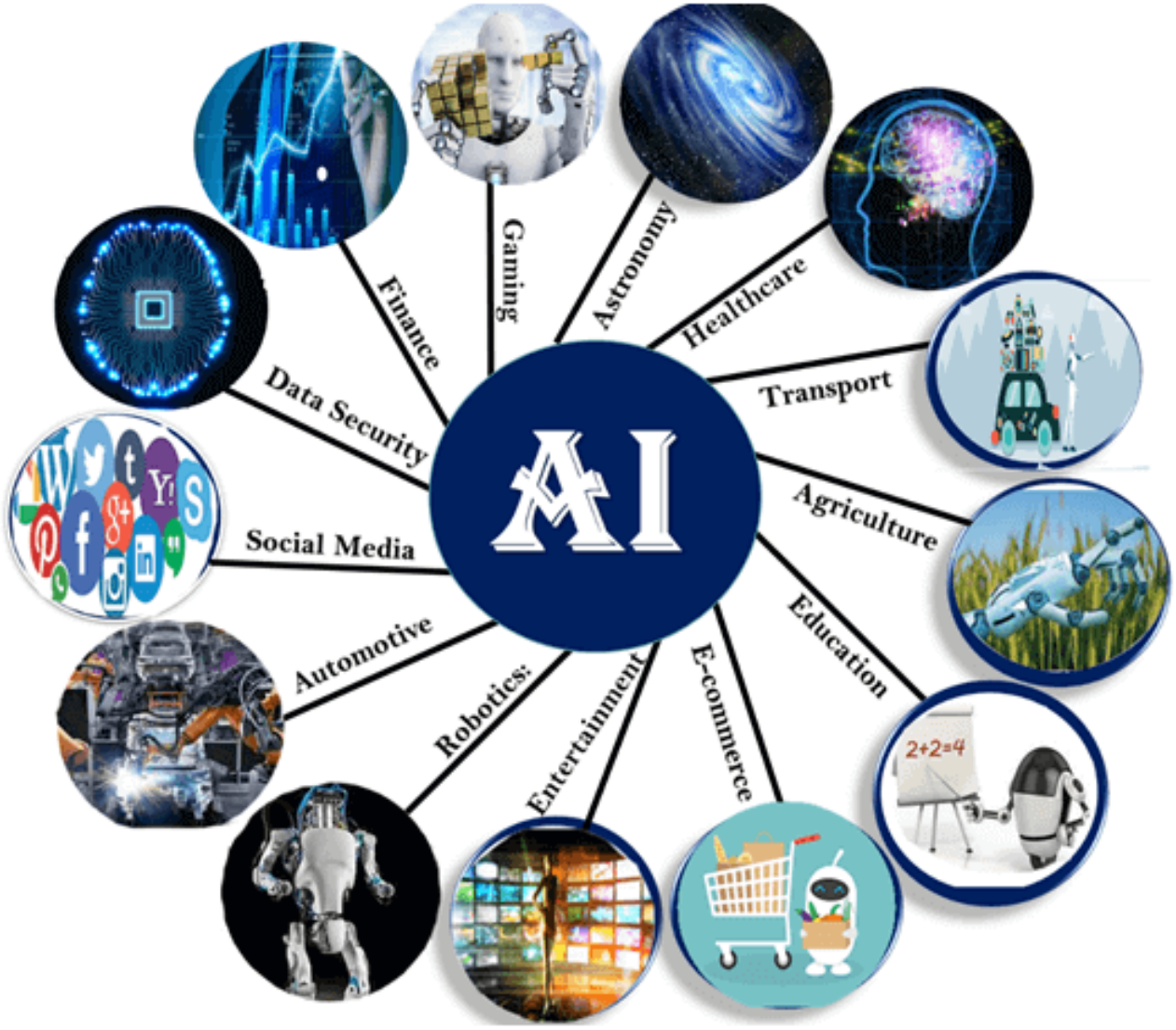
Artificial Intelligence with Lung Cancer

Jun Hyeok Lim
Division of Pulmonology
Department of Internal Medicine
Inha University Hospital

Advancement of AI



Advancement of AI

엔비디아 '딥시크 쇼크' 주가 17%↓... 시총 840조 증발

입력: 2025-01-28 08:58 수정: 2025-01-28 09:23

반도체 덮친 '딥시크 쇼크' 하이닉스 9%↓…네이버는 6%↑(종합)

中 DeepSeek 쇼크, 미국과 글로벌 AI 업계 강타

주형석 기자 입력 01.28.2025 08:45 AM 조회 2,717

> 설립 1년 조금 넘은 中 스타트업 기업 DeepSeek 돌풍
> 현 세계 최고 AI 제품 ChatGPT와 맞먹는 성능 제공
> DeepSeek 혁신 기대감 커지며 IT 기업들 주가 폭락세
> 10년 이상 뒤처진 것으로 알려진 中 고급 기술력에 큰 충격

