



AI-ON-Lab

# Intelligenza Artificiale e NSCLC

Arsela Prelaj, MD, PhD

AI-ON-Lab leader (Artificial Intelligence for Oncology lab), and

Medical Oncologist, at Istituto Nazionale Tumori di Milano

PhD in Bioengineering and AI, Politecnico di Milano, Italy

ESMO Real-World Data and Digital Health working group member



Responsabile Scientifico  
STEFANIA GORI

**AIGOM**  
ASSOCIAZIONE ITALIANA  
GRUPPI ONCOLOGICI MULTIDISCIPLINARI

Congresso Nazionale

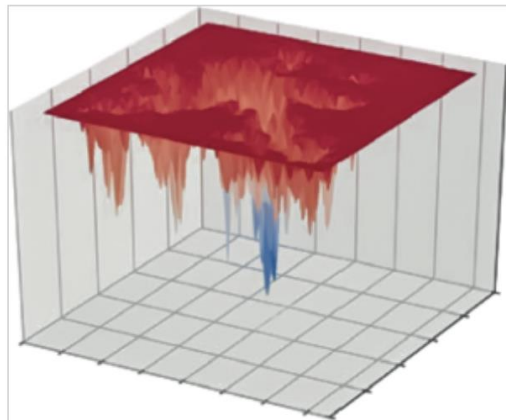
# CARCINOMA DEL POLMONE: QUALI NOVITÀ NEL 2024?

V EDIZIONE

28 OTTOBRE 2024

# Physics Nobel scooped by machine-learning pioneers

John Hopfield and Geoffrey Hinton pioneered computational methods that enabled the development of neural networks.



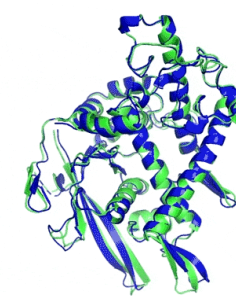
## Machine memory

In 1982, Hopfield, a theoretical biologist with a background in physics, came up with a network that described connections between virtual neurons as physical forces<sup>1</sup>..

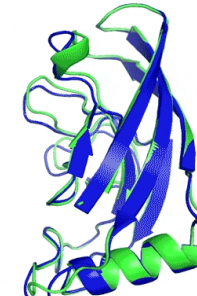
Optimizing the synergy between physics and machine learning. *Nat Mach Intell* **3**, 925 (2021). <https://doi.org/10.1038/s42256-021-00416-w>; Jumper, J. et al. *Nature* **596**, 583–589 (2021).

# Chemistry Nobel goes to developers of AlphaFold AI that predicts protein structures

This year's prize celebrates computational tools that have transformed biology and have the potential to revolutionize drug discovery.



T1037 / 6vr4  
90.7 GDT  
(RNA polymerase domain)



T1049 / 6y4f  
93.3 GDT  
(adhesion tip)

- Experimental result
- Computational prediction

**“I hope when we look back on AlphaFold, it will be the first proof point of AI’s incredible potential.”**

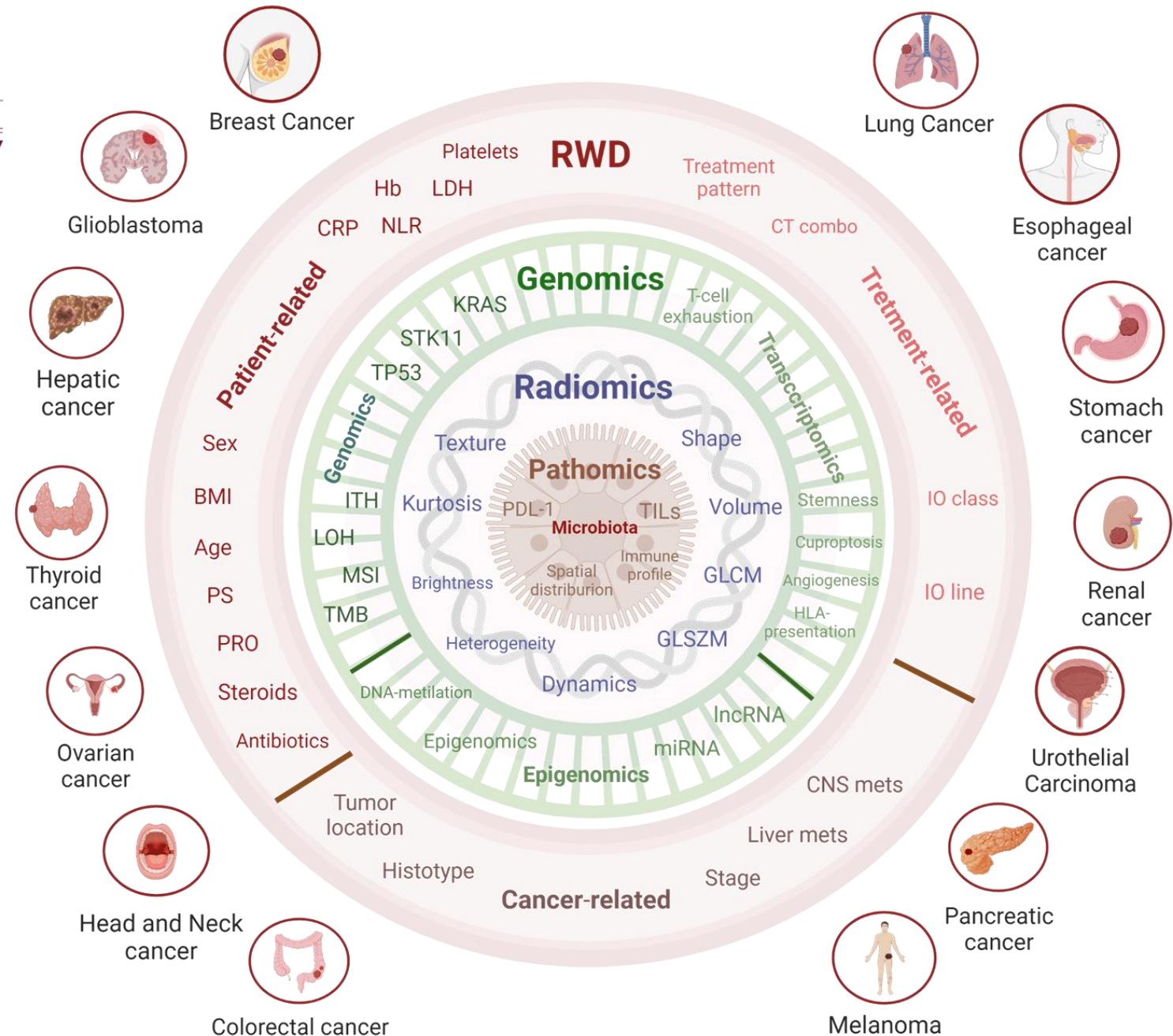
# What AI Can Do?



## REVIEW

Artificial intelligence for predictive biomarker discovery in immunoncology: a systematic review

1. use different data modalities, especially images
2. discover new patterns and biomarkers
3. drive new wet experiments





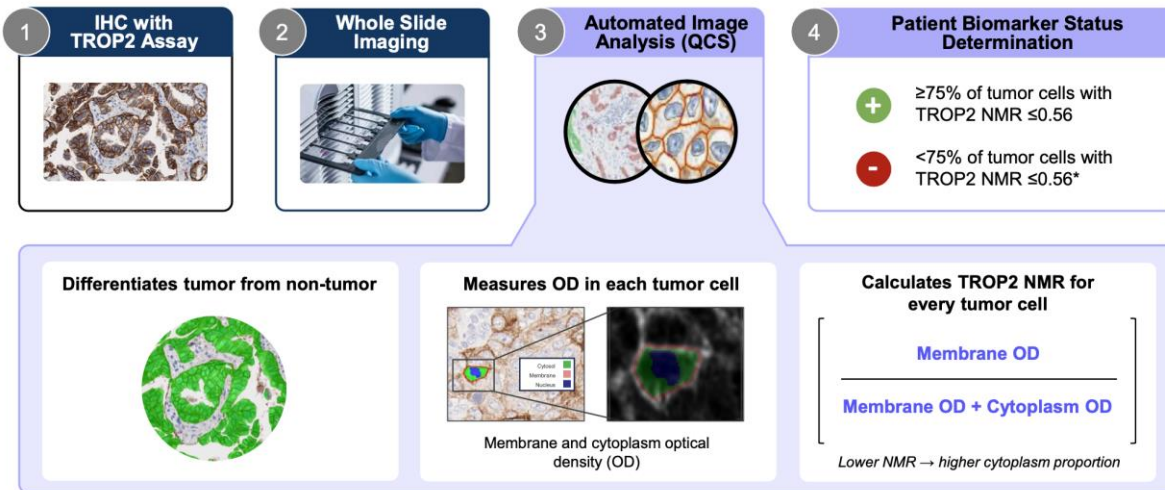
# IMMUNE BIOMARKER DISCOVERY WITH AI: Example

## Normalized Membrane Ratio of TROP2 by Quantitative Continuous Scoring is Predictive of Clinical Outcomes in TROPION-Lung01

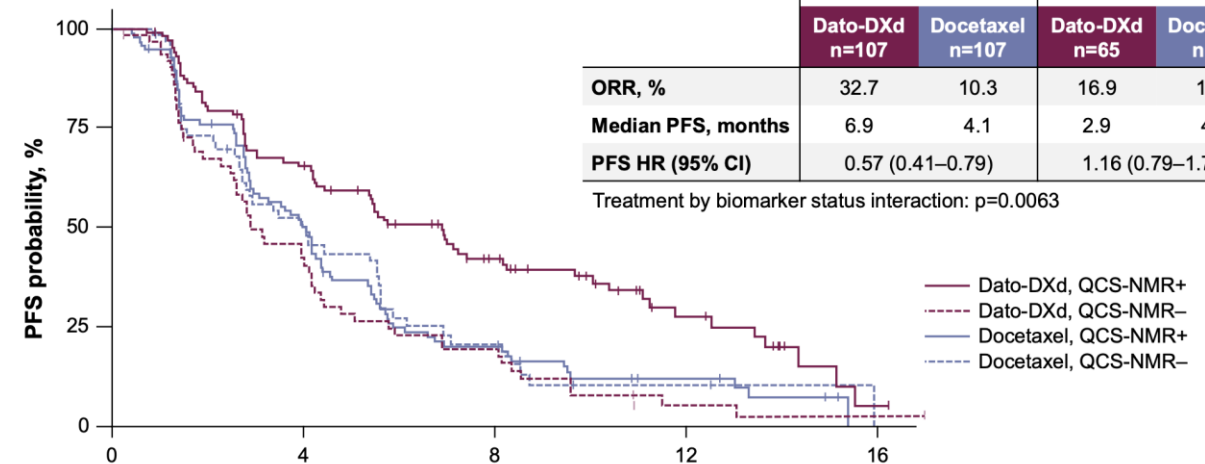
Marina Chiara Garassino,<sup>1</sup> Jacob Sands,<sup>2</sup> Luis Paz-Ares,<sup>3</sup> Aaron Lisberg,<sup>4</sup> Melissa Johnson,<sup>5</sup> Maurice Pérol,<sup>6</sup> Danielle Carroll,<sup>7</sup> Ansh Kapil,<sup>8</sup> Vincent Haddad,<sup>7</sup> Deise Uema,<sup>9</sup> Hadassah Sade,<sup>8</sup> Myung-Ju Ahn,<sup>10</sup>

Patients receiving Dato-DXd who were TROP2 QCS-NMR+ had a higher ORR and longer PFS compared with those who were TROP2 QCS-NMR-

QCS is a novel, fully-supervised computational pathology approach that precisely quantifies and locates targets like TROP2. TROP2 QCS-NMR positivity is predictive for longer PFS with Dato-DXd in the biomarker-evaluable population



Biomarker-evaluable population, n=352



**TROP2 QCS-NMR has the potential to be the first TROP2 biomarker and the first computational pathology biomarker for predicting clinical response to Dato-DXd in NSCLC**

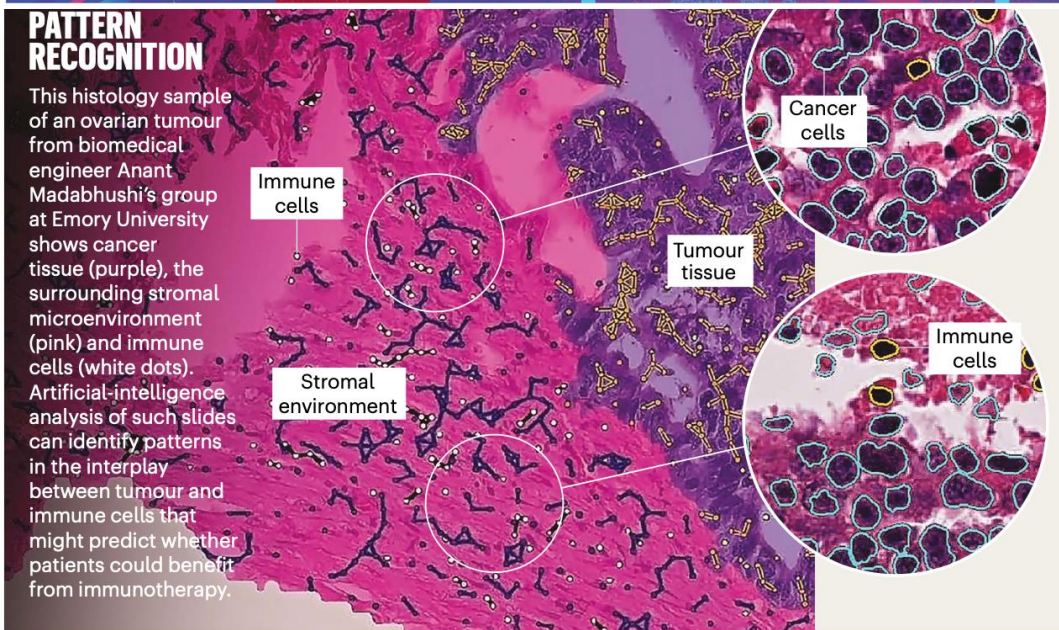
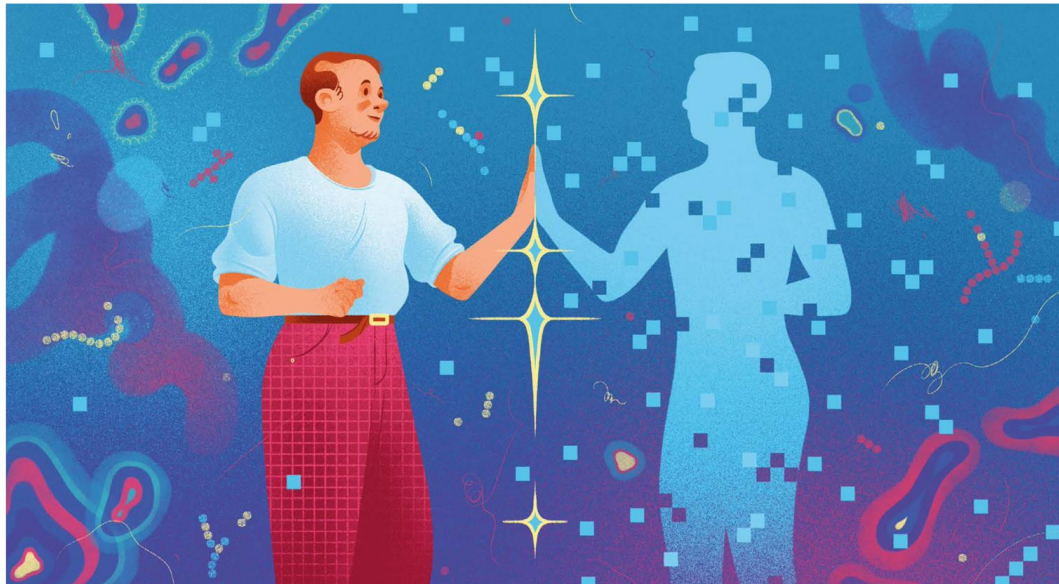


