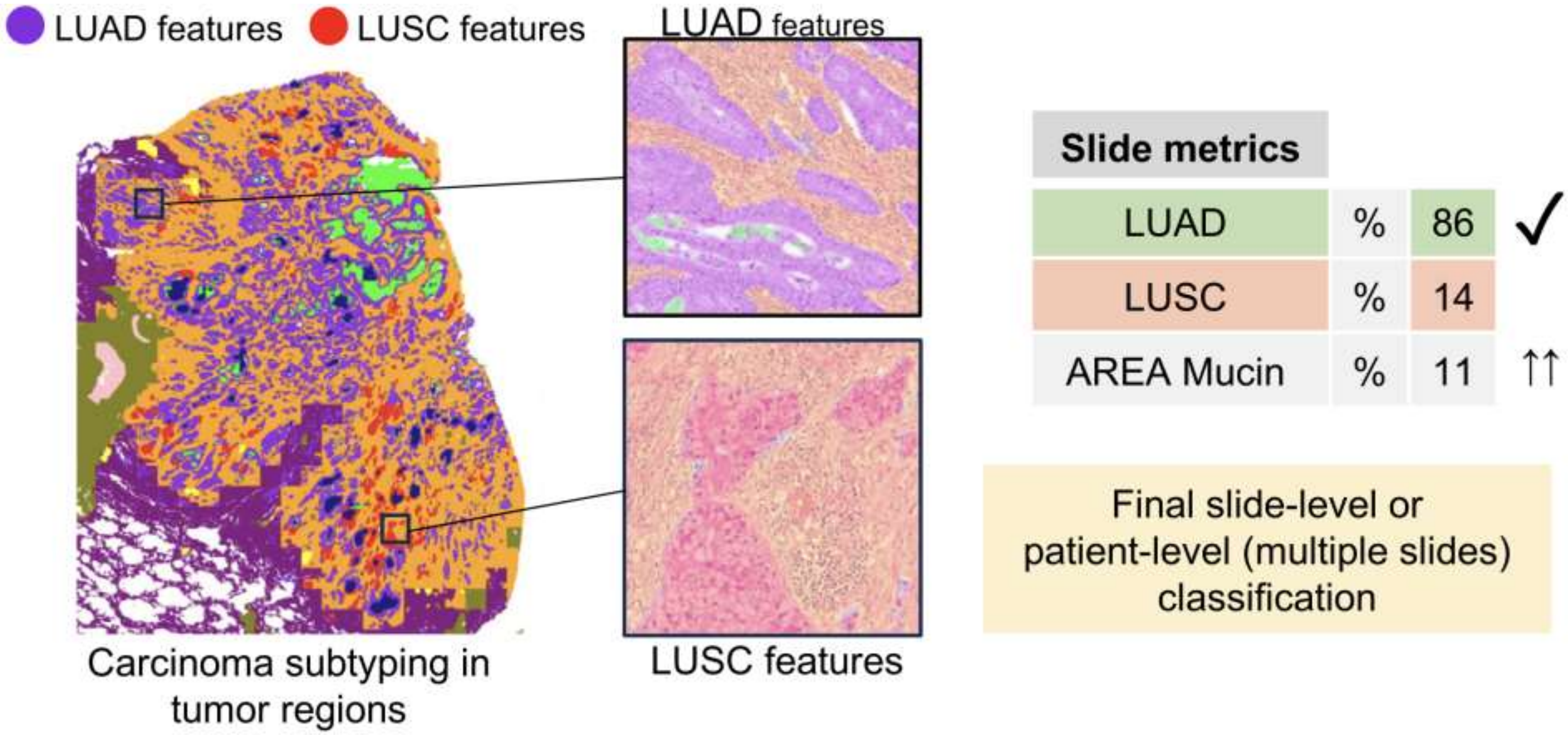


Kindly provided by Yuri Tolkach

Basic: Histologic subtyping

A multi-institutional model by the University Hospital Cologne

Histologic subtyping (LUAD vs LUSC) with a high overall accuracy of 0.929 to 0.978