Reuters

World Business Markets Sustainability Legal Breakingviews Technology Investigati

Trump announces private-sector $500 billion investment in AI infrastructure

By Steve Holland

January 22, 2025 12:42 PM GMT+9 · Updated 10 days ago

Advancement of AI

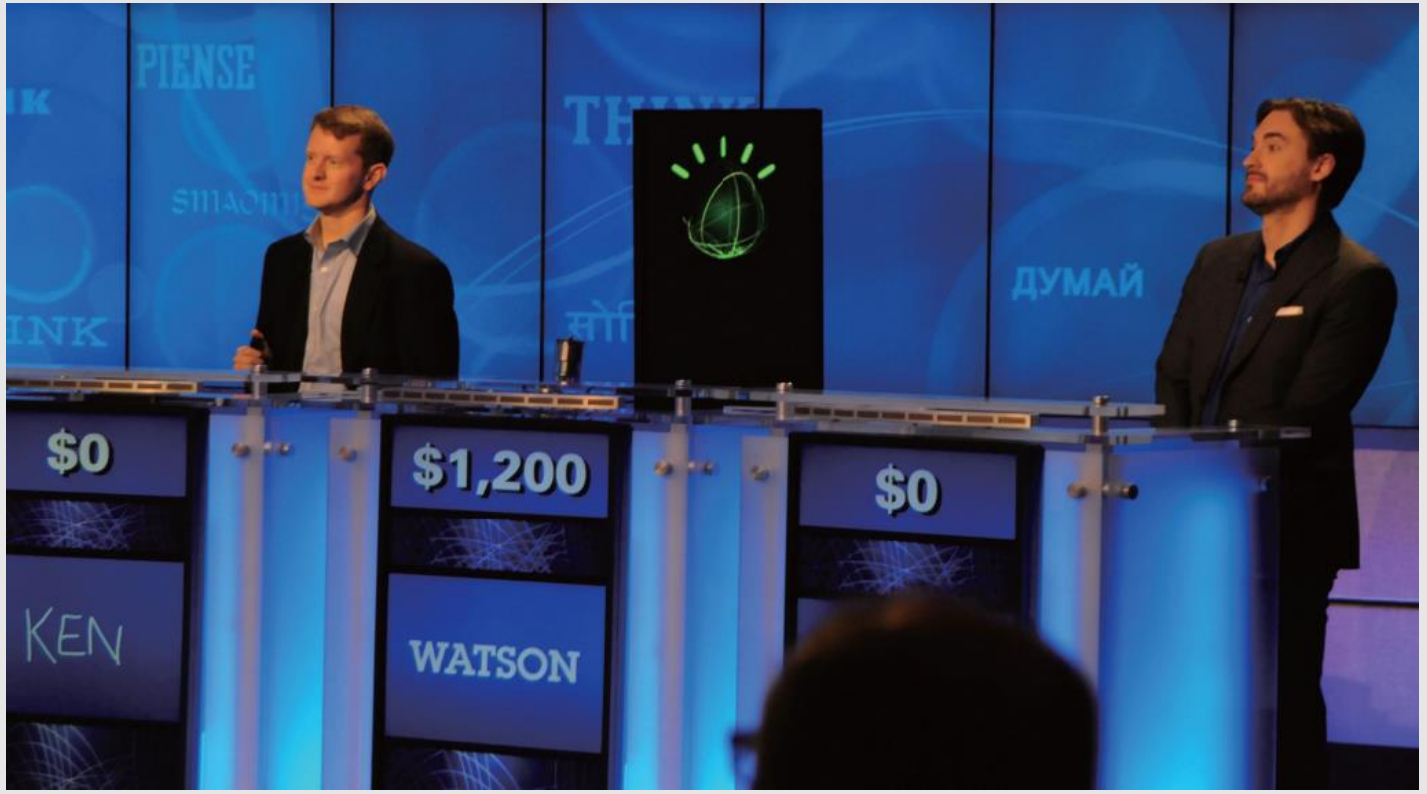
Pattern


더지니어스
GRAND FINAL
2R : 미스터리 사인
김경훈 1 3 장동민
5R 11 14 = ?
HINT 79 897687543217 = 749282615141312111
HINT 897687543217 89 = 839273615141312111
HINT 897687543217 897687543217 = 849276625242322212
HINT 99 897687543217 = 938273615141312111

Pattern


11 ? 14 = 1341
더지니어스
GRAND FINAL
2R : 미스터리 사인
김경훈 1 3 장동민
5R 11 ? 14 = ?
HINT 79 ? 897687543217 = 749282615141312111
HINT 897687543217 ? 89 = 839273615141312111
HINT 897687543217 ? 897687543217 = 849276625242322212
HINT 99 ? 897687543217 = 938273615141312111