# AI-based **DECISION SUPPORT SYSTEM**

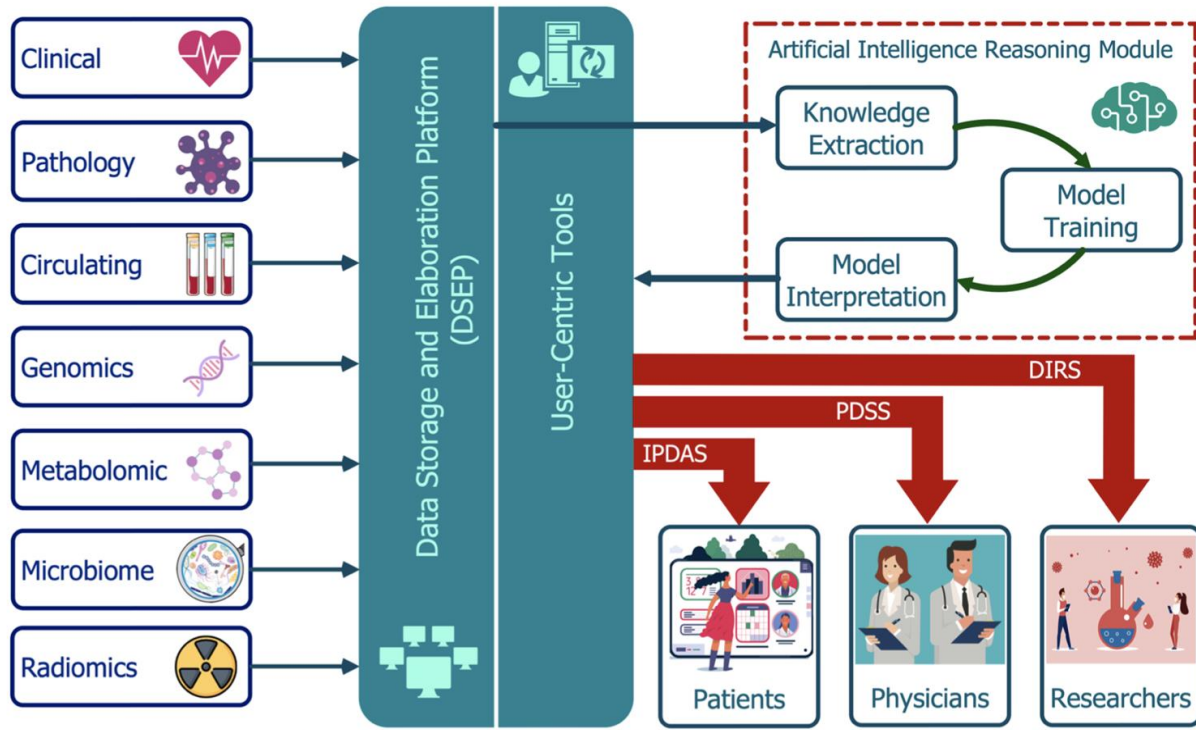
FROM RESEARCH TO CLINIC

## Clinical Decision Making With Ai For Treatment Selection

Prelaj's group has also evaluated AI-guided prediction of immunotherapy response in people with NSCLC, and has generally been impressed with its ability to deliver useful predictions. "These tools are trustworthy, and are working," she says. In 2022, her group spearheaded the I<sup>3</sup>LUNG Project, a five-year initiative that has recruited 2,200 people with NSCLC in Europe, the United States and Israel. I<sup>3</sup>LUNG aims to develop a deep-learning model for predicting the response to checkpoint inhibitors – either alone, or in combination with other therapies – on the basis of imaging, histology and data from clinical records. The researchers will then validate the model's ability to identify effective treatment strategies in a prospective cohort of people with cancer<sup>3</sup>.



# THE I3LUNG PROJECT - OVERVIEW



10M Euros Grant

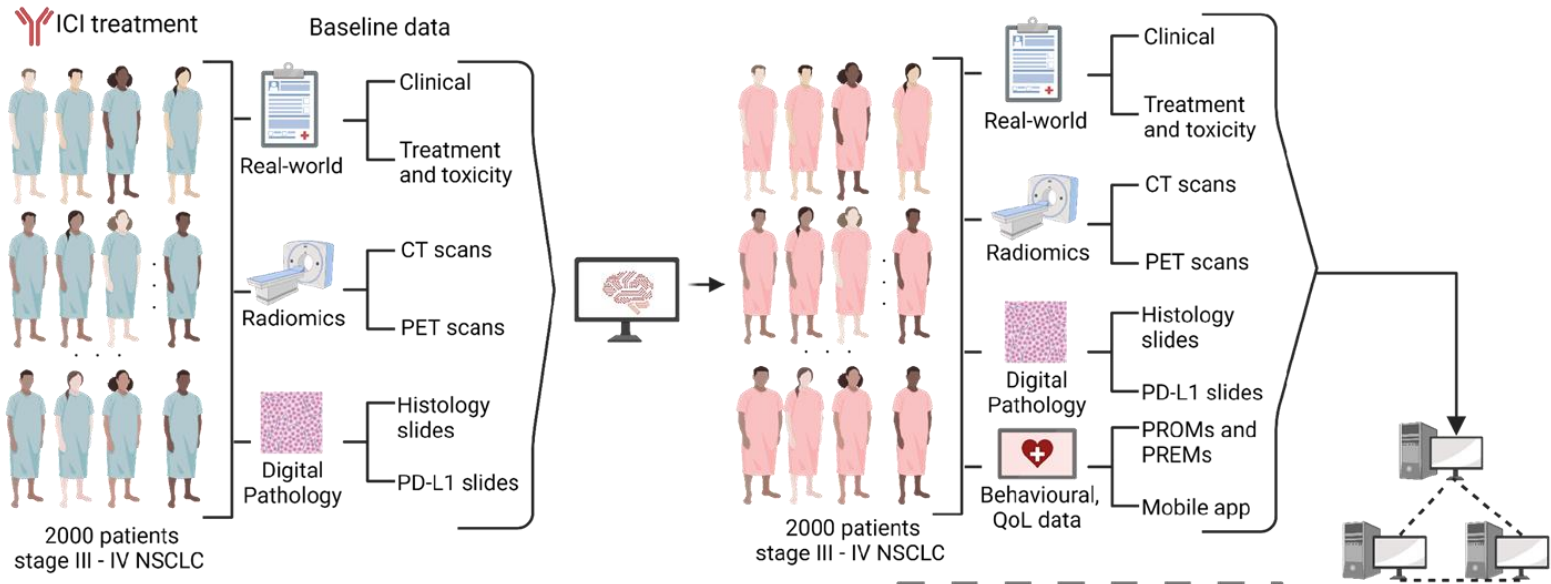


**Goal:** Develop a Data storage and Elaboration Platform (DESP) by integrating Real world and multiomics data in NSCLC patients treated with immunotherapy with the aim to produce a clinical decision making tool



# Multimodal and Multiomics AI

**RETROSPECTIVE** **PROSPECTIVE**

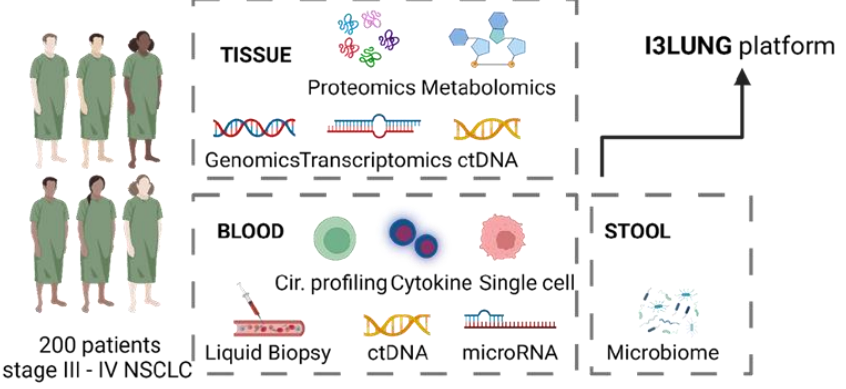


**PROSPECTIVE I**

2000 RW prospective patient's cohort

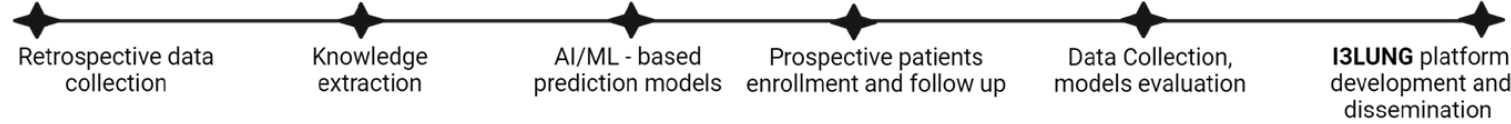
**RETROSPECTIVE**

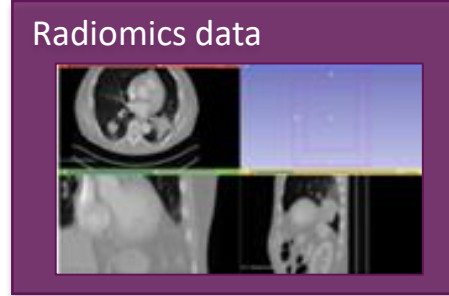
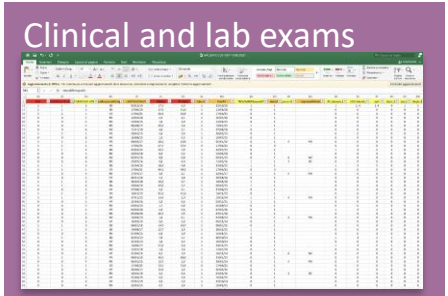
2000 RW retrospective patient's cohort



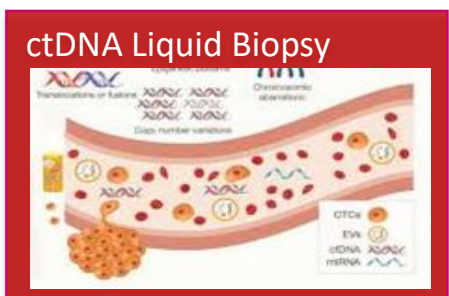
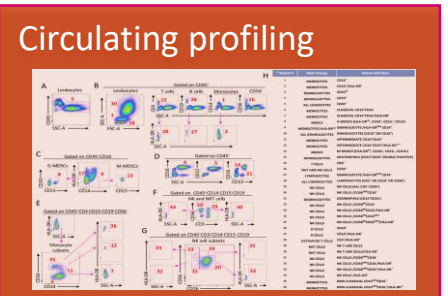
**PROSPECTIVE II**

200 OMICs prospective patient's cohort

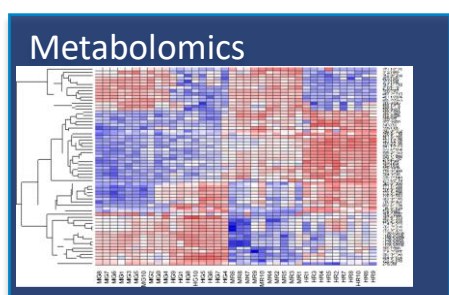
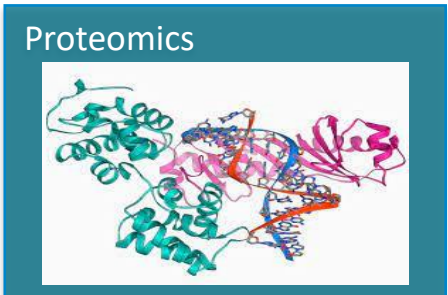




**NGRWD:**  
RWD  
IMAGES  
QoL

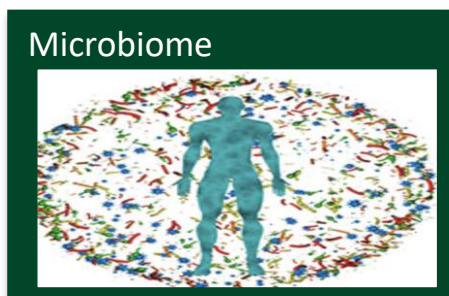
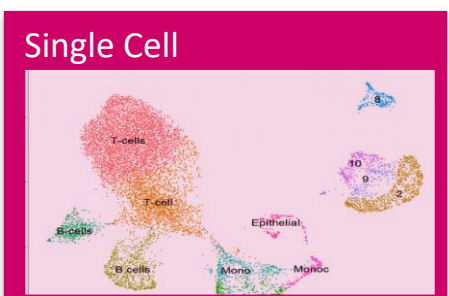


**BLOOD**  
level



**TISSUE**  
level

**SINGLE  
CELL**  
Blood level



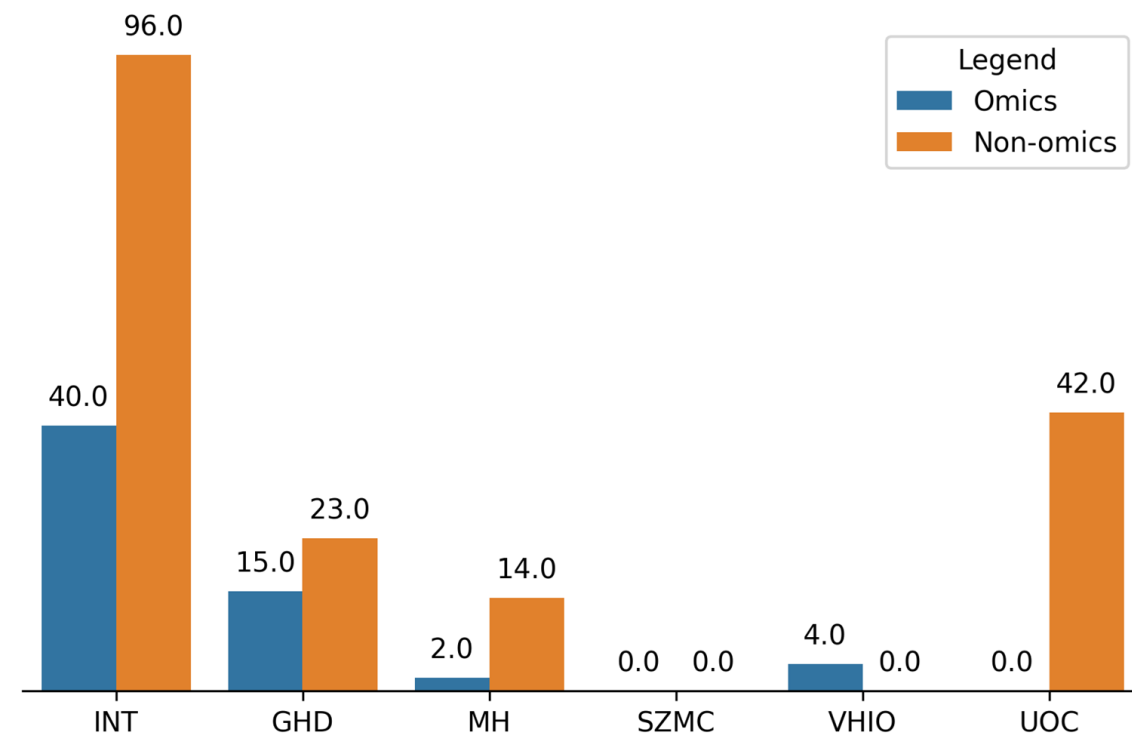
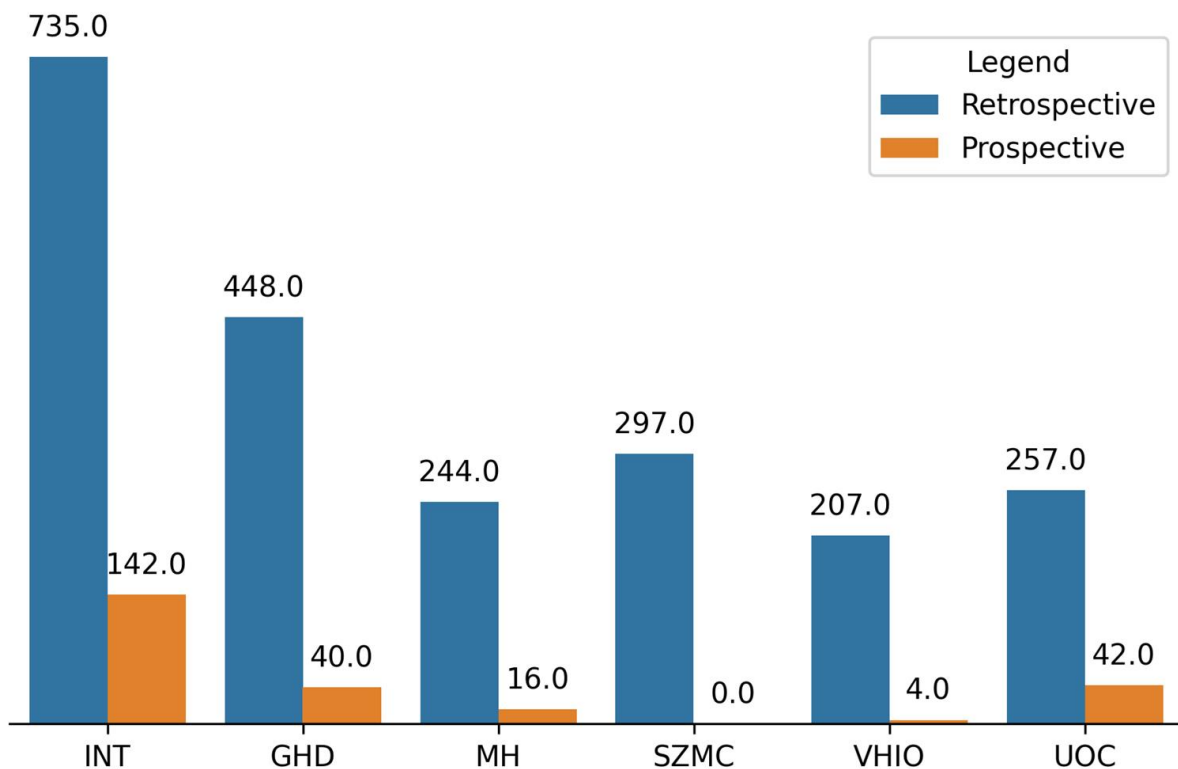
**STOOL**  
level