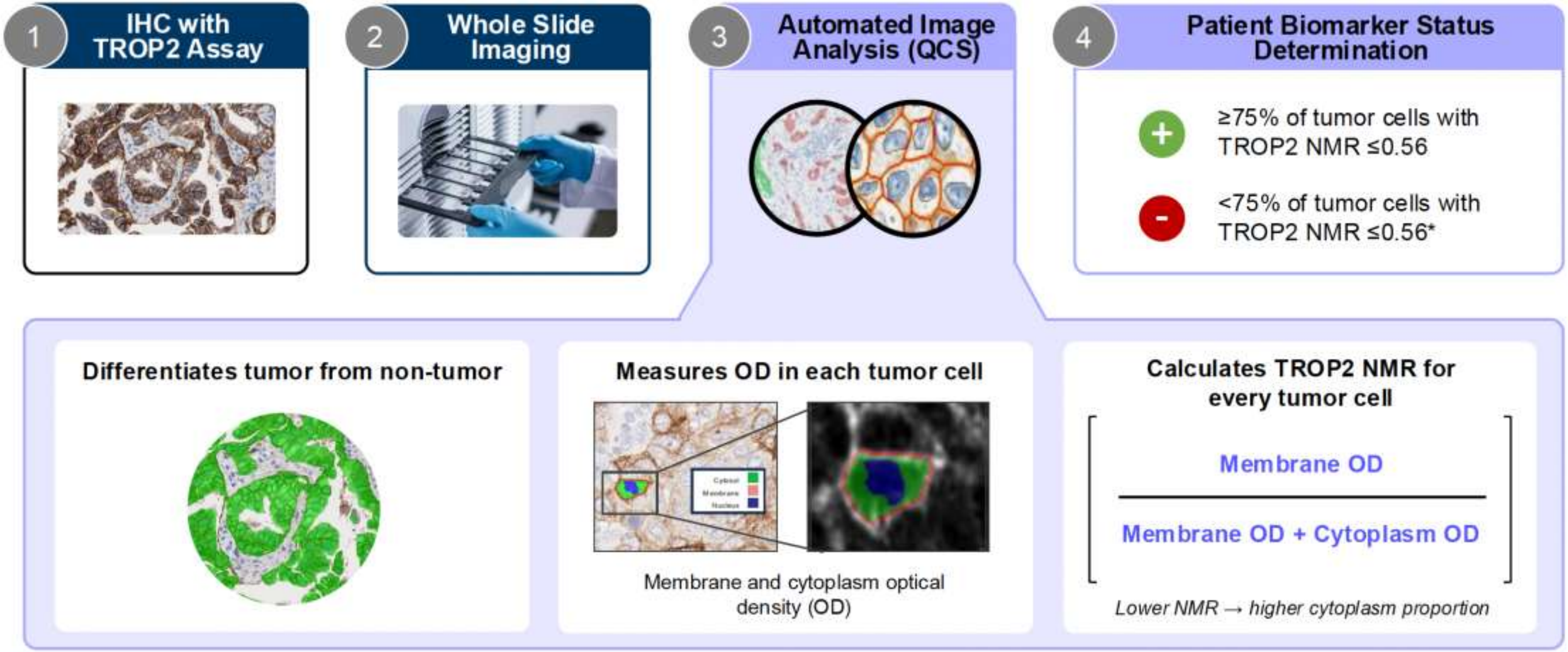


Kludt et al. Cell Rep Med 2024; doi: 10.1016/j.xcrm.2024.101697

Basic: Biomarker quantification

Quantitative continuous scoring (QCS) to quantify TROP2 expression

A supervised (explainable) model to determine the normalized membrane ratio (NMR) of surface markers

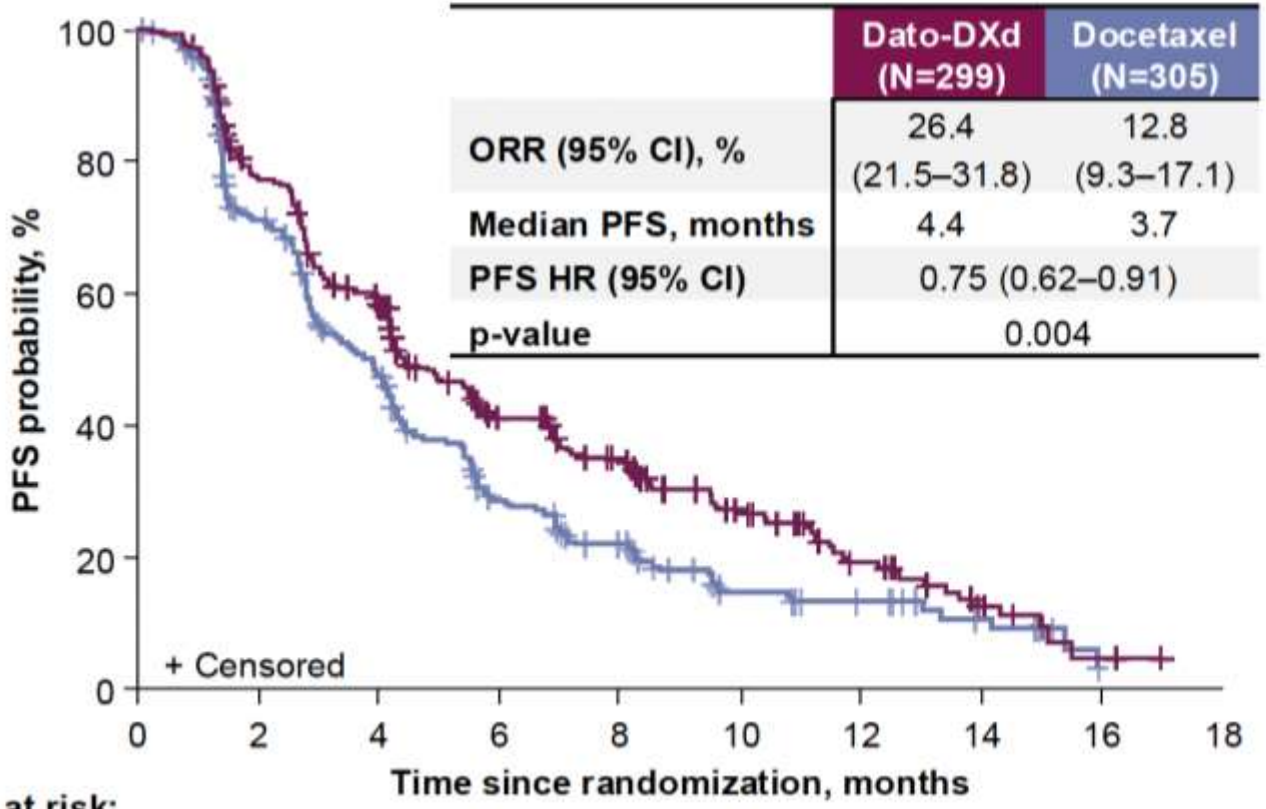


Garassino et al. WCLC 2024 – PL02.11

Basic: Biomarker quantification

Results from the TROPION-Lung01 study (NCT04656652)

PFS and ORR by BICR of the TROP2 ADC datopotamab-deruxtecan vs docetaxel



	Dato-DXd (N=299)	Docetaxel (N=305)
ORR (95% CI), %	26.4 (21.5–31.8)	12.8 (9.3–17.1)
Median PFS, months	4.4	3.7
PFS HR (95% CI)	0.75 (0.62–0.91)	
p-value	0.004	

No. at risk:

Dato-DXd	299	216	156	96	74	46	24	10	2	0
Docetaxel	305	186	120	63	42	19	14	7	0	0

- No difference in OS observed!
  - HR, 0.94 (0.78-1.14)[1]
- Conventional read out of TROP2 expression did not correlate efficacy[2]

Left, Garassino et al. WCLC 2024 – PL02.11
1. Ahn et al. J Clin Oncol 2024; doi.org/10.1200/JCO-24-01544; 2. Shimizu T, et al. J Clin Oncol 2023; doi: 10.1200/JCO.23.00059.