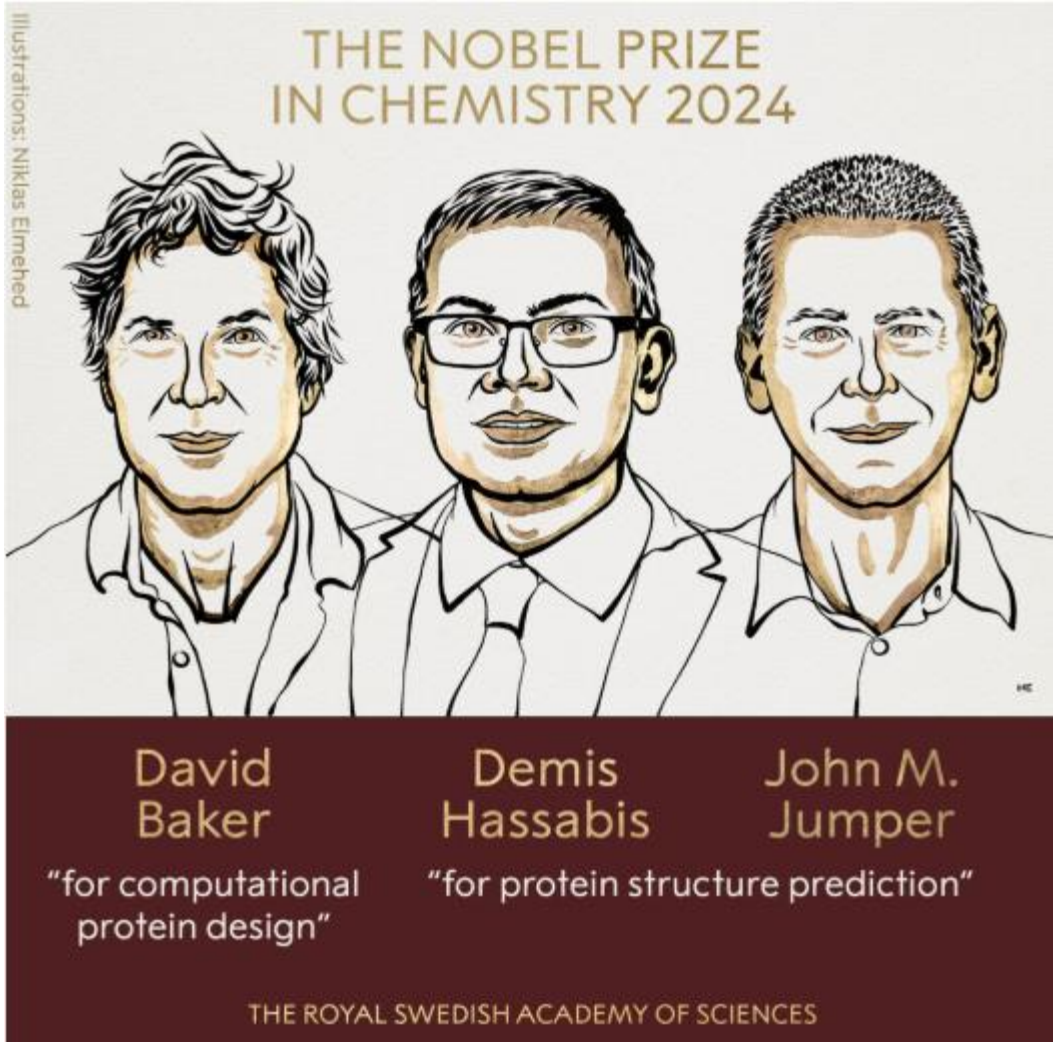
Pattern



Pattern

AlphaGo

Pattern

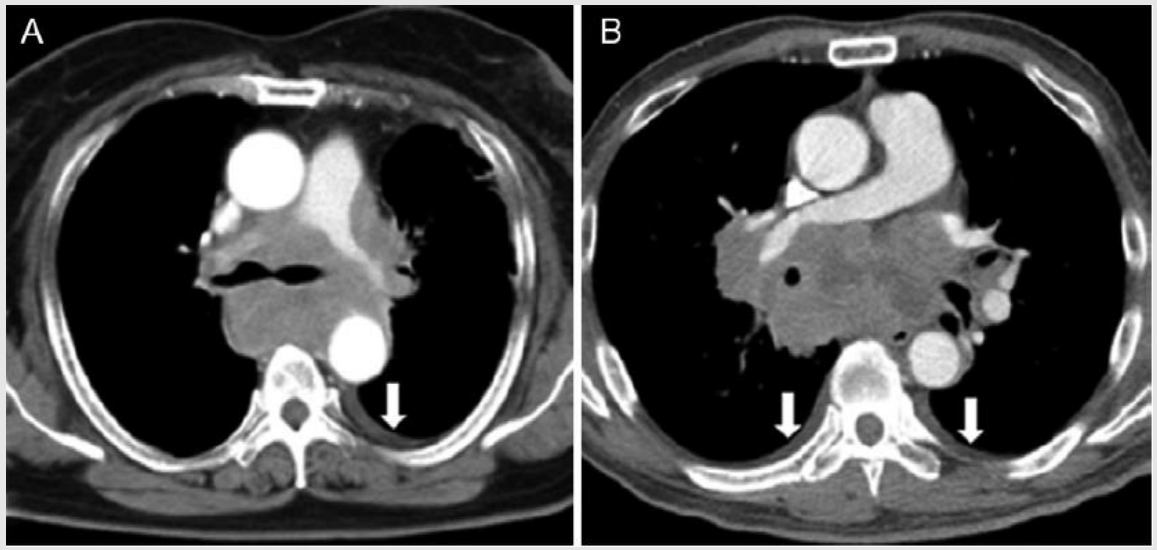


AlphaFold

Accelerating breakthroughs in biology with AI

Explore the AlphaFold Database

Sensitivity



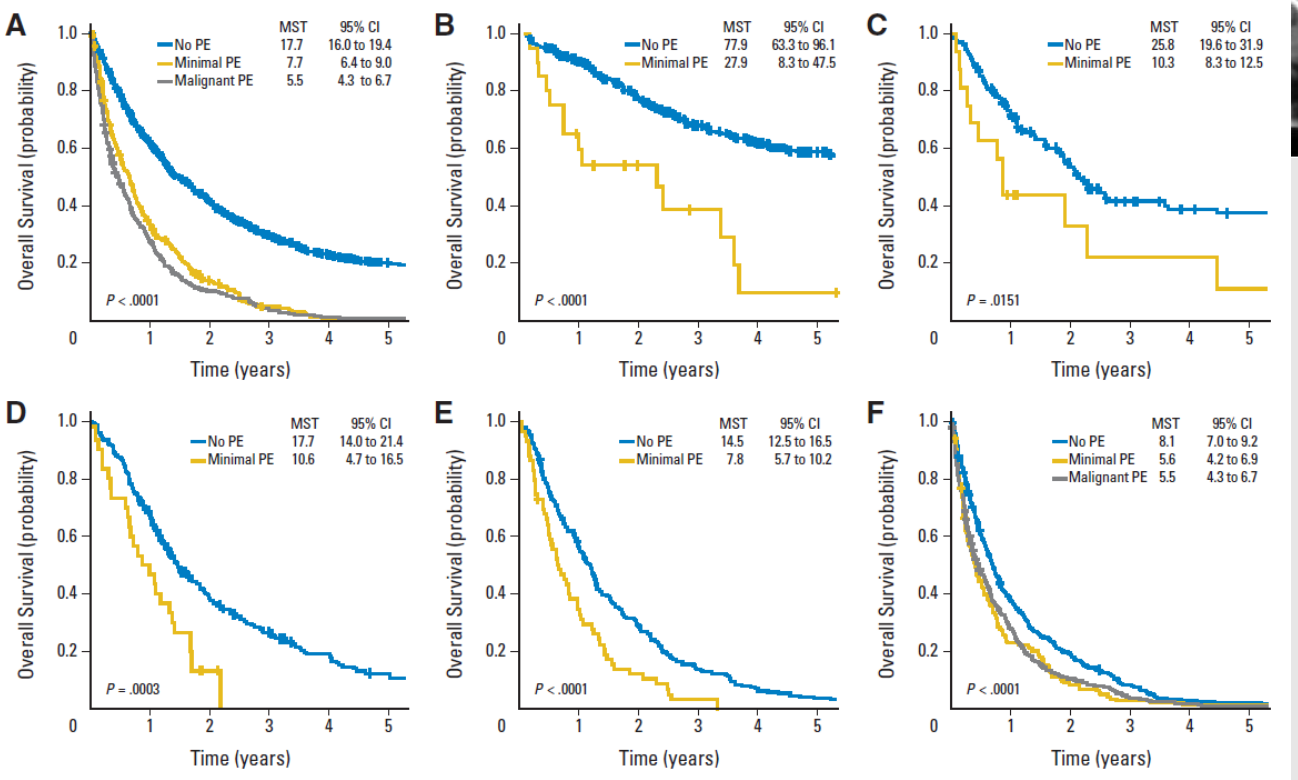
Sensitivity

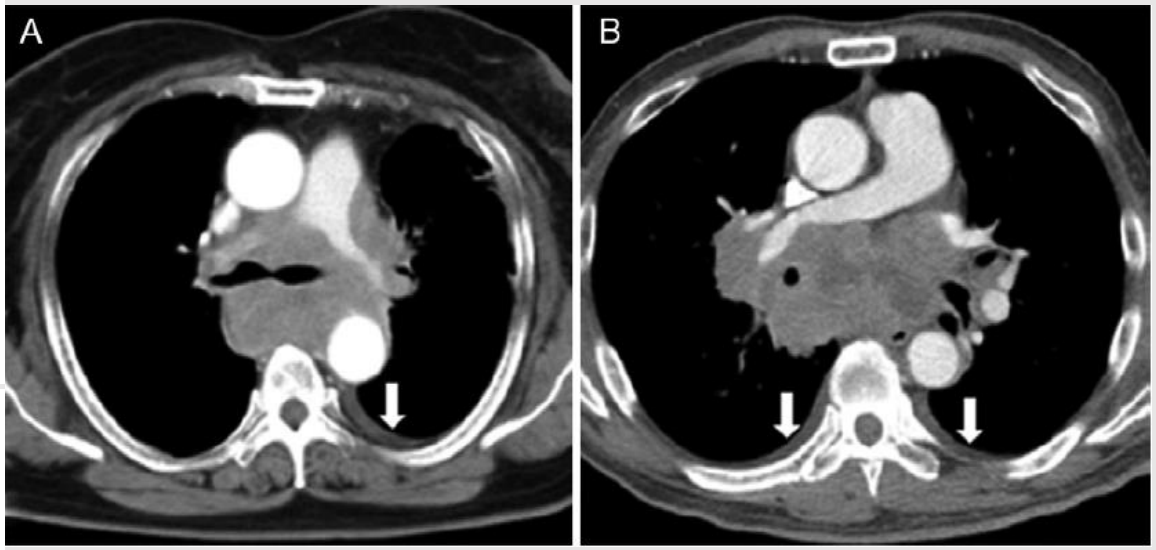
JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Prognostic Impact of Minimal Pleural Effusion in Non–Small-Cell Lung Cancer

Jeong-Seon Ryu, Hyo Jin Ryu, Si-Nae Lee, Azra Memon, Seul-Ki Lee, Hae-Seong Nam, Hyun-Jung Kim, Kyung-Hee Lee, Jae-Hwa Cho, and Seung-Sik Hwang