# I3LUNG: PATIENTS ENROLLMENT, MAY 2024



Retrospective and Prospective  
N=2188 patients





# I3LUNG: Digital pathology predicts PD-L1 expression in metastatic NSCLC patients treated with immunotherapy

# POSTER



Arsela Prelaj<sup>1,3</sup>, Matteo Sacco<sup>2</sup>, Vanja Miskovic<sup>1,3</sup>, Daniele Lorenzini<sup>1</sup>, Francesco Trovò<sup>3</sup>, Aleksandra Zec<sup>3</sup>, Laila Roisman<sup>4</sup>, Leonardo Provenzano<sup>1</sup>, Claudia Proto<sup>1</sup>, Andrea Spagnoletti<sup>1</sup>, Cecilia Silvestri<sup>1</sup>, Alba Meya<sup>6</sup>, Alessandra Pedrocchi<sup>3</sup>, Evangelos Sarris<sup>7</sup>, Enriqueta Felip<sup>8</sup>, Filippo G.M. De Braud<sup>1</sup>, Martin Reck<sup>9</sup>, Giuseppe Lo Russo<sup>2</sup>, Alexander T. Pearson<sup>2</sup>, Marina Chiara Garassino<sup>2</sup>

1. Fondazione IRCCS – Istituto Nazionale dei Tumori, Milan, Italy; 2. University of Chicago Department of Medicine - Section of Hematology/Oncology, Chicago, United States of America; 3. Politecnico di Milano, Milan, Italy; 4. Shaare Zedek Medical Center, Jerusalem, Israel; 5. UNIMI - Università degli Studi di Milano Statale, Milan, Italy; 6. MEDSIR - Medica Scientia Innovation Research, Barcelona, Spain; 7. Metropolitan Hospital, Athens, Greece; 8. Vall d'Hebron University Hospital, Barcelona, Spain; 9. Airway Research Center North, German Center for Lung Research, LungenClinic, Grosshansdorf, Germany



## I3LUNG

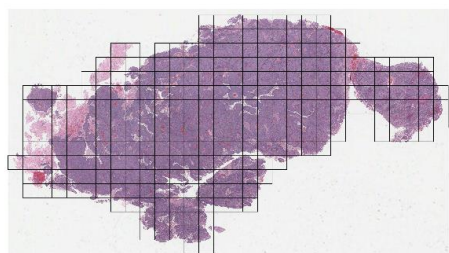


I3LUNG is a project funded by the European Union through the Horizon 2020 program that aims to develop Artificial Intelligence (AI)-based tools to predict the response of advanced non-small cell lung cancer (NSCLC) patients to immune checkpoint inhibitors (ICIs). The project brings together a consortium of 16 partners from 10 countries (Belgium, Denmark, Italy, Germany, Greece, Spain, Sweden, Switzerland, the United States, and Israel). Utilizing patient data, which includes digital pathology slides (DPS), genomics, radiomics, along with other patient characteristics, the overall goal is to develop a platform to guide therapeutic decisions in immuno-oncology for both healthcare professionals and patients.

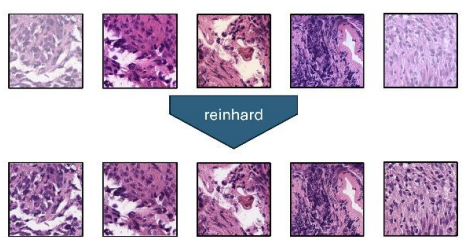
## INTRODUCTION

- Immunotherapy (IO) is the new standard of care for patients with advanced NSCLC, yet only 30-50% of patients benefit from it long-term.
- A better understanding of tumor features could help guide treatment decisions.
- To date, Programmed Death-Ligand 1 (PD-L1) remains the only biomarker used to predict IO efficacy, demonstrating its unique predictive ability, even if not perfect.
- A specific morphology has been found to be associated with PD-L1 expression, introducing new scenarios for the biological interpretation of the immune response.
- Utilizing AI and machine learning processes to analyze DPS could help create decision making tools for more individualized prediction of response.

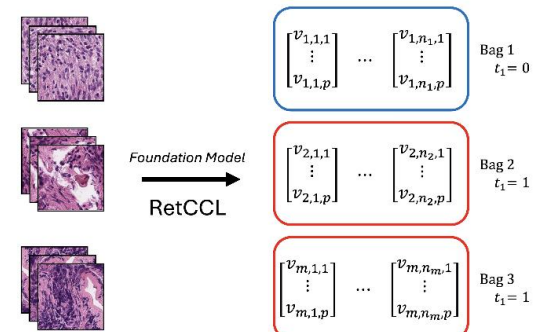
## METHODS



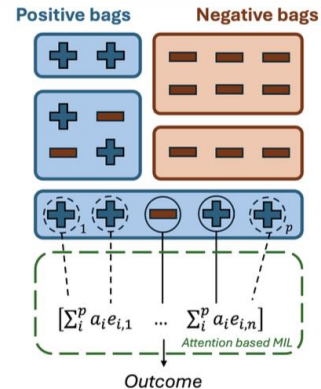
Digital Pathology whole slide images are very large images, and to feed them into Artificial Neural Network-based models we extracted square tiles [299x299px] at 10x magnification from the whole slide images.



Reinhard normalization was employed to reduce batch effect.



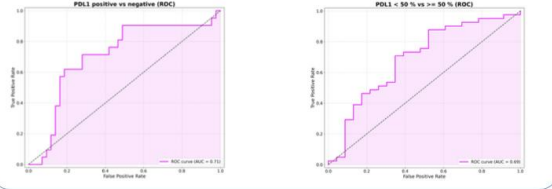
Large-scale AI models, known as foundation models, are developed using vast datasets and a training approach called self-supervised learning. This method does not rely on manually labeled data. Instead, it presents the model with a complex task inherent to the data itself. By solving this task, the model learns to identify and extract important features from the input information on its own. For this task, we used RetCCL, a model trained on TCGA and PAIP datasets. Once processed through RetCCL, tiles are converted into vectors of biologically relevant features.



Slides are now converted into bags of vectors. Assuming that not all tiles are equally relevant for our task, we need a model able to learn not only the patterns associated with a certain outcome, but also which tiles to focus on to find those patterns. For this goal, we used an Attention-Based Multiple Instance Learning model, which employs the attention mechanism to infer the importance of each tile vector.

## RESULTS

- Among the 2188 pt enrolled in the I3LUNG retrospective cohort, 474 patients had available DPS and PD-L1 status to be considered for the present analysis.
- PD-L1 expression was high (>50%), low (1-49%) and negative in 145 (37%), 129 (32%) and 127 (31%) patients within the training cohort, respectively, and 24 (33%), 23 (32%) and 26 (35%) among the validation cohort, respectively.
- PD-L1 high vs low/negative status through DPS were able to be predicted with an area under the curve (AUC) of 0.69; while for PD-L1 positive vs negative an AUC of 0.71 was achieved.



## CONCLUSIONS

- To our knowledge, this is the largest series to date demonstrating a correlation between morphological features and PD-L1 expression in lung cancer.
- Data suggests that PD-L1 high and negative have different morphological phenotype.
- This rapid and generalizable model underscores the potential for morphological features to serve as valuable biomarkers in elucidating the mechanisms of immune responses.
- Within I3LUNG integration of genomic and radiomic data will probably allow to improve the ability to assess patient prognosis at diagnosis.

## ACKNOWLEDGEMENTS

We sincerely thank the patients and their families for their invaluable participation in this study. This project has received funding from the European Union's Horizon Europe research and innovation programme. We are grateful for the collaborative efforts of our consortium partners. Their expertise has been crucial to the success of the I3LUNG project.

Scan here to view a PDF of this poster. Copies of this poster obtained through QR (Quick Response) code are per personal use only and may not be reproduced without written permission of the authors.

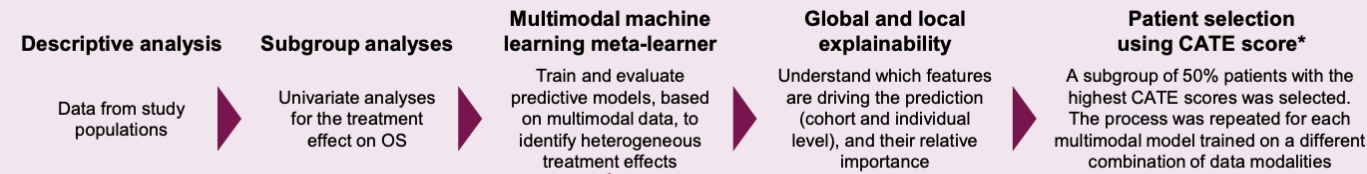


Scan here to view a plain language summary of this study.

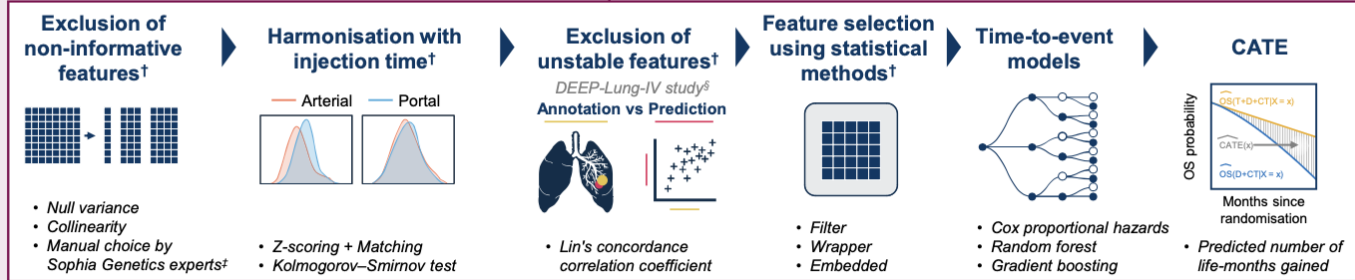


# TRIDENT: Machine learning multimodal signatures to identify patients that would benefit most from tremelimumab addition to durvalumab plus chemotherapy with data from the POSEIDON trial

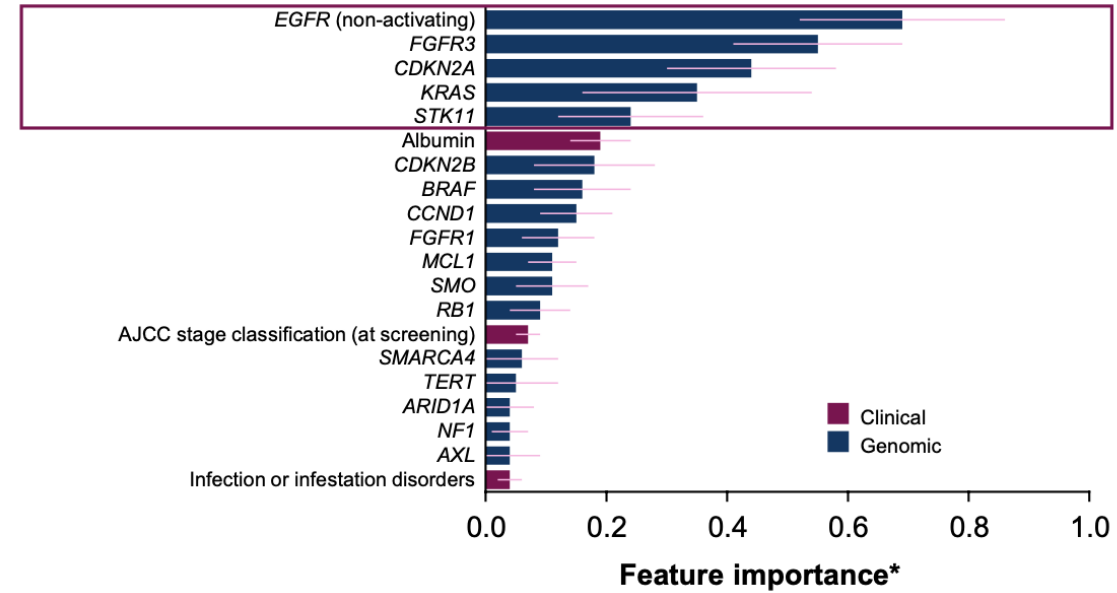
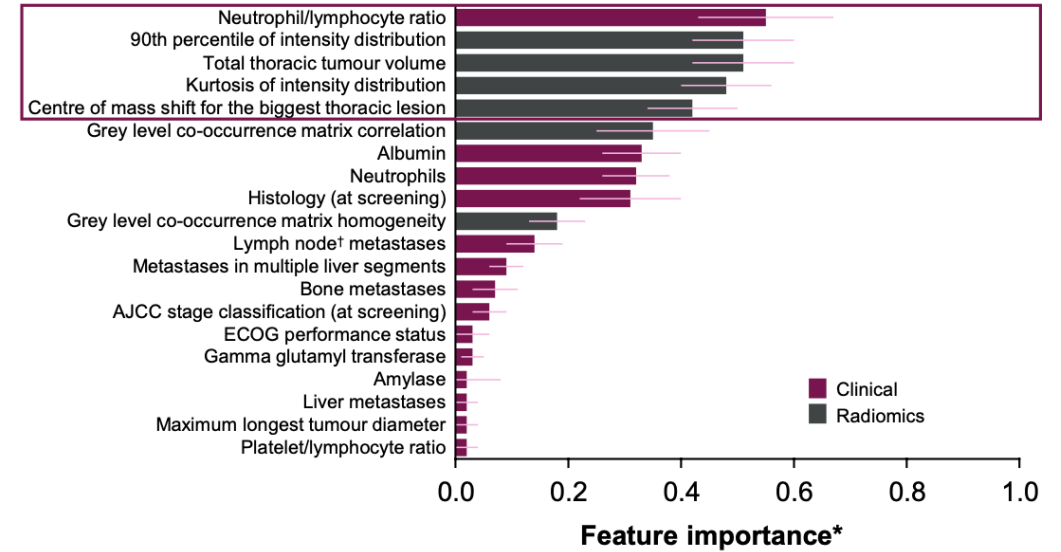
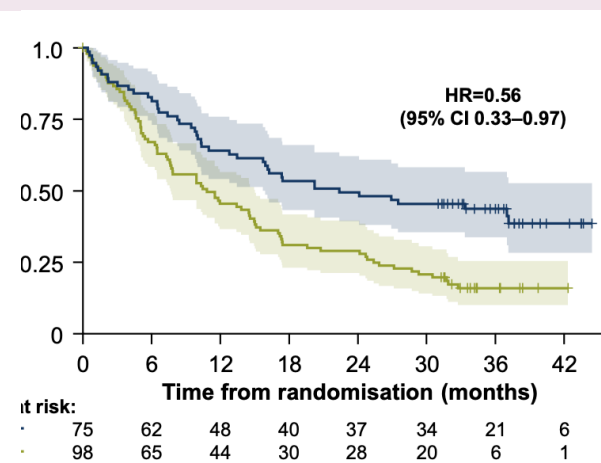
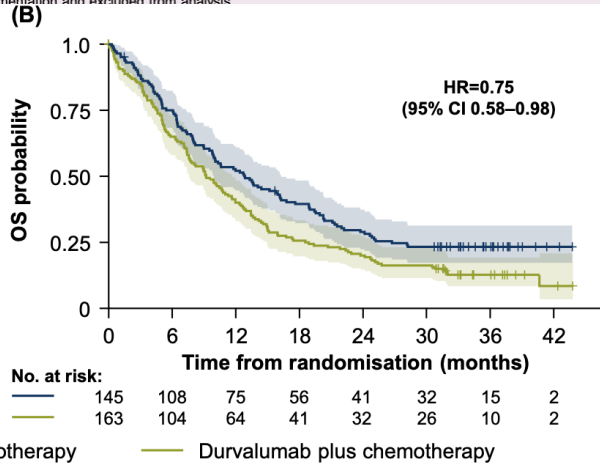
Figure 1. Overview of the analysis.