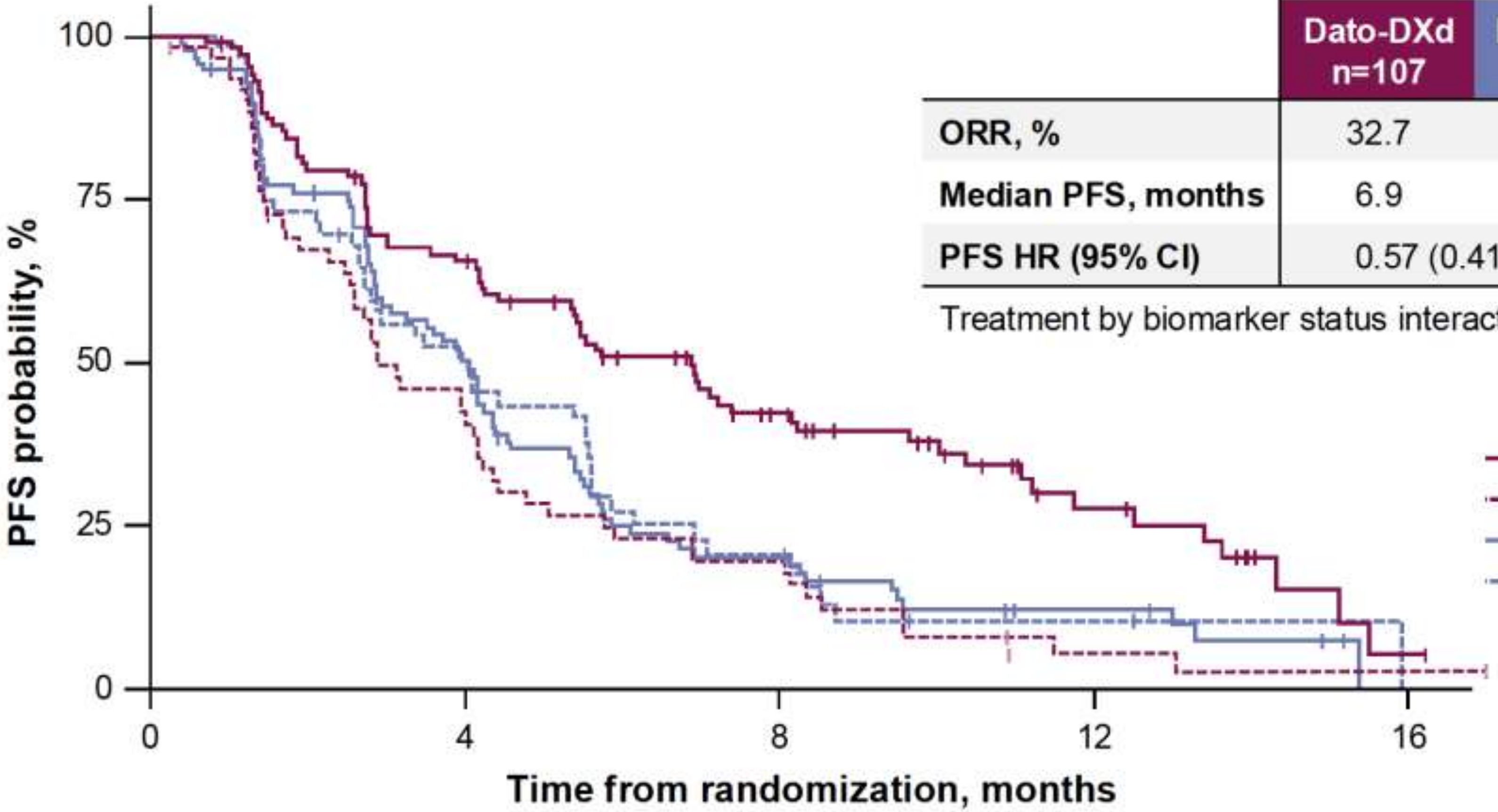
Basic: Biomarker quantification

Results from the TROPION-Lung01 study (NCT04656652)

PFS and ORR by BICR in patients stratified by TROP2 NMR (N=352)

	TROP2 QCS-NMR+		TROP2 QCS-NMR–	
	Dato-DXd n=107	Docetaxel n=107	Dato-DXd n=65	Docetaxel n=73
ORR, %	32.7	10.3	16.9	15.1
Median PFS, months	6.9	4.1	2.9	4.0
PFS HR (95% CI)	0.57 (0.41–0.79)		1.16 (0.79–1.70)	