Sensitivity

Radiology

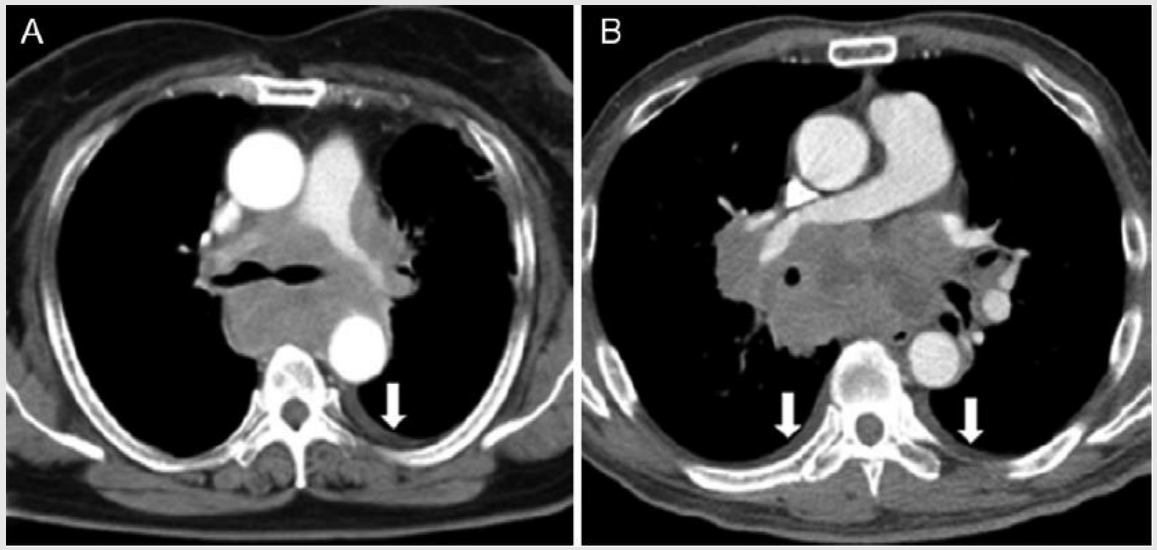
Minimal Pleural Effusion in Small Cell Lung Cancer: Proportion, Mechanisms, and Prognostic Effect[1]

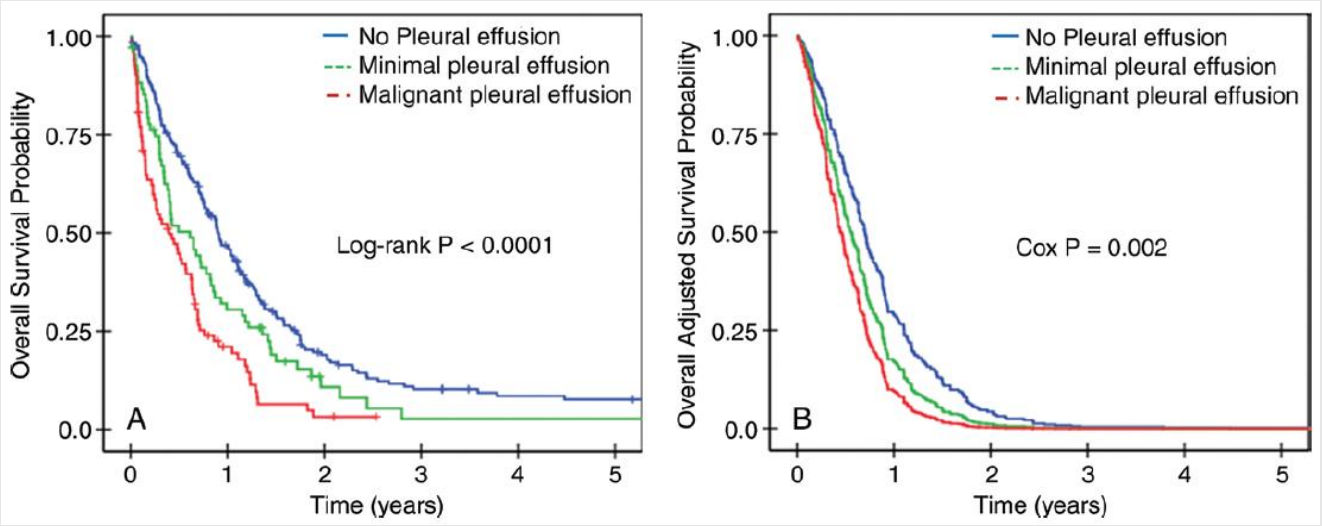
Jeong-Seon Ryu, MD, PhD
Jun Hyeok Lim, MD
Jeong Min Lee, MD
Woo Chul Kim, MD
Kyung-Hee Lee, MD, PhD
Azra Memon, MD
Seul-Ki Lee, MD
Bo-Rim Yi, MD
Hyun-Jung Kim, RN
Seung-Sik Hwang, MD, PhD

ORIGINAL RESEARCH ■ THORACIC IMAGING

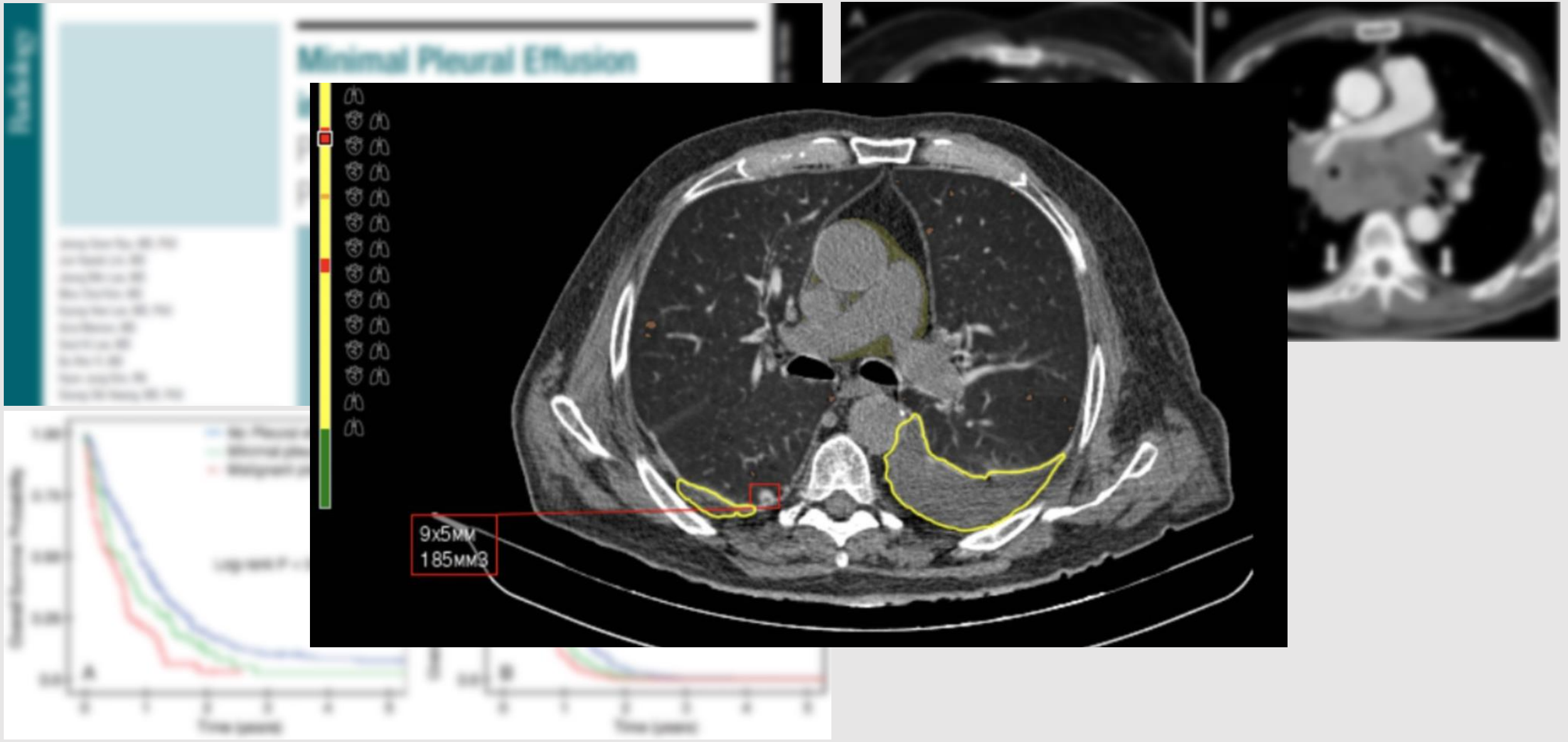
Purpose: To determine the frequency and investigate possible mechanisms and prognostic relevance of minimal (<10-mm thickness) pleural effusion in patients with small cell lung cancer (SCLC).