**Treatment effect prediction model**

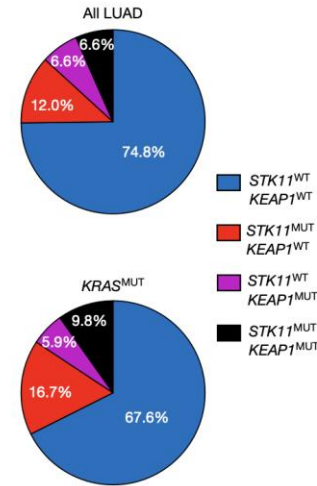
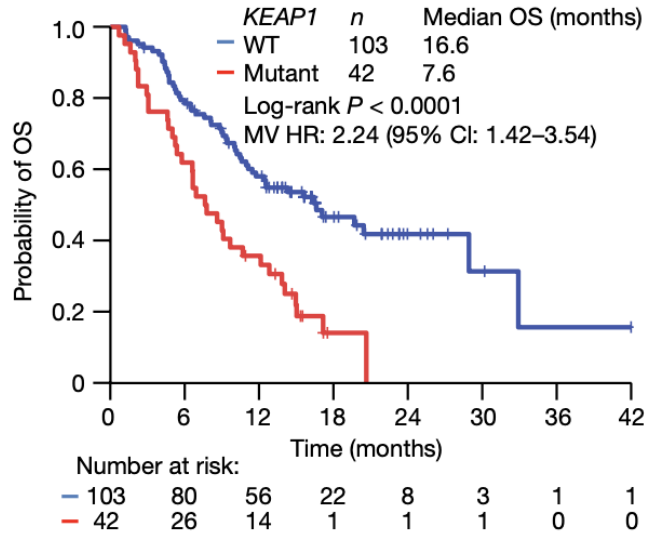


\*CATE score (defined as the difference of RMSTs at 24 months after randomisation) is calculated per-patient. The higher the score for the patient, the higher the expected benefit of tremelimumab plus durvalumab and chemotherapy vs durvalumab plus chemotherapy. †Only relevant for derivation of the radiomics modality. ‡Non-relevant features in the context of metastatic NSCLC tumours and non-interpretable features were removed. §DEEP-Lung-IV (NCT04994795) was used to identify unstable features by comparing the feature value when extracted from annotation to deep learning prediction. Features with low correlation were considered sensitive to segmentation and excluded from analysis.



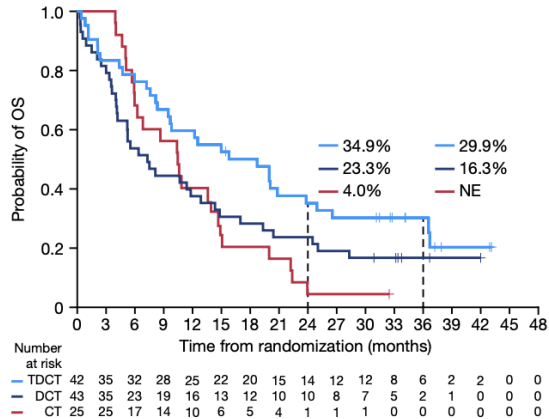
Poster ESMO 2024

## CTLA4 blockade abrogates *KEAP1*/*STK11*-related resistance to PD-(L)1 inhibitors

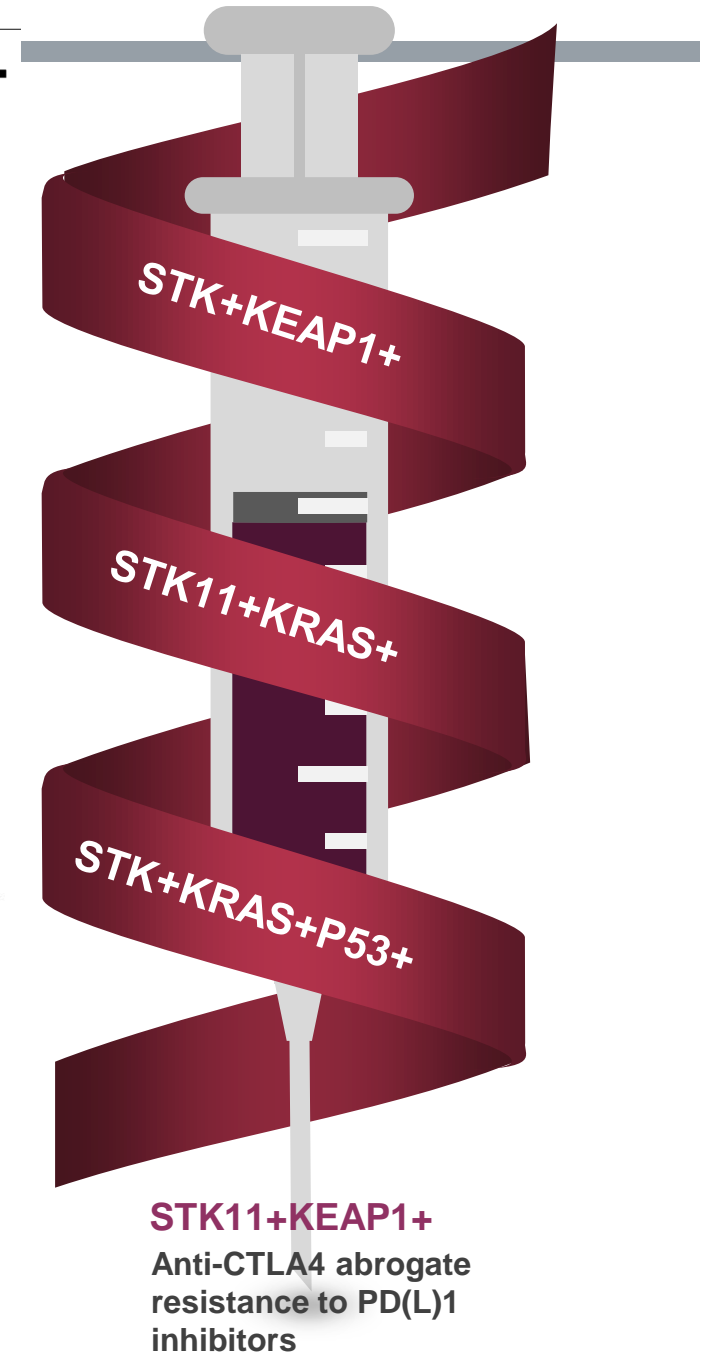
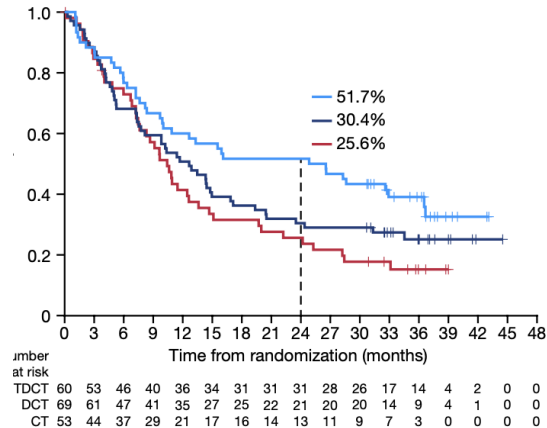


**b**

	<i>STK11</i> <sup>MUT</sup> and/or <i>KEAP1</i> <sup>MUT</sup>		
	TDCT	DCT	CT
mOS (months) (95% CI)	15.8 (9.5–23.8)	7.3 (4.2–12.9)	10.5 (6.0–14.7)
HR versus CT (95% CI)	0.50 (0.29–0.87)	0.90 (0.53–1.52)	–
HR versus DCT (95% CI)	0.64 (0.40–1.04)	–	–



	<i>KRAS</i> <sup>MUT</sup>		
	TDCT	DCT	CT
mOS (months) (95% CI)	25.7 (9.9–36.5)	12.6 (7.5–16.9)	10.4 (7.5–13.6)
HR versus CT (95% CI)	0.56 (0.36–0.88)	0.80 (0.53–1.20)	–
HR versus DCT (95% CI)	0.72 (0.47–1.10)	–	–





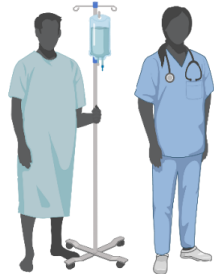
# Papers under submission: PDSS and IPDAS tools



## Patient oriented study

**Goal:** Quality of shared DM and physician-patient communication

Cohort 1: Without IPDAS  
100 patients



Cohort 2: With IPDAS  
100 patients



### Outcome:

Impact of IPDAS on shared medical DM



## Physician oriented study

**Goal:** AI impact on accuracy and decisional process (HI vs HI and AI)

Cohort 1: HI group  
100 patients



Accuracy  
Decision fatigue  
Process

Cohort 1: HI and AI  
100 patients



Accuracy  
Decision fatigue  
Process

### Outcome:

Confirmation whether the combination of HI and AI perform better respect to HI alone (target +15%)

# MOVING FROM PREDICTIVE AI TO GENERATIVE AI

## Predictive based MODELS

- Using task specific – AI
- Using FM to extract meaningful information
- Predict the outcome

## Generative based MODELS

- Using LLMs and FMs to generate decisions
- Create synthetic data and images

**LARGE LANGUAGE MODELS**

**FOUNDATION MODELS**



# Generative AI exists because of the Transformer

Google Brain 2017



---

## Attention Is All You Need

---

**Ashish Vaswani\***  
Google Brain  
avaswani@google.com

**Noam Shazeer\***  
Google Brain  
noam@google.com

**Niki Parmar\***  
Google Research  
nikip@google.com

**Jakob Uszkoreit\***  
Google Research  
usz@google.com

**Llion Jones\***  
Google Research  
llion@google.com

**Aidan N. Gomez\* †**  
University of Toronto  
aidan@cs.toronto.edu

**Łukasz Kaiser\***  
Google Brain  
lukaszkaizer@google.com

**Illia Polosukhin\* ‡**  
illia.polosukhin@gmail.com

### Abstract

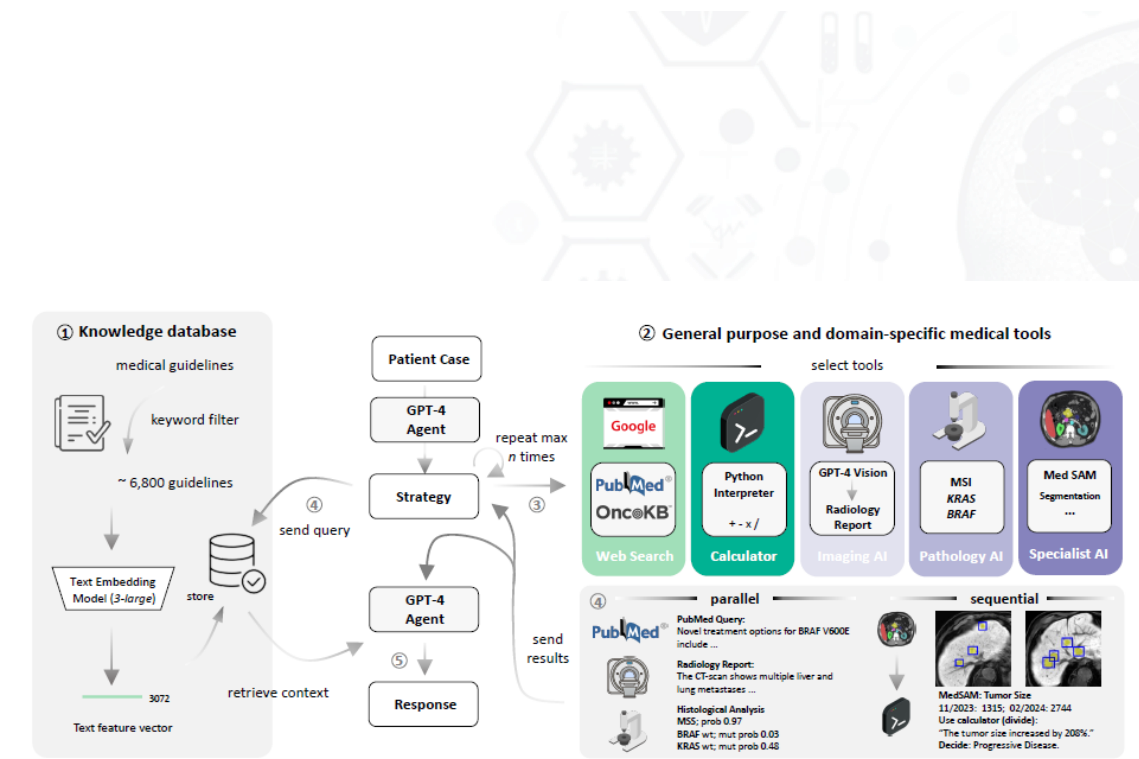
The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.0 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature.

[Submitted on 6 Apr 2024]

# Autonomous Artificial Intelligence Agents for Clinical Decision Making in Oncology

Dyke Ferber, Omar S. M. El Nahhas, Georg Wölflein, Isabella C. Wiest, Jan Clusmann, Marie-Elisabeth Leßman, Sebastian Foersch, Jacqueline Lammert, Maximilian Tschochohei, Dirk Jäger, Manuel Salto-Tellez, Nikolaus Schultz, Daniel Truhn, Jakob Nikolas Kather

# Autonomous AI x CDM



**Clinical Context**

Ms. Xing, a 23-year-old patient, was diagnosed with cholangiocellular carcinoma in early 2023. Following her diagnosis, she underwent a complete surgical resection (R0) and was treated with adjuvant chemotherapy using capecitabine for 11 months. Subsequent imaging showed no evidence of disease recurrence until September of the same year, when MRI imaging ("September2023.png") revealed a new, solitary metastasis. In response, a treatment regimen of gemcitabine and cisplatin was initiated. Considering Ms. Xing's young age, a liver biopsy was performed for comprehensive panel diagnostics. Both imaging results and panel findings are made available for review:

\*\*\*

**Radiology Report (\* -)**  
 Liver: There is one single hypointense lesion in the left lobe of the liver, measuring approximately 0.4 cm in its greatest dimension (Location: [475, 250, 490, 275]). Highly suggestive for recurrence of the known cholangiocellular carcinoma.  
 (\* -)  
 Lymph Nodes: There is no enlargement of the abdominal or pelvic lymph nodes.  
 Peritoneum: No evidence of peritoneal implants or significant fluid accumulation is present.  
 Strong suspicion of a recurrence of the known CCC due to a solitary metastasis in the left liver.