Treatment by biomarker status interaction: p=0.0063



Garassino et al. WCLC 2024 – PL02.11

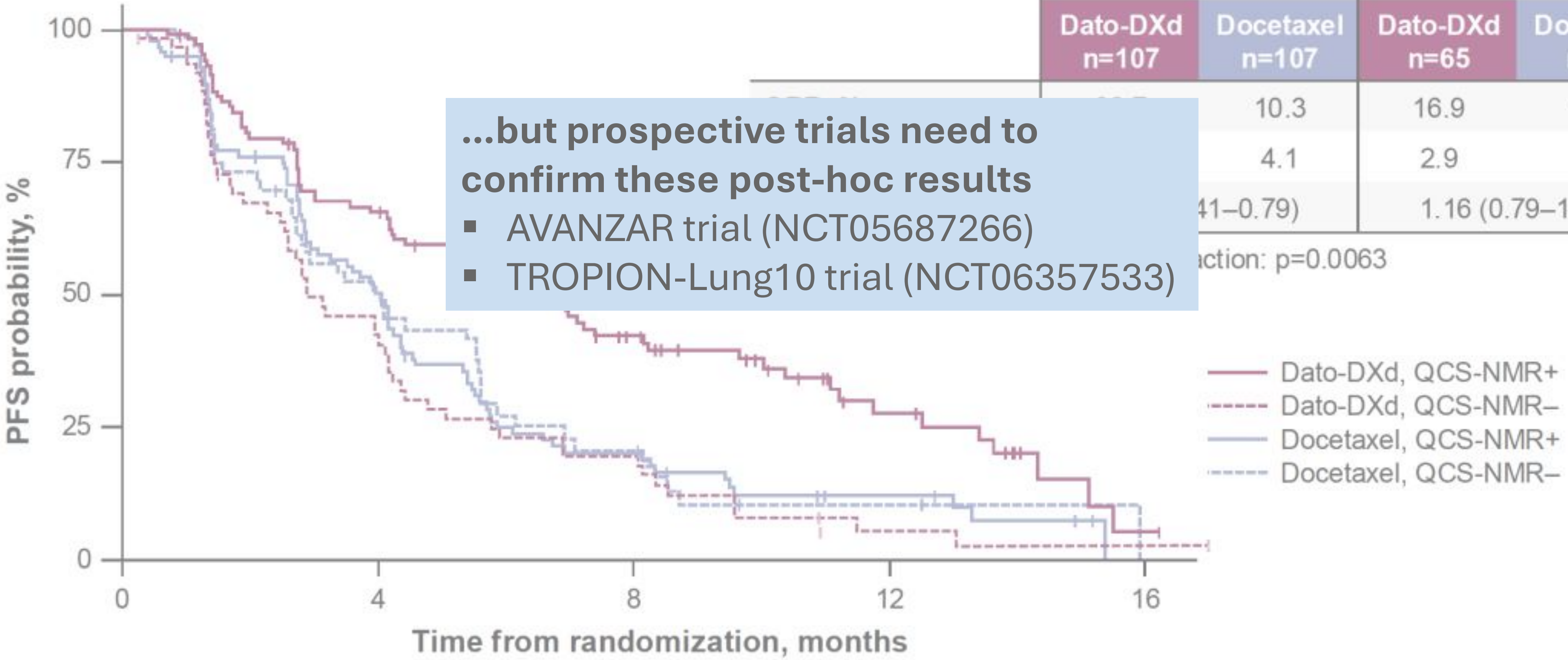
Basic: Biomarker quantification

Results from the TROPION-Lung01 study (NCT04656652)

PFS and ORR by BICR in patients stratified by TROP2 NMR (N=352)

	TROP2 QCS-NMR+		TROP2 QCS-NMR–	
	Dato-DXd n=107	Docetaxel n=107	Dato-DXd n=65	Docetaxel n=73
[illegible]	[illegible]	10.3	16.9	15.1
[illegible]	[illegible]	4.1	2.9	4.0
[illegible]	[illegible]41–0.79)		1.16 (0.79–1.70)	

[illegible]action: p=0.0063



...but prospective trials need to confirm these post-hoc results

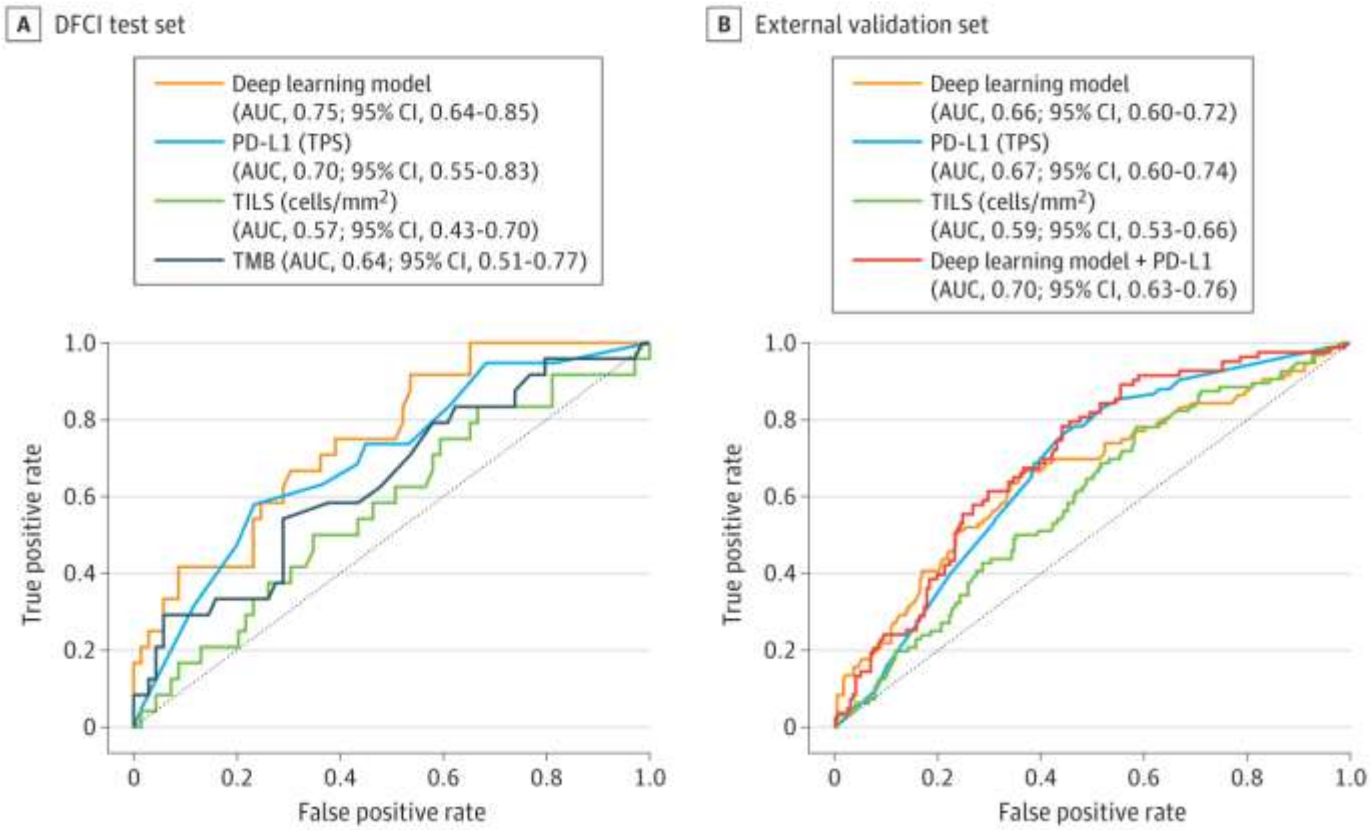
- AVANZAR trial (NCT05687266)
- TROPION-Lung10 trial (NCT06357533)

Garassino et al. WCLC 2024 – PL02.11

Advanced: Prediction of treatment response

A multi-institutional model from the Dana-Faber Cancer Institute

Deep-IO: A supervised (explainable) stratification model to predict response (responders vs non-responders) to immune checkpoint inhibitors (ICIs) monotherapy in NSCLC