Materials and Methods: The single-center retrospective study was approved by the institutional review board of the hospital, and informed consent was waived by the patients. A cohort of 360 consecutive patients diagnosed with SCLC by using histologic analysis was enrolled in this study. Based on the status

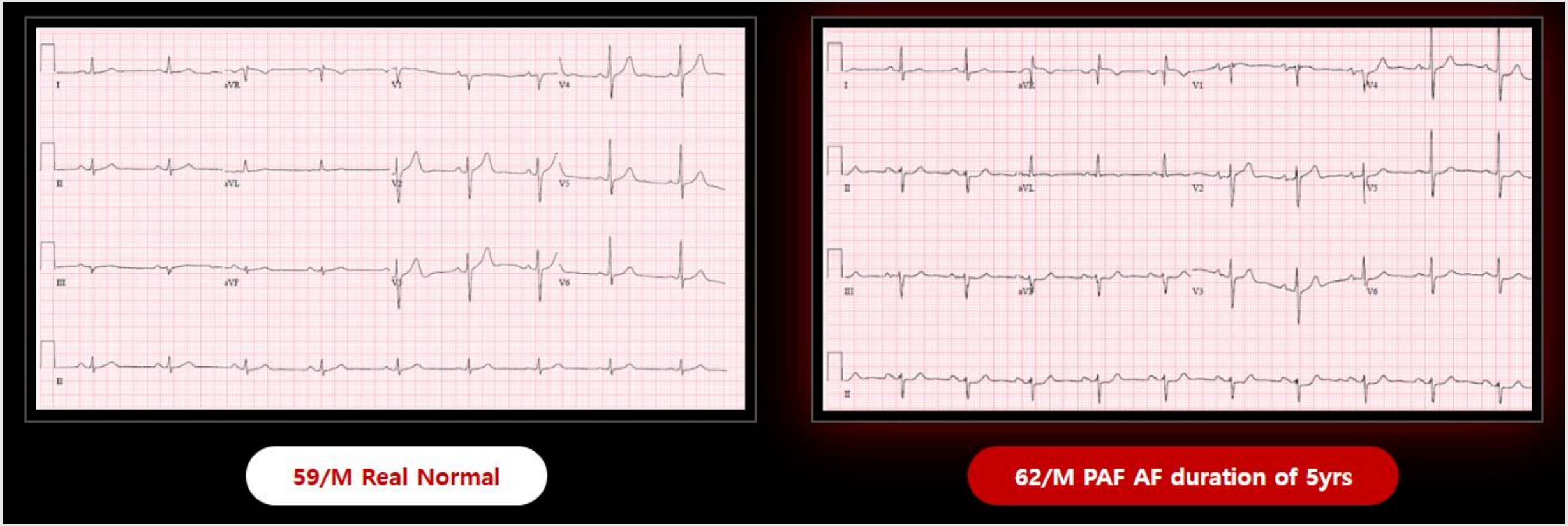




Sensitivity



Sensitivity



Good research?

High-Quality Data!

Big database
Prospective
Multicenter

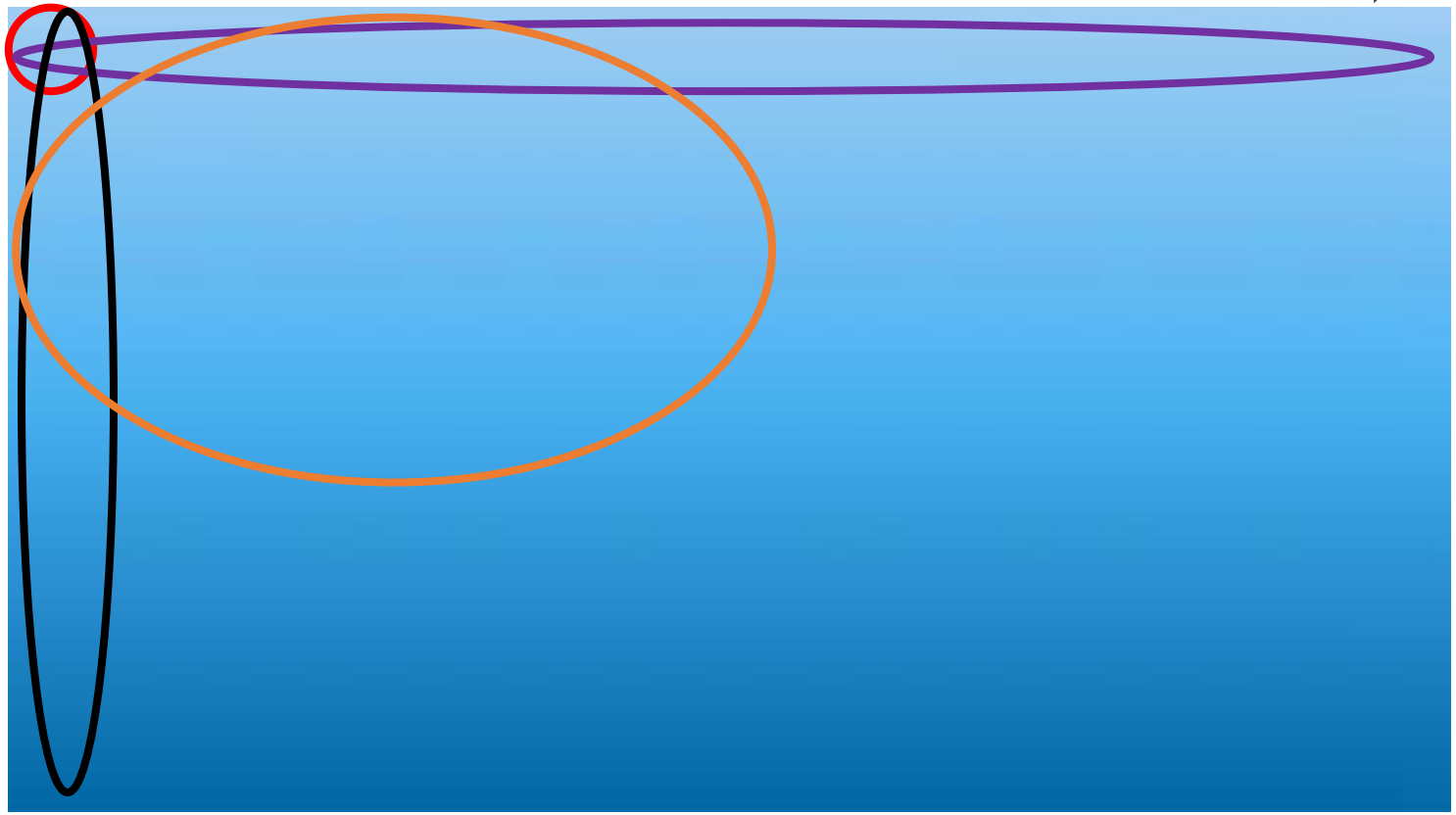
Multi-omics
Experimental validation (animal, organoid…)

Good research?