**Molecular Report Summary:**  
 Microsatellite-Instability High, Mutation in BRAF V600E and CD74-R0S1 variant fusion.  
 No alterations in NTRK or FGFR.  
 \*\*\*

Currently, the patient experiences recurrent ascites with notable abdominal tension. Also, a new MRI-scan was performed yesterday ("February2024.png") by our in house radiologists. According to the latest radiology report, the MRI scan performed on the patient with the system ID X-09.22 and documented as "February2024.png" revealed a single metastatic lesion in the left lobe of the liver, delineated by the coordinates 455, 270, 505, 220. The report emphasizes the importance of the findings with the image from a previous examination in September 2023 during the forthcoming tumor board meeting for a detailed analysis. Also, the imaging findings suggest the presence of peritoneal carcinomatosis. Her system ID is X-09.22.

**Question**

Please investigate the current state of the disease. What does the new MRI scan show? Did the metastatic grow? In case of progress, what (targeted) treatment options could we suggest for the molecular tumor board tomorrow? Be aware that the patient is extremely young and we need all options we could find. Also check any options you can find on google or pubmed.

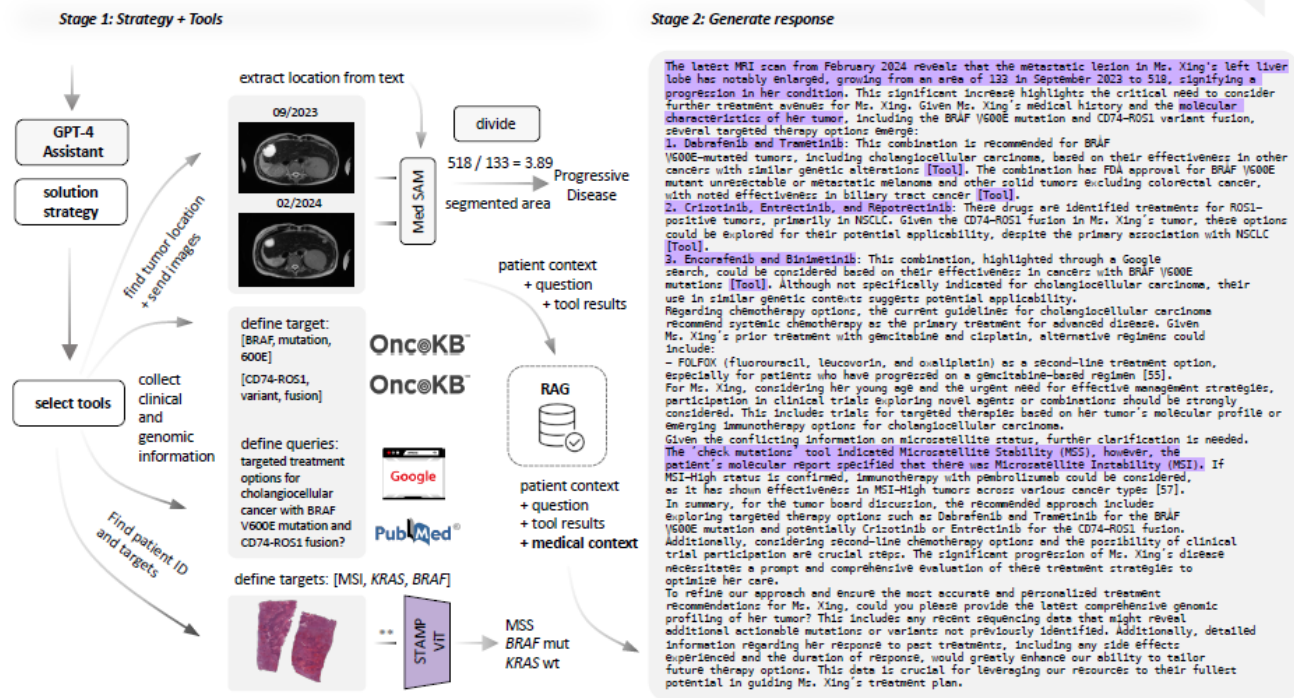
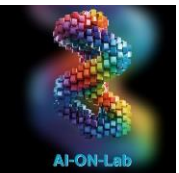


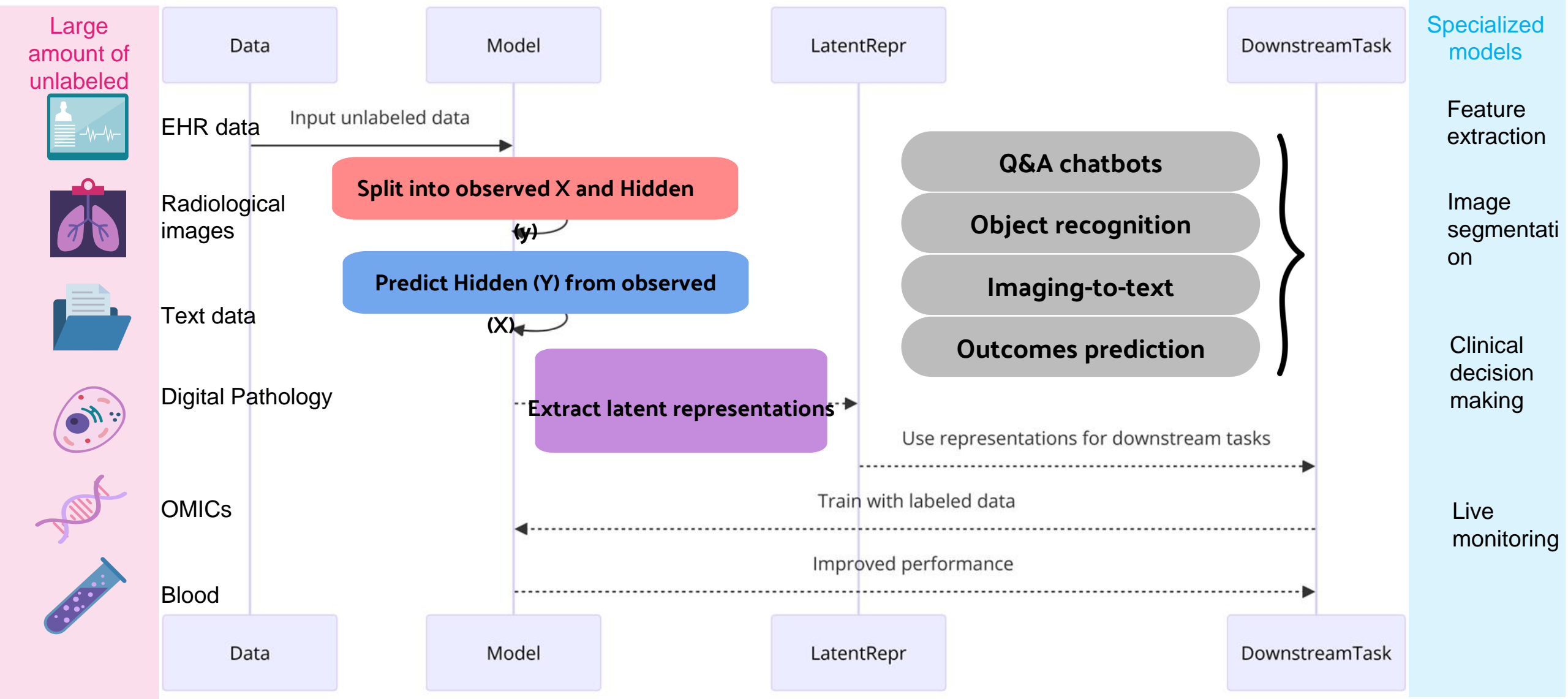
Figure 1: High-level overview of the RAG-Agents framework.



MAXIMIZING DATA usage improving technology

# **FOUNDATION MODELS**

# What the F. Models do with self-supervised learning





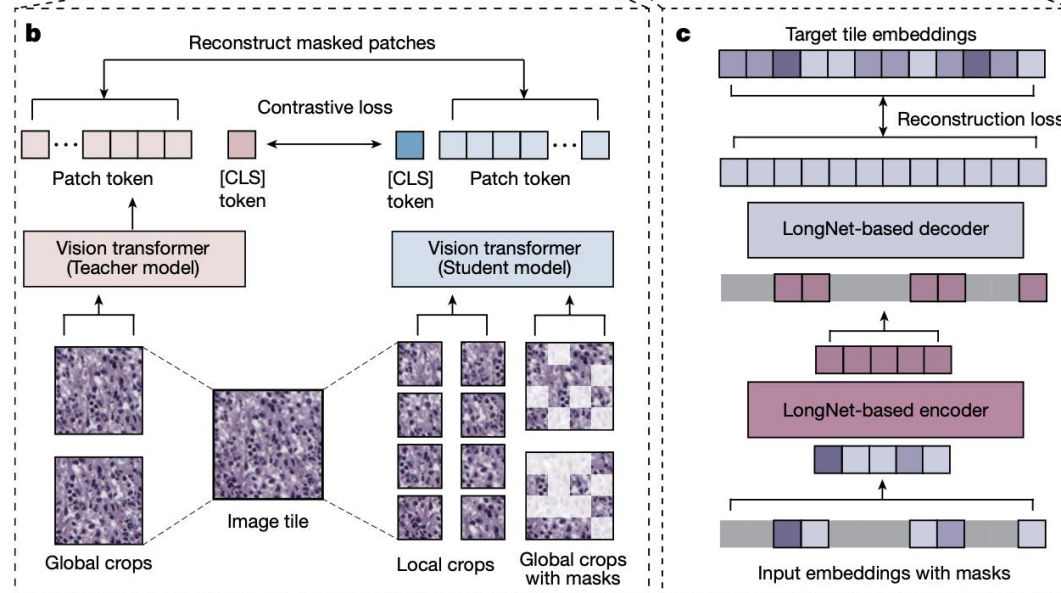
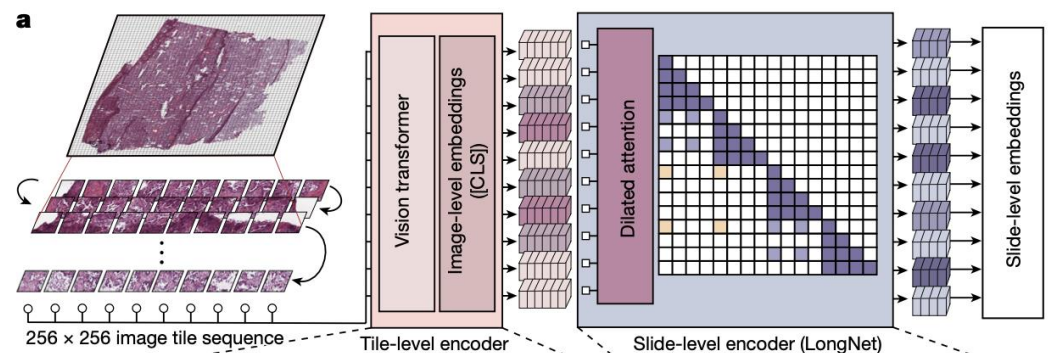
# A whole-slide foundation model for digital pathology from real-world data

**Whole-Slide Modelling:** Prov-GigaPath overcomes subsampling limitations by training on 1.3 billion image tiles from 171,000 slides, preserving full slide context.

**State-of-the-Art Performance:** Achieves top results in 25 of 26 tasks, with significant improvements in 18 tasks, through large-scale pretraining.

**AI-Powered Innovation:** Uses real-world data from 30,000 patients and 31 tissue types, setting a new standard in digital pathology with vision–language models.

# FM for digital pathology: At WS level



# FM for single-cell multi-omics

nature methods

Article

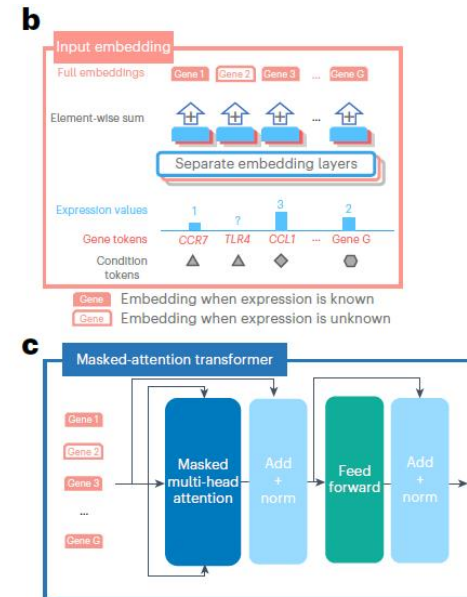
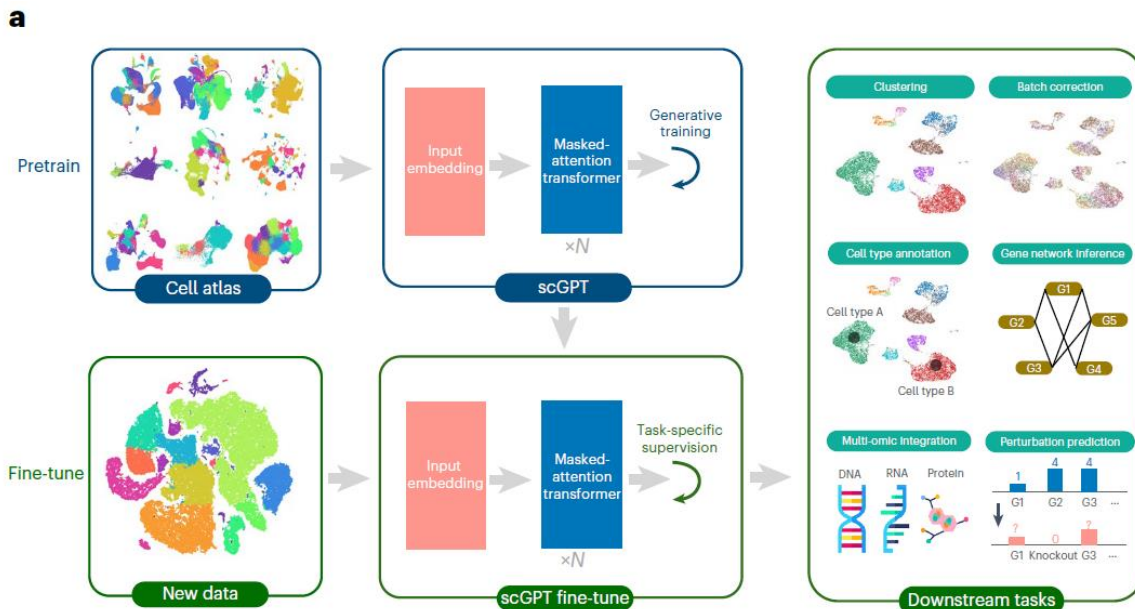
<https://doi.org/10.1038/s41592-024-02201-0>

## scGPT: toward building a foundation model for single-cell multi-omics using generative AI

Received: 12 July 2023

Haotian Cui<sup>1,2,3,8</sup>, Chloe Wang<sup>1,2,3,8</sup>, Hassaan Maan<sup>1,3,4</sup>, Kuan Pang<sup>2,3</sup>,  
Fengning Luo<sup>2,3</sup>, Nan Duan<sup>5</sup> & Bo Wang<sup>1,2,3,4,6,7</sup>✉

Accepted: 30 January 2024



The model **scGPT** is **generatively pretrained on large-scale scRNA-seq data from cell atlases**.

For downstream applications, the pretrained model parameters can be fine-tuned on new data.

They applied scGPT in a variety of tasks including cell type annotation, batch correction, multi-omic integration, genetic perturbation prediction and gene network inference.