Performance compared to other predicting markers

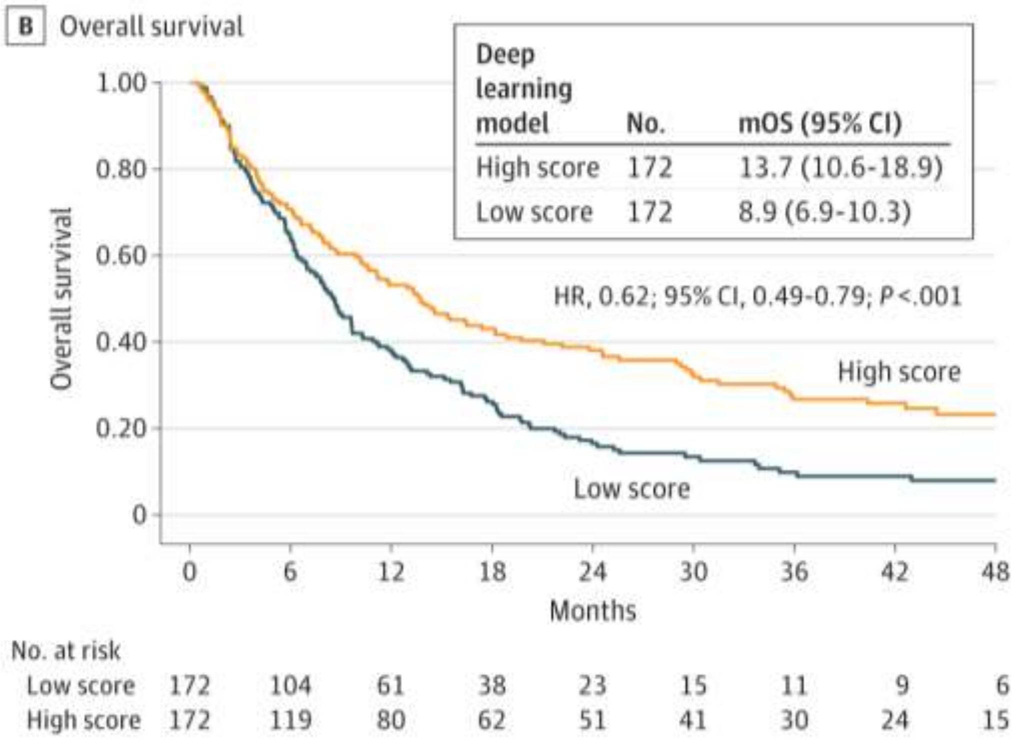
- Diverging results between test and validation cohort (robustness and applicability?)
- Overall performance <0.8 (ll ideal threshold) = 0.66 – 0.75

Rakee et al. JAMA Oncol 2024; doi:10.1001/jamaoncol.2024.5356

Advanced: Prediction of treatment response

A multi-institutional model by the Dana-Faber Cancer Institute

Deep-IO: A supervised (explainable) stratification model to predict response to immune checkpoint inhibitors (ICIs) monotherapy in NSCLC



OS of ICI monotherapy in NCLC stratified by Deep-IO score
Cut-off: median score

Ready for decision making?

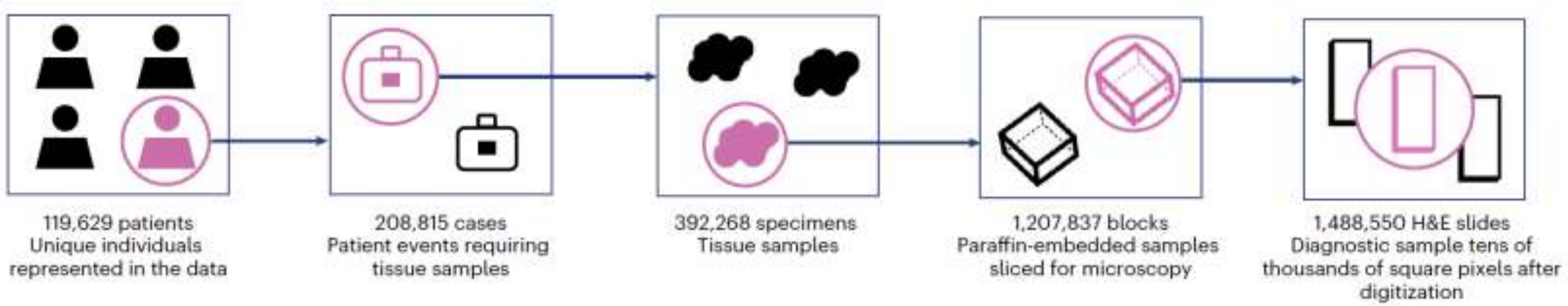
- No!
- Interesting results, which may compete with PD-L1 IHC!
- But, prospective randomized trials are necessary!

Rakee et al. JAMA Oncol 2024; doi:10.1001/jamaoncol.2024.5356

Current developments: Foundation models

One fits all models?

Virchow – A pan-cancer, multi-purpose foundation model

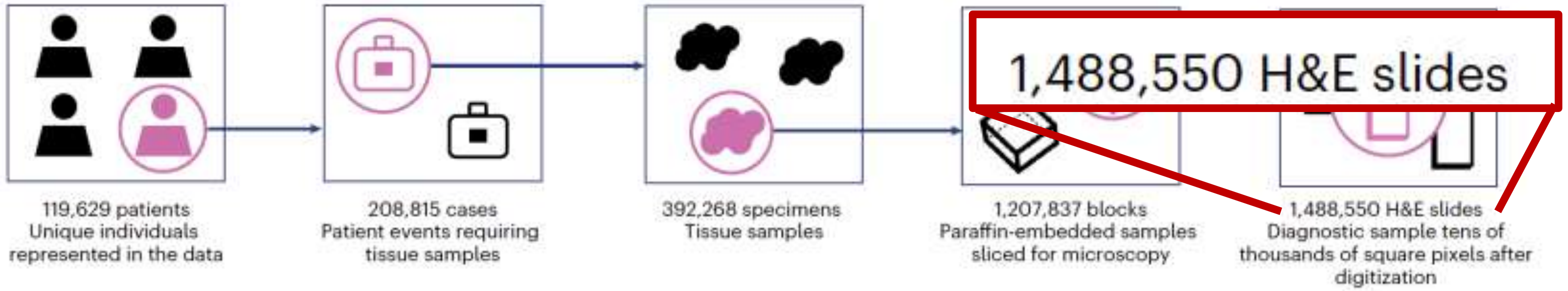


Vorontsov et al, Nat Med 2024; doi.org/10.1038/s41591-024-03141-0

Current developments: Foundation models

One fits all models?