High-Quality Data!

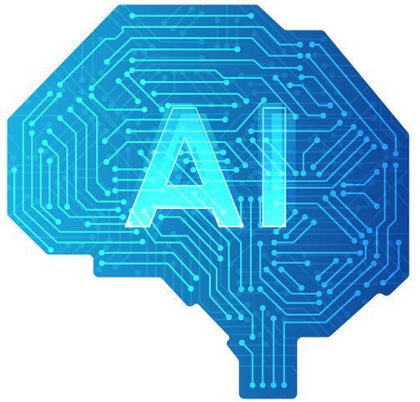
데이터의 크기

데이터의 깊이



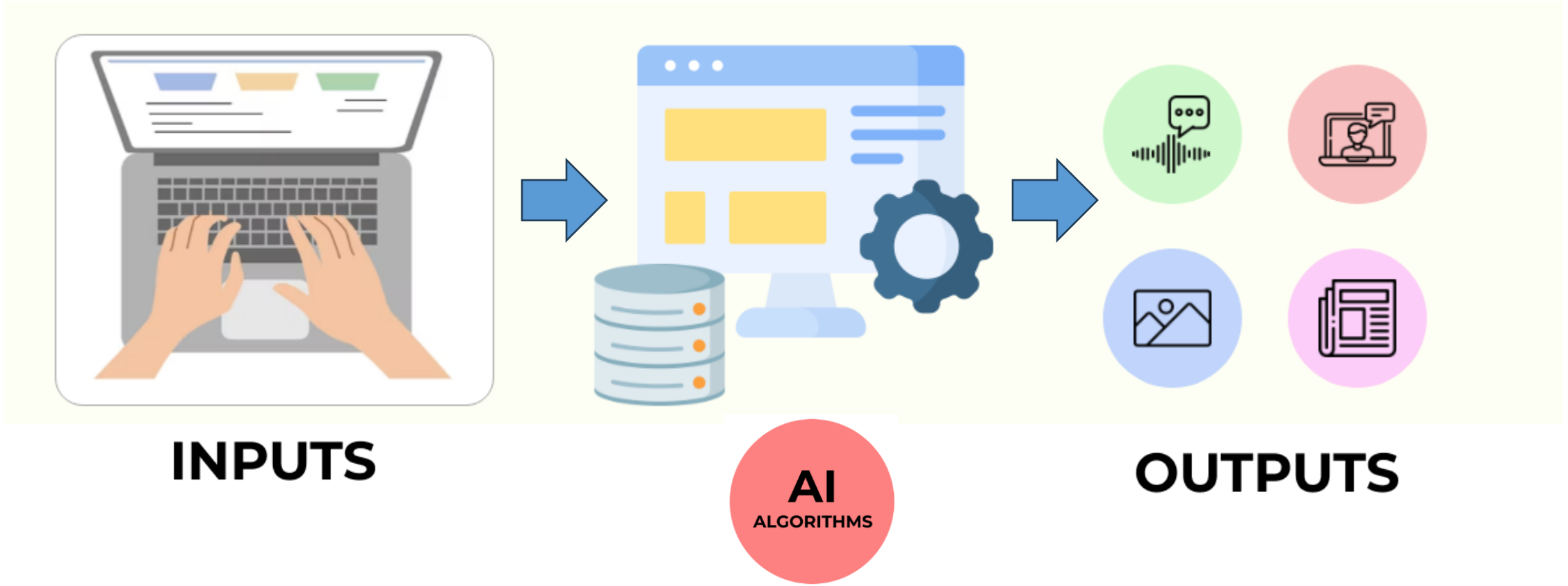
Big database
Prospective
Multicenter

Multi-omics
Experimental validation (animal, organoid…)