# Foundation Model for cancer imaging biomarkers

- Foundation model **pre-training**
- Clinical application** of the Foundation model (3 use cases)
- Foundation model **implementation** (2 approaches)
- Performance** evaluation

Google DeepMind and Google Research

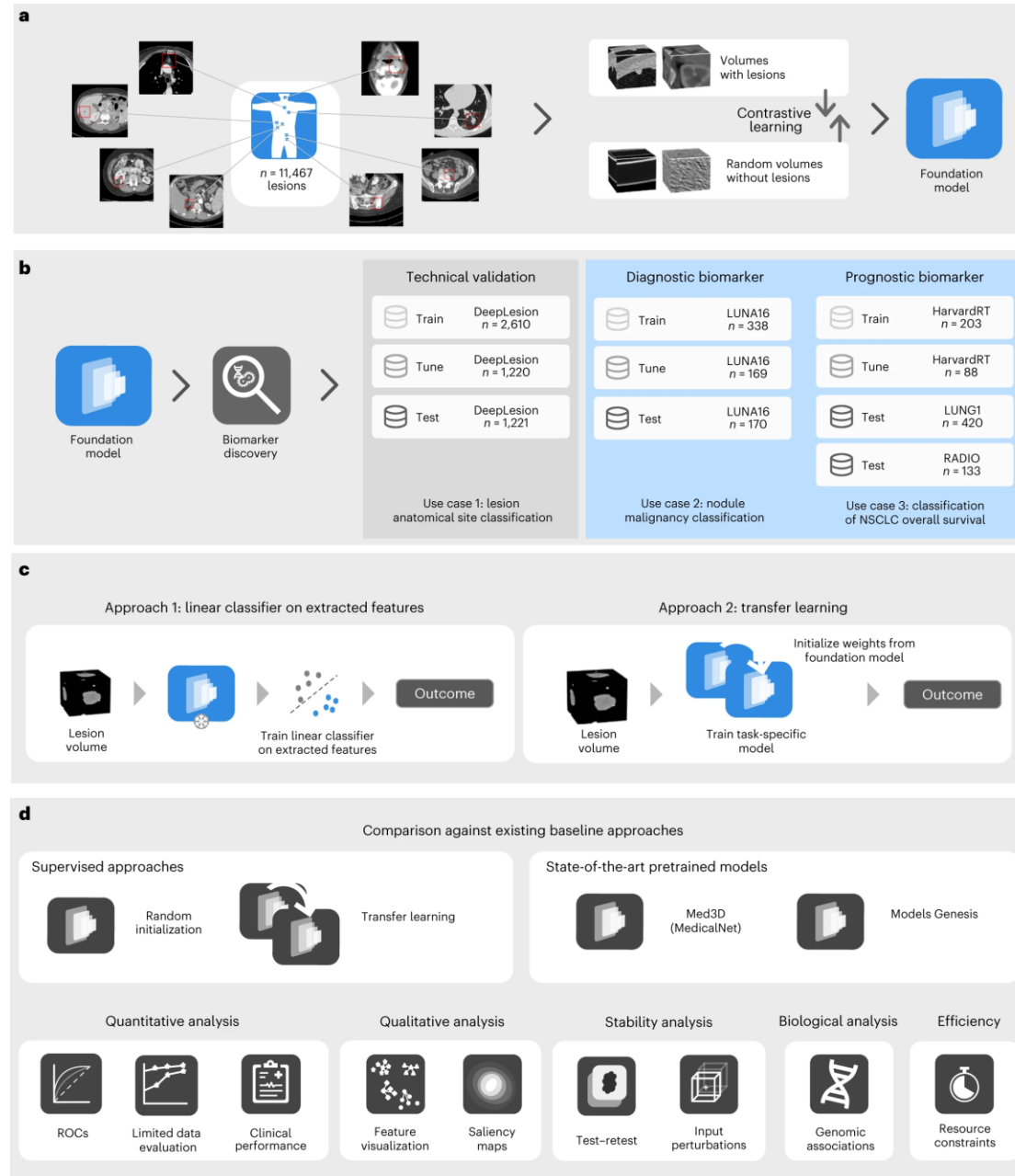
2024-5-7

## Advancing Multimodal Medical Capabilities of Gemini

Google Research and Google DeepMind †

Pai, Suraj, et al. *Nature machine intelligence*, 2024; <https://github.com/Google-Health/imaging-research/tree/master/ct-foundation%0A>

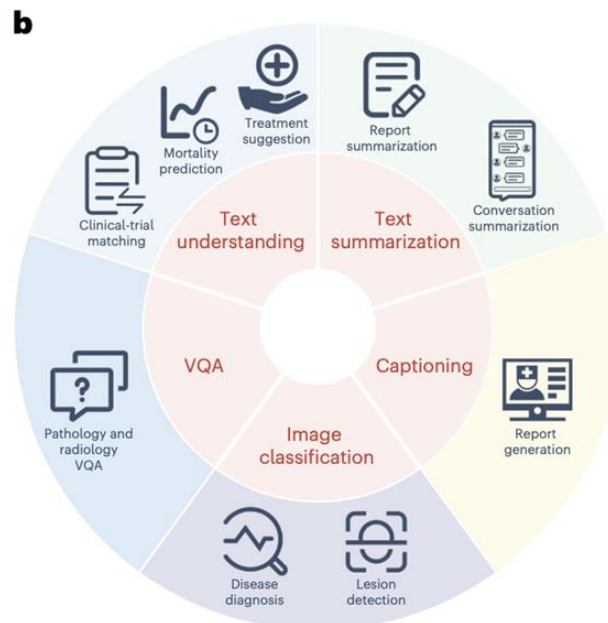
### nature machine intelligence





# A generalist vision language foundation model

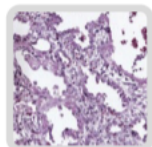
## nature medicine



**BiomedGPT** handles multimodal inputs and performs diverse downstream tasks.

The expected form of output for each task is determined by feeding the specific instruction to the model.

### Pathology and radiology VQA



**Q:** What is seen at this stage, associated with regeneration and repair?

**A:** Numerous reactive type II pneumocytes.

**Q:** Are bite cells like this one in the smear associated with regeneration and repair at this stage?

**A:** No.

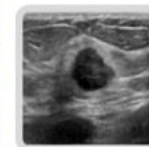
### Report generation



**Q:** What are the findings based on the image?

**A:** The nasogastric tube is in adequate position, and there is a resolution of the gastric distention. There is still mild bibasilar atelectasis. There are no pneumothorax no pleural effusion.

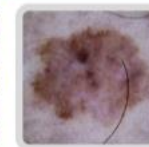
### Disease diagnosis



**Q:** What disease does this image depict?

**A:** Breast cancer.

### Lesion detection



**Q:** What skin lesion does this image depict?

**A:** Melanoma.

Use-cases

# Multimodal generative AI pathchat – Copilot pathchat

PathChat flexibility:

- Multi-choice questions (accuracy image only **78.1%**, when provided with clinical context **89.5%**)
- answering open ended questions (accuracy **78.7%**)
- interactive multi-turn conversation



potentially serve as a  
consultant for human-in-  
the-loop differential



What morphological features are present in this image?

Human

PathChat

The image shows a tumour composed of spindle cells, which are elongated and have a fusiform shape. These cells are arranged in fascicles, which are bundles or groups of cells aligned in a parallel or whorled pattern. Additionally, there are scattered melanin-containing cells within the tumour.

This tumour was found in the liver of a patient. Is it a primary tumour or a metastasis?

Human

PathChat

Based on the information provided, this tumour is most likely a metastasis rather than a primary tumour. The presence of spindle cells and melanin-containing cells further supports the possibility of a metastatic melanoma. The liver is a common site for metastasis of melanoma, especially when it has spread from the skin.



Viewer



Select Category

All

Select Image

TCGA-4P-AA8J-01Z-00...

Annotate

Parse Slide



Theme

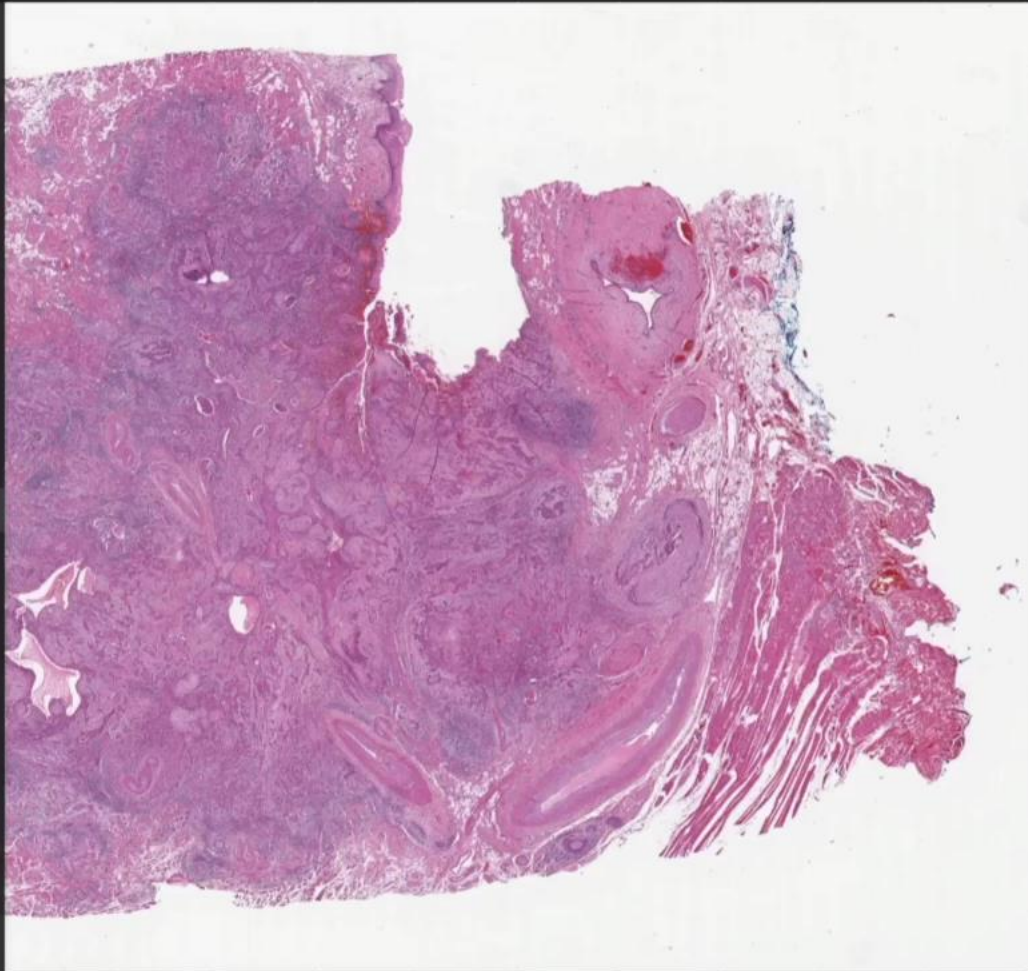


Settings

Logout

P

pathchat.user



# Welcome to PathChat!



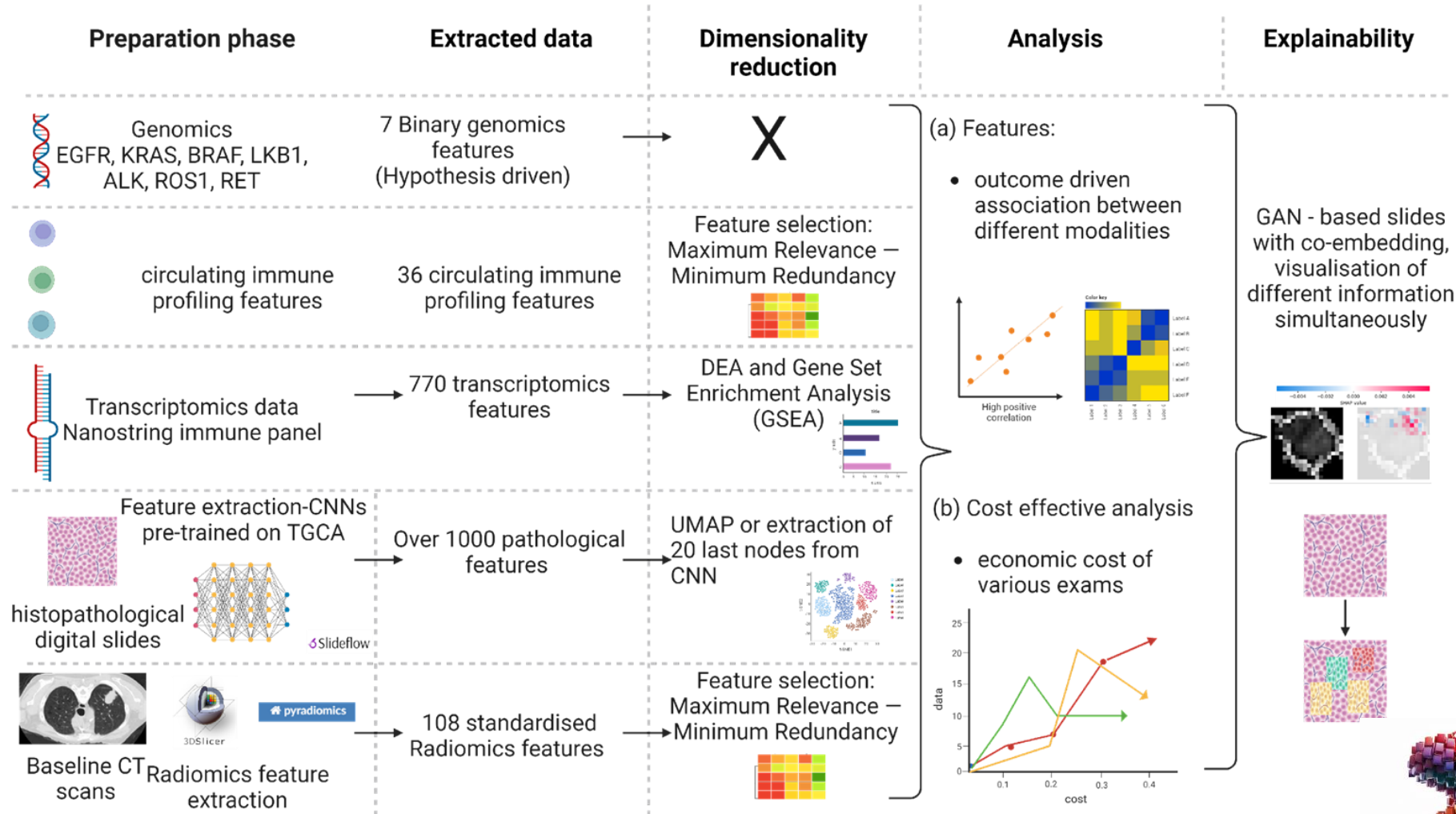
Type your message here...





# HOW we CAN BETTER EXPLAIN WITH FOUNDATION MODELS

# MULTIOMIC- FM-based Explainability in NSCLC patients treated with IO first line (PEOPLE study)



WE NEED TO BUILD DATA FOR AI



# Unity is Strength

Centri di ricerca italiani per strategie terapeutiche avanzate per il tumore al polmone



<https://apollo11.network/>

# 2016 → APOLLO 2020 → APOLLO II FEDERATED NETWORK



## 48 Centri di ricerca avanzata sul tumore al polmone

Creazione di una rete di centri italiani che si occupano di pazienti affetti da tumore del polmone NSCLC avanzati già trattati o candidati a ricevere una terapia a base di ICI



## Biobanche

Sviluppo di una biobanca multilivello nazionale registrata con impostazione e armonizzazione delle procedure operative per la raccolta, la conservazione e la spedizione dei campioni biologici.