Virchow – A pan-cancer, multi-purpose foundation model



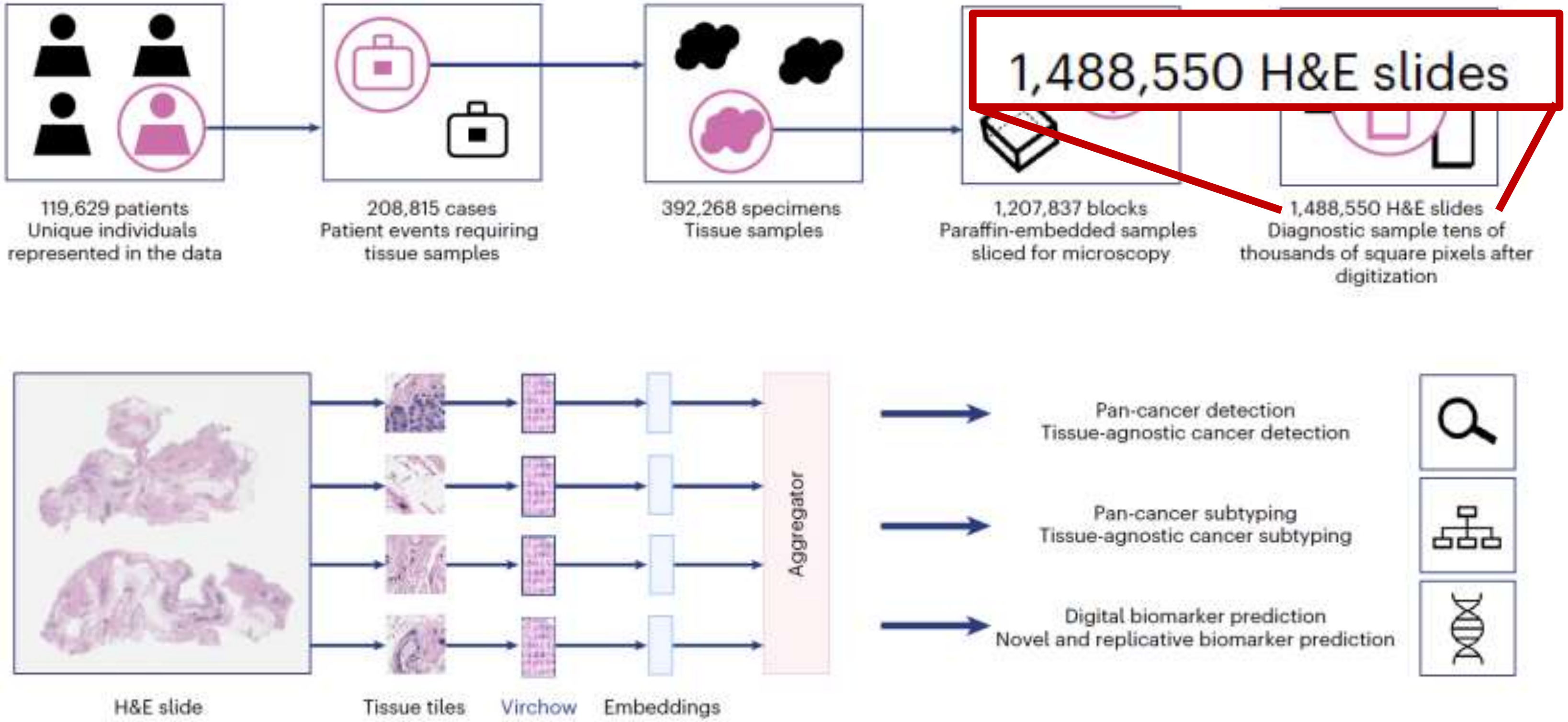
Learning model
Self-supervised
Unlabelled data

Vorontsov et al, Nat Med 2024; doi.org/10.1038/s41591-024-03141-0

Current developments: Foundation models

One fits all models?

Virchow – A pan-cancer, multi-purpose foundation model



Learning model
Self-supervised (SSL)
Unlabelled data

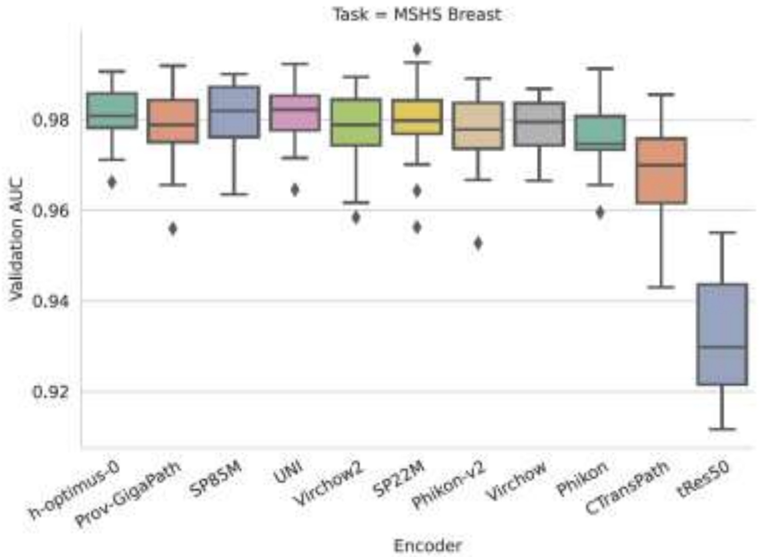
Tasks
Multi-purpose
Multi-task
Multi-modal

Vorontsov et al, Nat Med 2024; doi.org/10.1038/s41591-024-03141-0

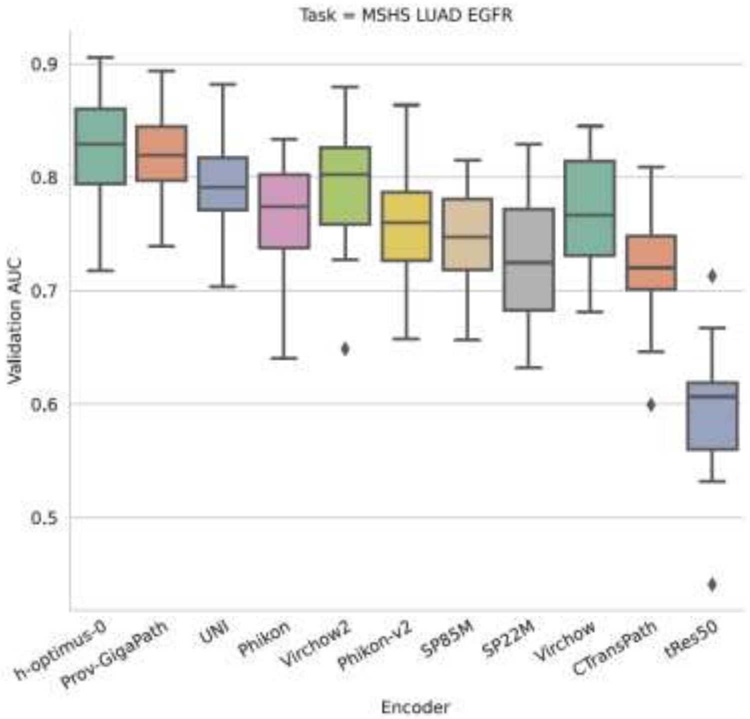
Current developments: Foundation models

One fits all models?