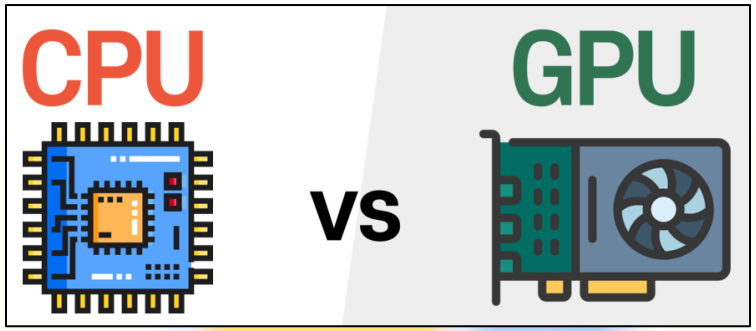


Pattern
Sensitivity

Application of AI in Lung Cancer Research



Application of AI in Lung Cancer Research

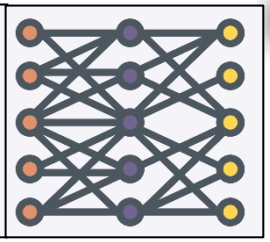


python™

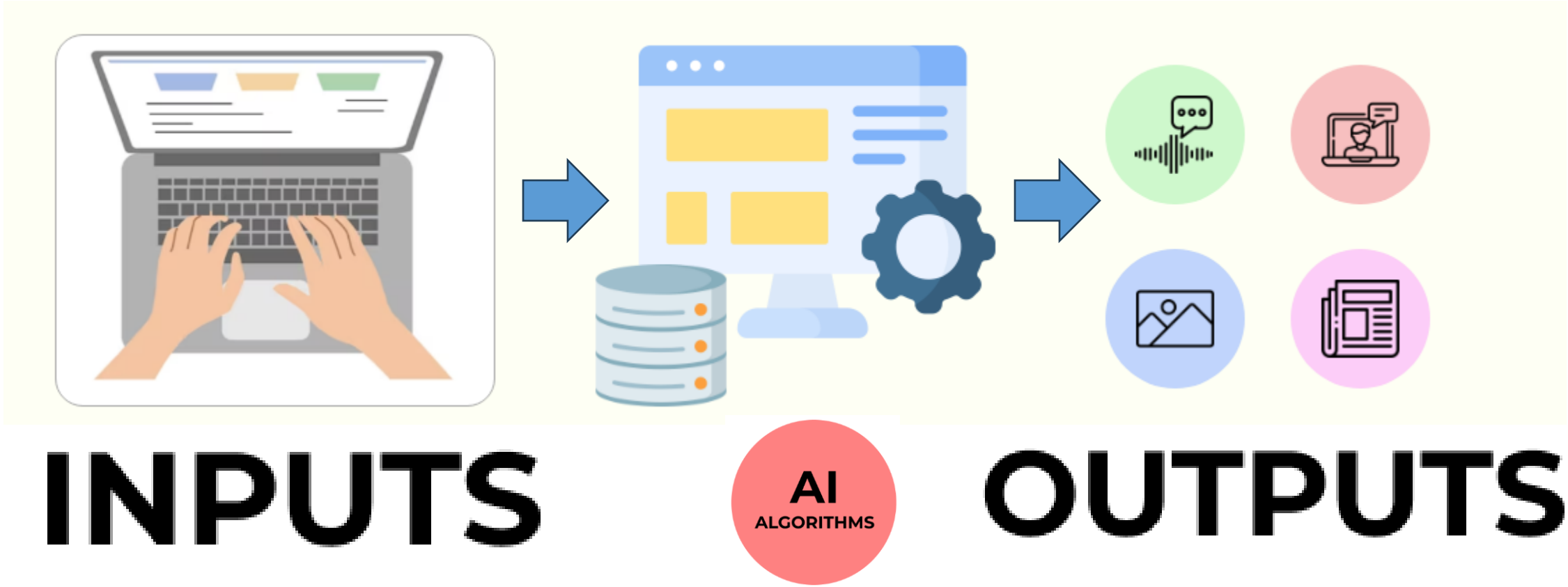
TensorFlow

PYTORCH

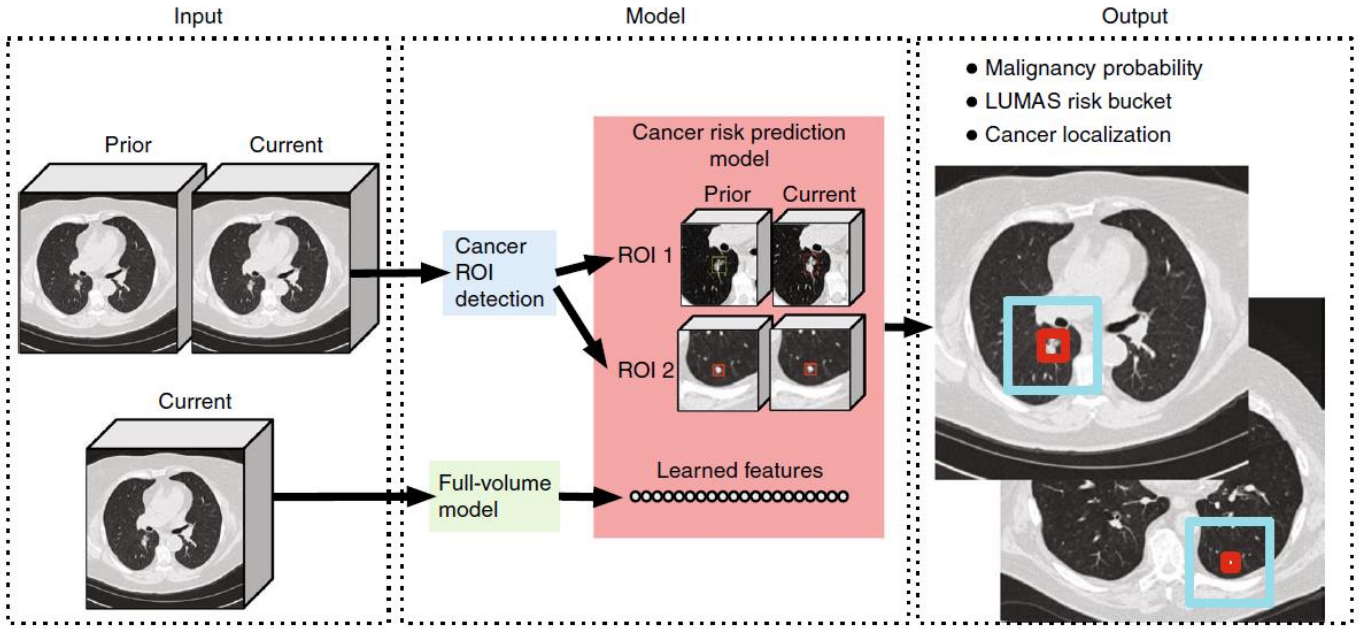
Neural Networks:
CNN vs ANN vs RNN

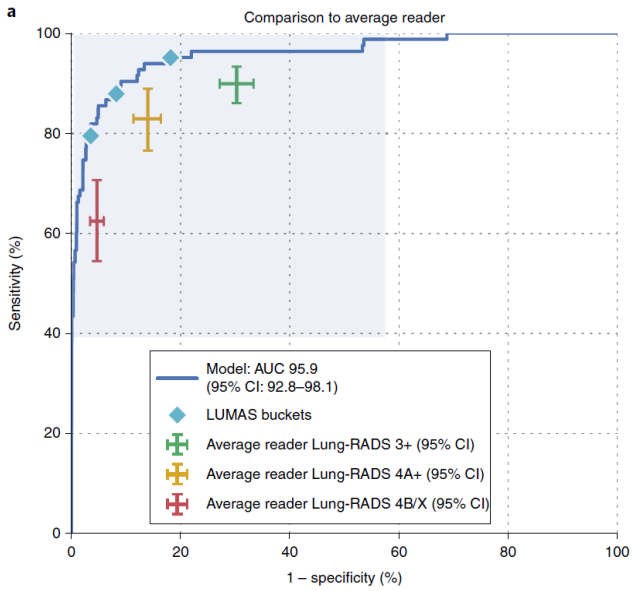


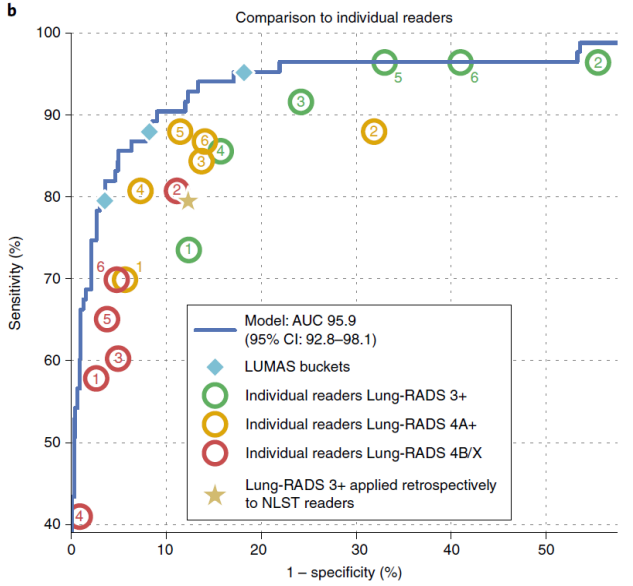
Application of AI in Lung Cancer Research



Clinical Research: Radiology







c

Risk buckets		Sensitivity	Delta
1,2 versus 3+	Average reader	90.0 (86.1, 93.4)	+5.2*
	Model	95.2 (89.9, 98.9)	(.4, 9.8) $P = 0.0386$
1,2,3 versus 4A+	Average reader	82.9 (76.6, 89.4)	+7.4*
	Model	90.4 (83.3, 96.3)	(1.7, 12.9) $P = 0.0114$
1,2,3,4A versus 4B/X	Average reader	62.5 (54.4, 70.7)	+17.1*
	Model	79.5 (70.8, 88.2)	(9.0, 24.5) $P < 1 \times 10^{-4}$