## Real world data da tutti i centri attivi

Sviluppo di un database nazionale Real World per i pazienti affetti da tumore del polmone trattati con TERAPIE INNOVATIVE (ad es. immunoterapia, terapie target, anticorpi coniugati)

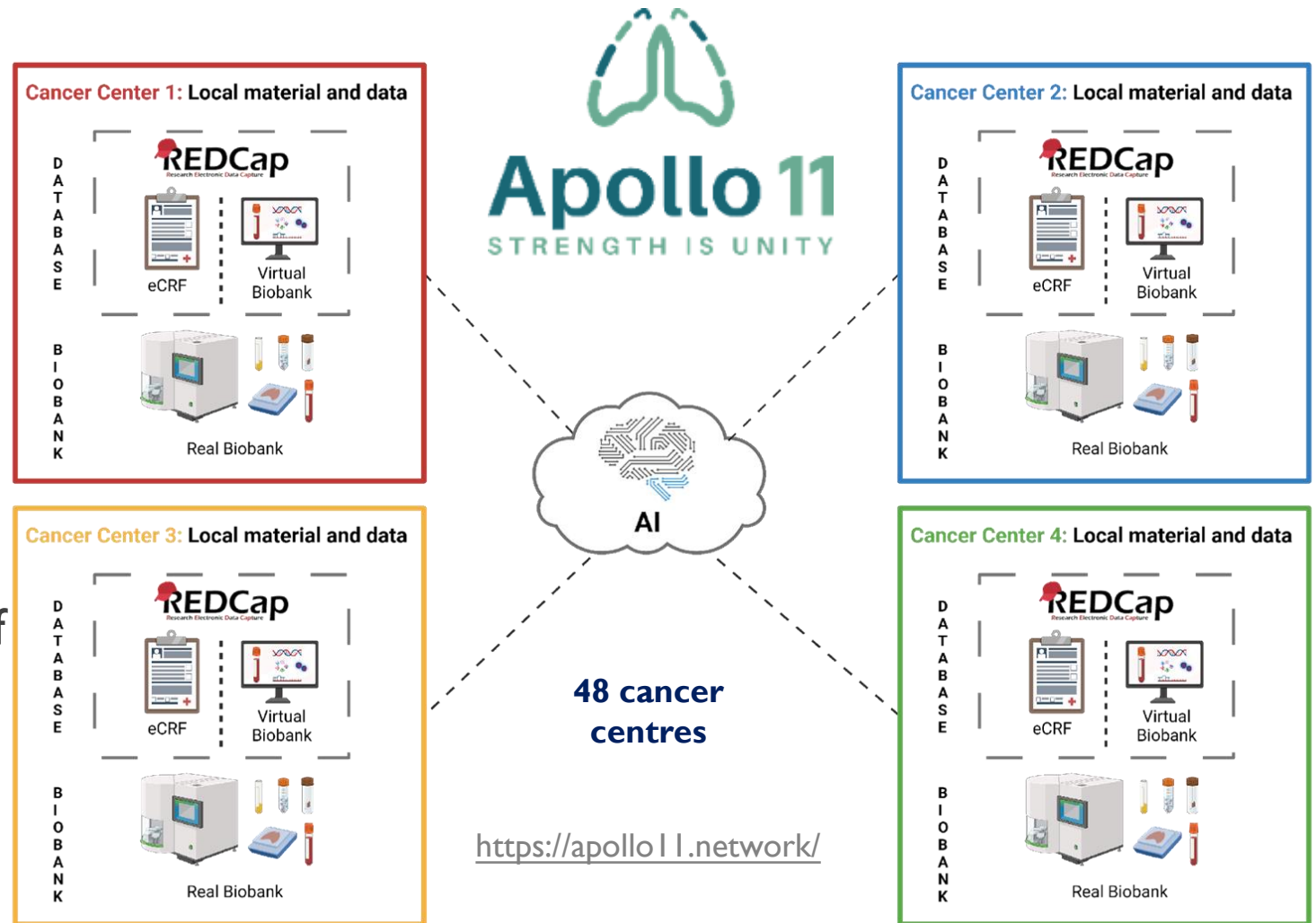


## Intelligenza Artificiale

Creazione di un modello predittivo di intelligenza artificiale (AI) per migliorare la previsione della risposta, portando in ultima analisi a una migliore sopravvivenza e qualità di vita dei pazienti oncologici

# How we can build BIG DATA: Apollo II use case

- RWD/NGS data &
- Image/Digital path collection
- Decentralized Platform
- Biobanking material
- Federated Learning (validation of Swarm Learning architecture)
- AI analysis

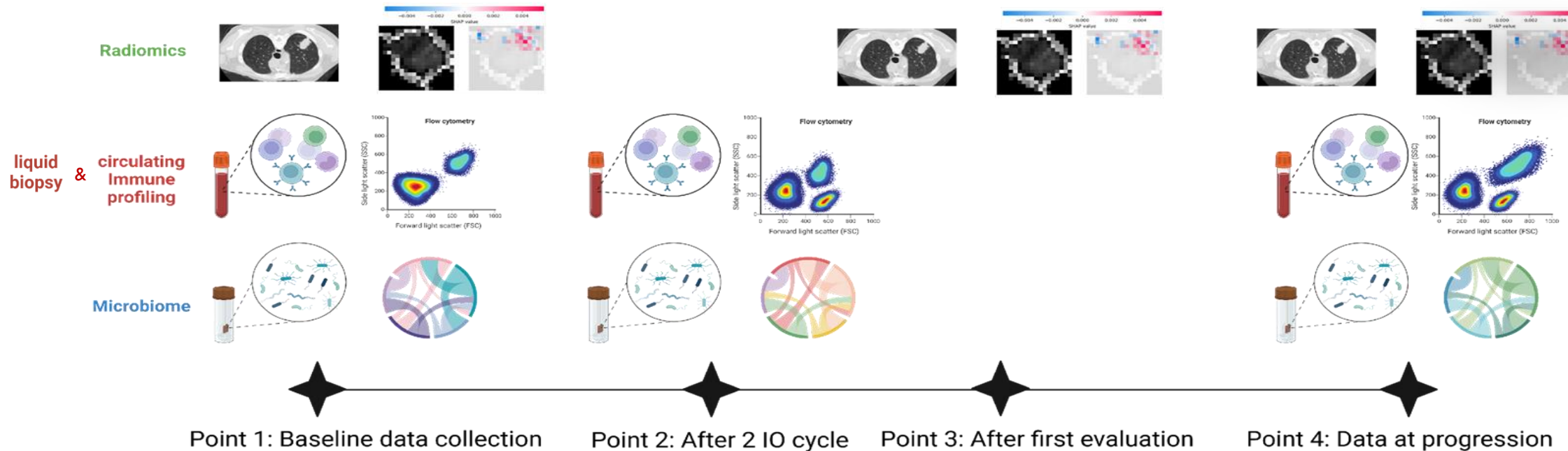


## BIOData Driven Model In Lung Cancer

# PEOPLE interventional prospective trial: 2 host vs 2 tumor time-series biomarkers to explore secondary resistance in IO first-line NSCLC patients treated with immunotherapy



## Evaluation of the influence of longitudinal interventional therapies: e.g., Steroids, antibiotics



No photography

# TRAINING AND EDUCATION

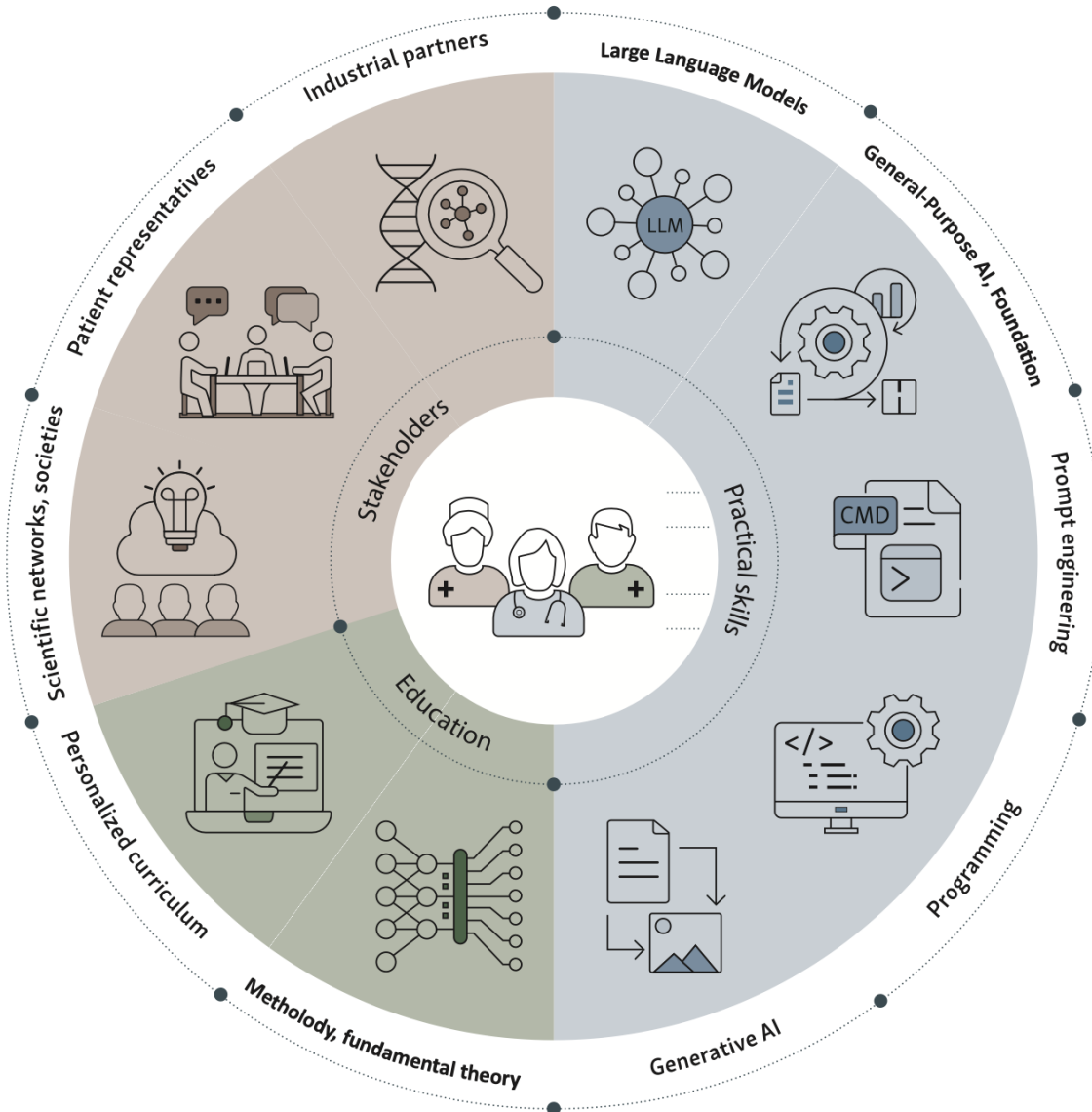


# AI is FUNDAMENTAL

## PERSPECTIVES

### Oncology education in the age of artificial intelligence

A. Prelaj<sup>1,2\*</sup>, G. Scoazec<sup>3</sup>, D. Ferber<sup>4,5</sup> & J. N. Kather<sup>6,7,8</sup>



- Knowledge creation in the medical field
- Information overload mitigation
- Streamlined daily activities
- Increased independence and active development of models

- 1. AI Principles:** Understanding methodology, feature selection, model validation, and bias management.  
**Interpreting**
- 2. AI Data:** Oncologists need to critically assess AI outputs to avoid over-reliance.
- 3. AI-Driven Research:** Oncologists require advanced skills to collaborate with data scientists on AI models.
- 4. Patient-Centric AI:** Oncologists should guide patients using AI tools for self-assessment and health information.

# AI CONGRESS

SAVE THE DATE

3<sup>rd</sup> edition


# ARTIFICIAL INTELLIGENCE FOR ONCOLOGY

PRESIDENT OF THE CONFERENCE  
**ARSELA PRELAJ**

MILAN, ITALY  
**8-9 MAY 2025**  
& ONLINE



## Hibrid Event organized by:

 Fondazione IRCCS  
Istituto Nazionale dei Tumori

Sistema Socio Sanitario  
 Regione  
Lombardia

 **POLITECNICO**  
MILANO 1863

## Invited speakers from:

 **THE UNIVERSITY OF CHICAGO**

 Memorial Sloan Kettering  
Cancer Center

 **TECHNISCHE UNIVERSITÄT DRESDEN**

 **VHIO**  
VALL D'HEBRON  
Institute of Oncology

 **GUSTAVE ROUSSY**  
CANCER CAMPUS  
GRAND PARIS



AALBORG UNIVERSITY

 **UNIVERSITY OF CAMBRIDGE**

 **UCI**  
School of  
Medicine



**HARVARD UNIVERSITY**



## Endorsed by:

**IASLC**



**ASCO**<sup>®</sup>

American Society of Clinical Oncology  
Making a world of difference in cancer care

# FIRST ESMO AI CONFERENCE

ESMO > Meeting Calendar

## ESMO AI & Digital Oncology Congress 2025

### 2025 ESMO AI & DIGITAL ONCOLOGY

Annual Congress

BERLIN GERMANY  
12-14 NOVEMBER 2025



### Scientific Co-Chairs

- Mireia Crispin Ortuzar, Cambridge, UK
- Rudolf Fehrmann, Groningen, Netherlands
- Jakob Kather, Dresden, Germany

390 : 15 : 59 : 44

Days

Hours

Minutes

Seconds



# ACKNOWLEDGEMENTS



Fondazione IRCCS  
Istituto Nazionale dei Tumori

Sistema Socio Sanitario



Regione  
Lombardia



POLITECNICO  
MILANO 1863

## AI-ON Lab



Spetses, Greece June 2024



## The Thoracic Oncology Team

### Medical Oncologist:

Giuseppe Lo Russo, Claudia Proto  
Marta Brambilla, Mario Occhipinti  
Laura Mazzeo, Teresa Beninato  
Rita Leporai, Marco Meazza  
And all the Resident Doctors  
Biologist: Monica Ganzinelli



AI-ON-Lab

### Miriam Fink, **Case Manager**

Diana Andra Dumitrascu,  
Teresa Serra, **Study Nurses**  
Rosa Maria Di Mauro, **Study Coordinator**  
Alice Abate and Other, **Data Manager**  
Sestina Spanò, Fabrizio Baggio, **Project Managers**





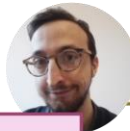
Arsela Prelaj, MD, PhD  
Lab Leader



AI-ON-Lab



Vanja Mišković, PhD  
Post-doc researcher



Giovanni Scoazec  
Project Manager

POLIMI Faculty



Prof. Marcello Restelli



Prof. Alessandra  
Pedrocchi



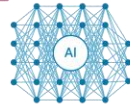
AI-ON-Lab

1. AI-tools

2. Radiomics

3. Digital Pathology

5. App and devices



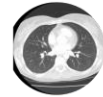
Alberto Ferrarin, MsC  
Research fellow



Laura Mazzeo, MD  
PhD student



Margherita Favali, MsC  
PhD student



Aleksandra Zec, MsC  
PhD student



Chiara Giangregorio,  
MsC, PhD student



Alessandro Chiesi,  
MsC, PhD student



Cecilia Silvestri,  
MD



Leonardo Provenzano,  
MD, PhD student



Andrea Spagnoletti, MD



Cristina Licciardello,  
Master's student

6. Explainable AI

4. Multiomic-AI



Luca Invernizzi, PhD  
Post-doc researcher



Elisa Totis, AI intern



Marco Meazza Prina, MD



Sara Ferri, Master Student



Ghazal Farhikhteh,  
Research fellow



Chiara Cavalli, MD



Beshoy Guirges, AI  
intern



Daniele Lorenzini, MD

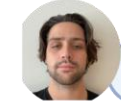
7. LLMs



Monica Ganzinelli,  
Biologist



Cecilia Silvestri,  
MD



Alessandro De Florio,  
Master's student



Marcello Declich,  
Master's student



Federica Corso, PhD  
Post-doc researcher



Vittoria Peppoloni,  
AI intern



Joao P. Ferreira,  
Master's student



Sestina Maria Spanó  
Lab Manager



POLITECNICO  
MILANO 1863



3 LUNG

A project funded by  
the European Union



Apollo 11  
STRENGTH IS UNITY



Fondazione IRCCS  
Istituto Nazionale dei Tumori

Sistema Socio Sanitario



Regione  
Lombardia

# RWD: AUTOMATIC QUALITY CHECK BUILDING

- Data cleaning and automatic quality check coding

1. Free-text inputs: all misspellings corrected
2. Correct date inconsistencies
3. Remove inconsistencies from laboratory values and physical parameters  
(...waiting for automated data curation development...)