Challenges – systematic analysis of performance



Example for detection
Breast cancer detection performance



Example for prediction
EGFR mutation prediction performance

Key findings

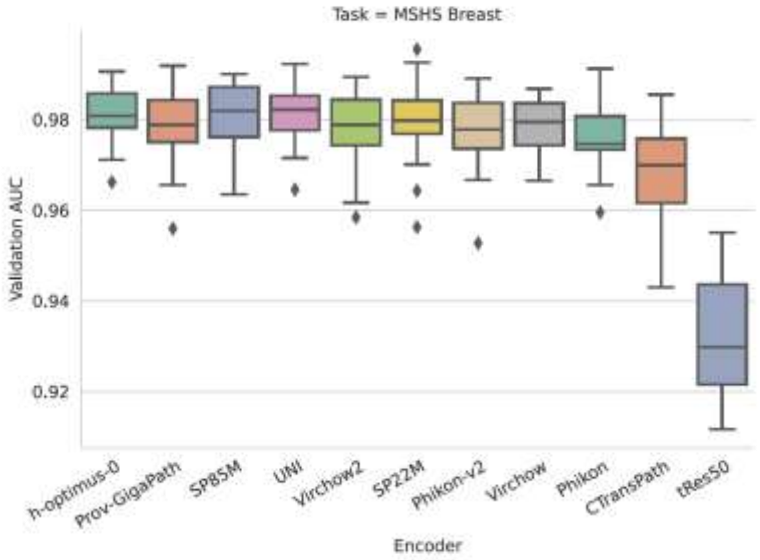
- Overall good performance in cancer detection
- Heterogenous performance in biomarker prediction
- No strong correlation between size of pre-training data set and performance – improvement with more samples?

Campanella et al, Nat Commun 2025; doi.org/10.1038/s41467-025-58796-1

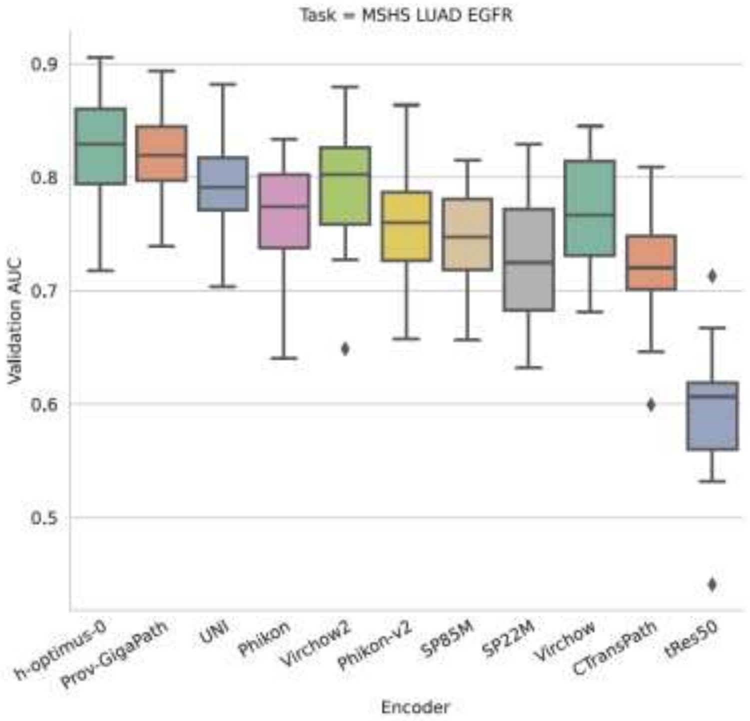
Current developments: Foundation models

One fits all models?

Challenges – systematic analysis of performance



Example for detection
Breast cancer detection performance



Example for prediction
EGFR mutation prediction performance

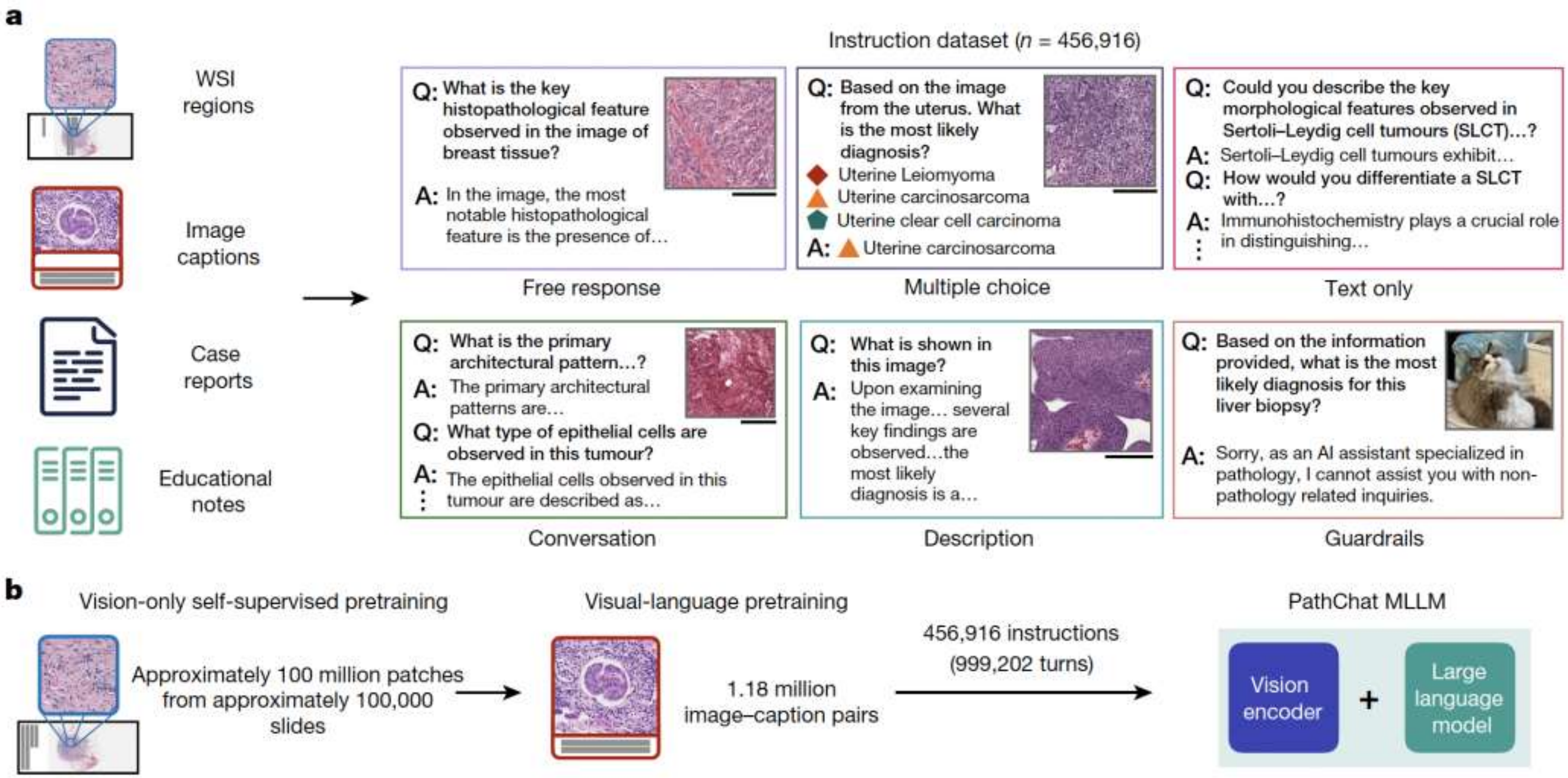
Would I base a clinical decision on the AI-based mutation prediction?