2 independent groups (2 oncologists each) blind feature selection  
extracted from eCRF      ? hypothesis-driven selection  
Review of discordancies by 1 experienced medical oncologist

10.000 features



US

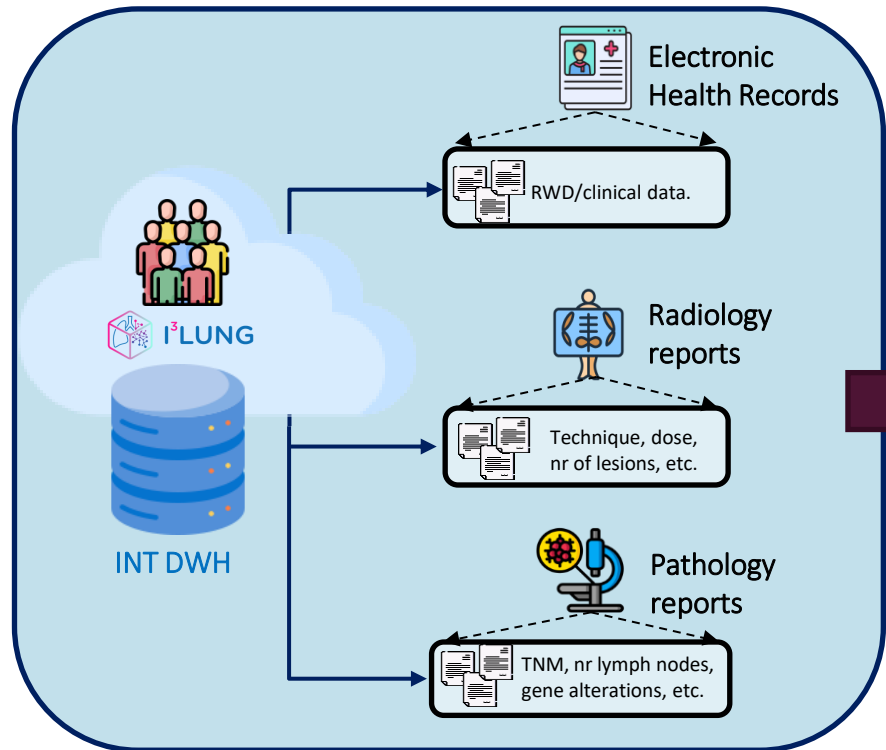
1. Descriptive features

2. Features for model training



Hot-encoding, new features , e.g.: 1) patient's age at different time points; 2) DFI, Time between treatments; 3) number of concomitant diseases, 4) patient therapies, and 5) allergies, 6) body mass index (BMI); 7) composite blood cell counts (NLR); 8) genomic pathways

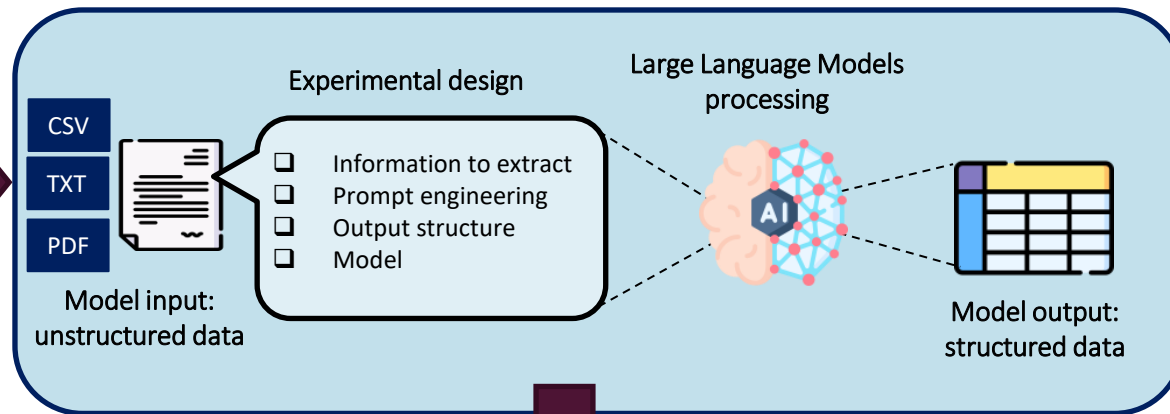
## REAL-WORLD SCENARIO



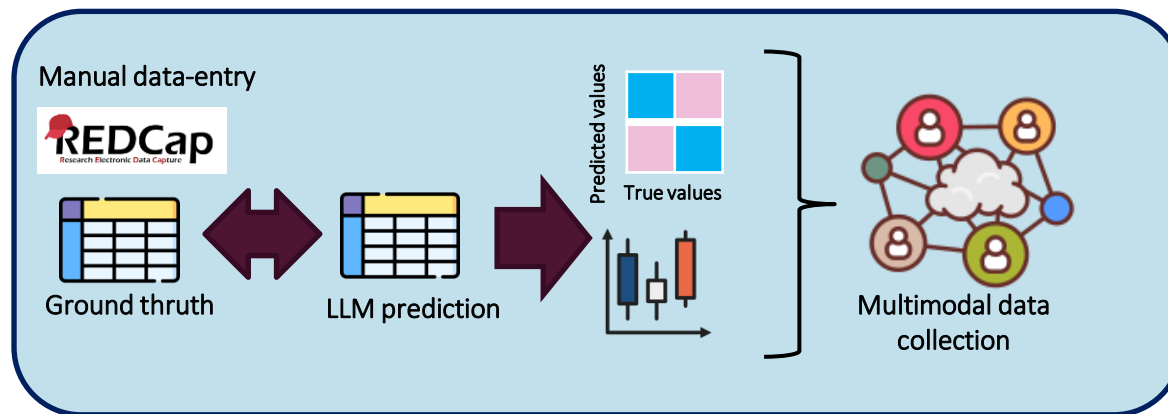
USING LLAMA  
GPT4

# LLMs for RWD structure

## INFORMATION EXTRACTION PIPELINE



## INFORMATION EXTRACTION EVALUATION



GENERALIZABILITY OF THE MODEL

**FAIRNESS AUDITING**



---

# Algorithmic fairness in artificial intelligence for medicine and healthcare









---

Received: 1 October 2021

Accepted: 13 April 2023

Published online: 28 June 2023

---

Richard J. Chen <sup>1,2,3,4</sup>, Judy J. Wang <sup>1,5</sup>, Drew F. K. Williamson <sup>1,3</sup>,  
Tiffany Y. Chen <sup>1,3</sup>, Jana Lipkova <sup>1,2,3</sup>, Ming Y. Lu <sup>1,3,4,6</sup>, Sharifa Sahai<sup>1,2,3,7</sup>  
& Faisal Mahmood <sup>1,3,4,8,9</sup> 

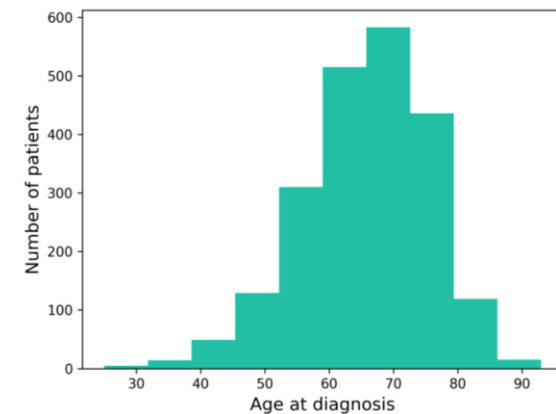
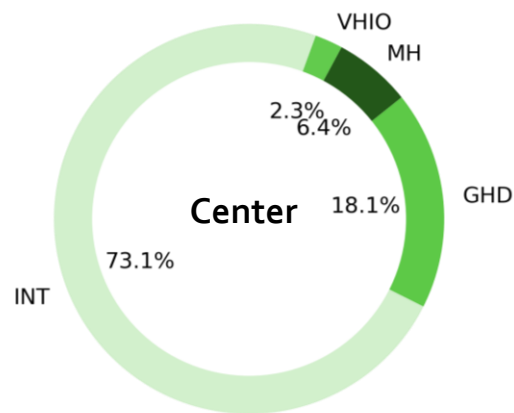
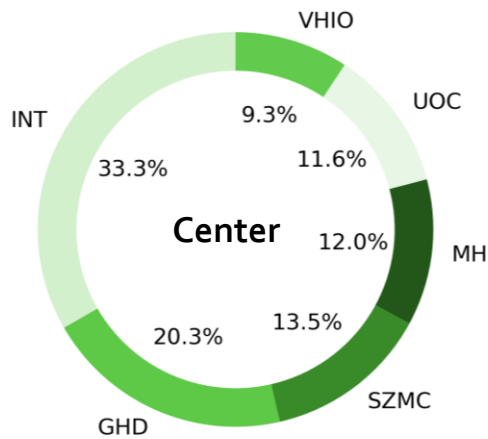
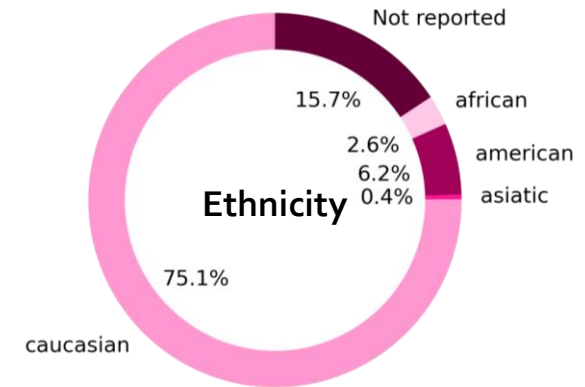
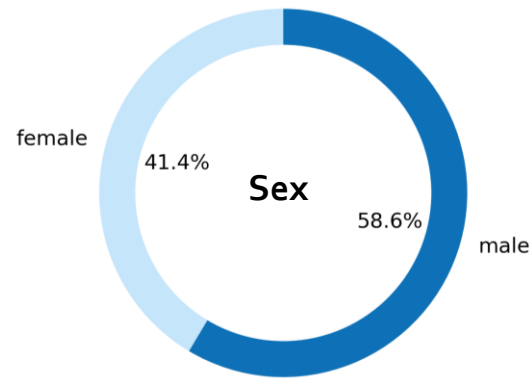
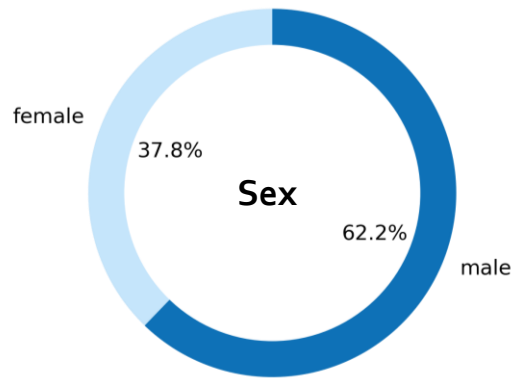
# Preliminary bias analysis



Retrospective (N=2190)

Prospective (until 15 February 2024) N=171

Retrospective



# Comparison of multimodal AI for image analysis

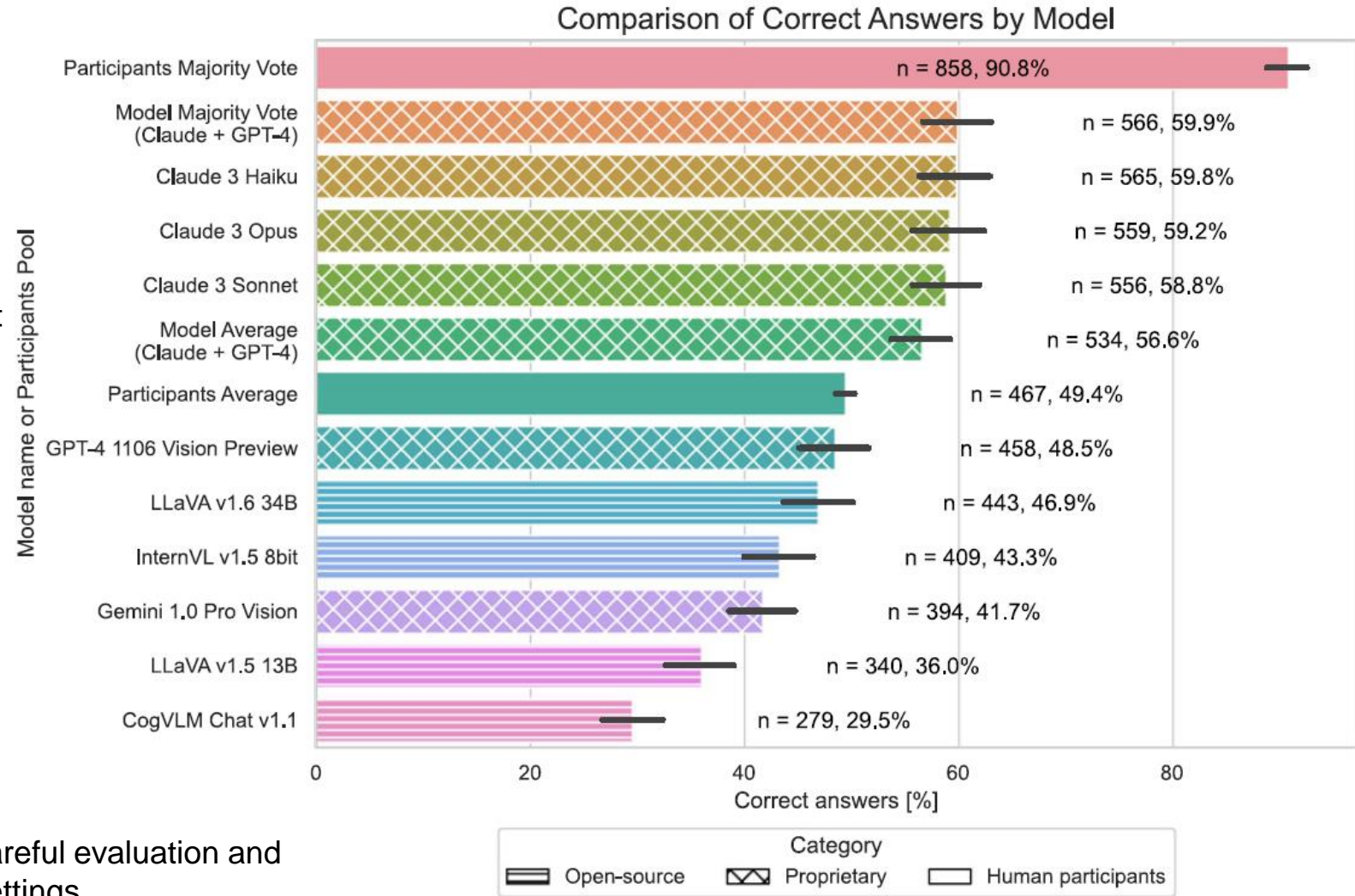
945 NEJM Image Challenge cases.

Nine multimodal AI models, were evaluated using a standardized prompt to answer the medical image challenges.


Anthropic's Claude 3 models achieved the highest accuracy (58.8% to 59.8%).


The study highlights issues such as AI models' selectivity in answering questions, the need for regulatory oversight, and the importance of transparency in AI decision-making.


AI has potential in medical diagnostics but requires careful evaluation and ethical considerations before deployment in clinical settings.





# From Research to Clinic: Perspective


 **EHR**  
Electronic health recorders


 **Genomics NGS**  
Whole genome exome sequencing  
RNA sequencing


 **Radiomics**  
Data from:  
CT and PET

 **Digital pathology**  
Digitized tissue  
Whole slide images

 **Liquid biopsy**  
analysis of non-solid biological tissue

 **Immune cell profiling**  
a snapshot of a state of immune health


 **Single cell analysis**  
cell-specific genetic information


 **Standard Statistics  
Standard AI**

**Clinical data**  
Medical records  
Demographic data

**Laboratory exams**  
Blood exams


 **Standard AI  
Foundation models**

 **Foundation models  
Quantum**

 **EHR**  
Electronic health recorders

**Clinical data**  
Medical records  
Demographic data

**Laboratory exams**  
Blood exams

 **Genomics NGS**  
Whole genome exome sequencing  
RNA sequencing

 **Radiomics**  
Data from:  
CT and PET

 **Digital pathology**  
Digitized tissue  
Whole slide images