- **No – not yet**
- **Refinement need**
- **Including other modules like clinical characteristics and radiomics**

Campanella et al, Nat Commun 2025; doi.org/10.1038/s41467-025-58796-1

Outlooks: AI co-pilots

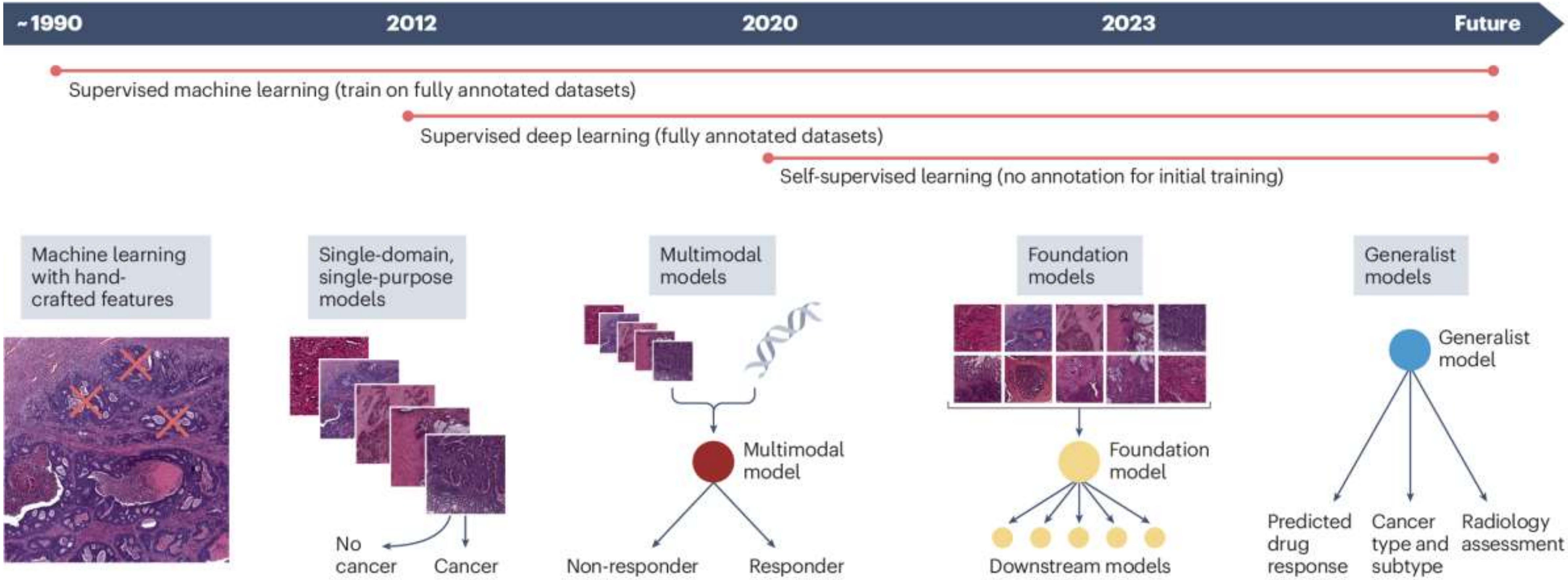
Integrating foundation models and specialized LLMs = Multimodal LLM



Lu et al. Nature 2024; doi.org/10.1038/s41586-024-07618-3

AI development – summary

From single purpose models to foundation models and beyond



Perez-Lopez et al., Nat Rev Cancer 2024 - doi.org/10.1038/s41568-024-00694-7

Challenges and hurdles

Why is AI in digital pathology still not routine?

- Task-dependent performance and reliability
  - Currently no approved one-fits-all foundation model
- Generalizability, repoducability and external validity
  - Clinical evidence gap
  - Lack of prospective trials
  - Lack of clinical meaningful robustness of prediction models
  - Biopsies still represent major challenges due to low tissue quality
- Reimbursement
  - Uncertain and different national reimbursement strategies
- Complex regulatory processes (IVDR)
  - However around 50 CE-IVDR tools are available in the EU, only 2 in the US[1]
- Workflow integration and infrastructure
  - High implementation costs and uncertain reimbursement
  - Uneven global implementation

1. Digital Pathology AI Companies – List of Commercial AIs; Fraunhofer Institute; https://websites.fraunhofer.de/smart-sensing-insights/digital-pathology-ais/?utm_source=chatgpt.com

16th CONGRESS Lung ON CANCER

BARCELONA 27 / 28 NOVEMBER 2025

THANK YOU

GECP lung cancer research

GECP
lung cancer
research