d

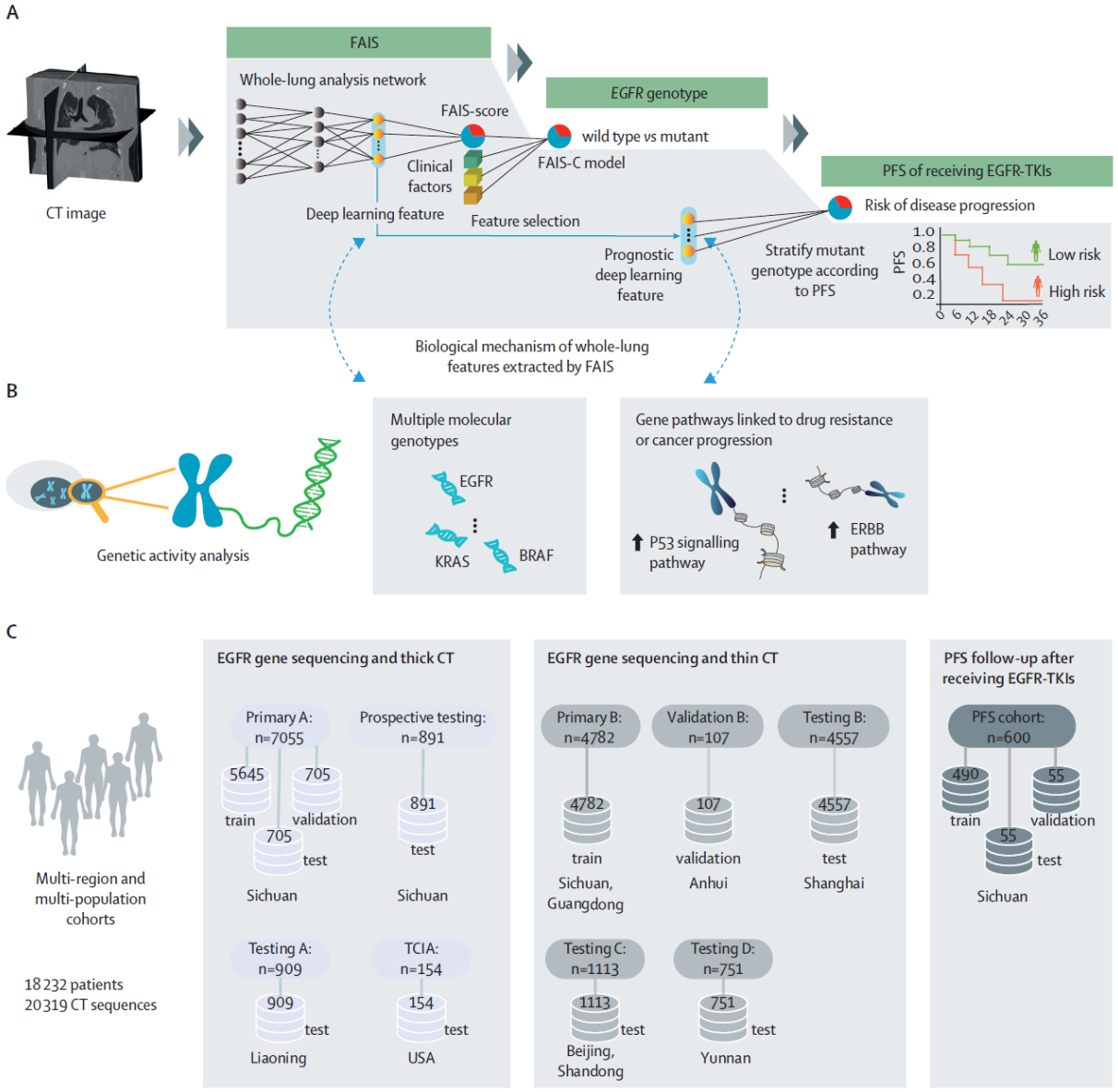
Risk buckets		Specificity	Delta
1,2 versus 3+	Average reader	69.7 (66.6, 72.8)	+11.6*
	Model	81.3 (77.3, 84.9)	(7.8, 15.1) $P < 1 \times 10^{-4}$
1,2,3 versus 4A+	Average reader	86.0 (83.4, 88.4)	+5.0*
	Model	91.0 (88.1, 93.9)	(1.7, 12.9) $P = 0.0008$
1,2,3,4A versus 4B/X	Average reader	95.3 (94.0, 96.6)	+1.1*
	Model	96.5 (94.6, 98.2)	(–0.4, 2.6) $P = 0.143$

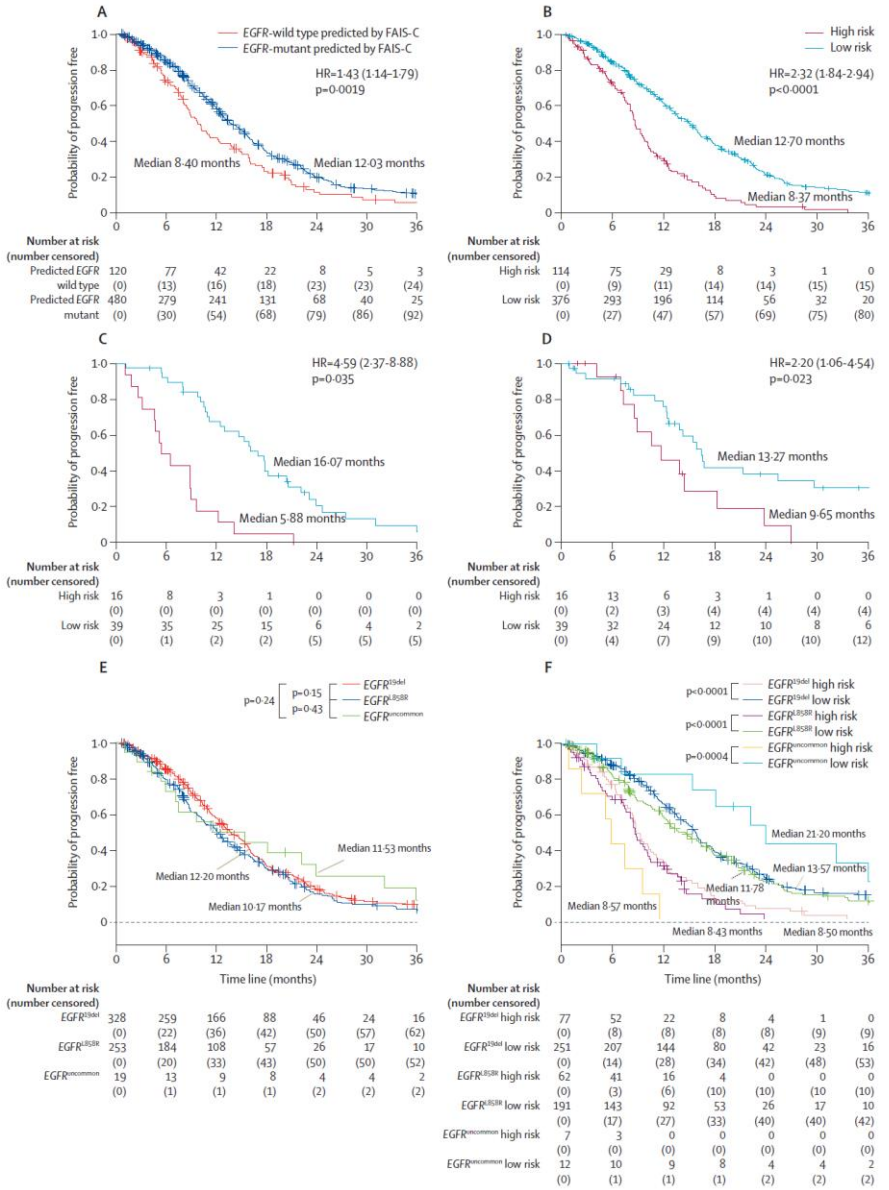
e

Risk buckets		
1,2 versus 3+	Hit@1	73/74
	Hit@2	74/74
1,2,3 versus 4A+	Hit@1	72/73
	Hit@2	73/73
1,2,3,4A versus 4B/X	Hit@1	62/63
	Hit@2	63/63

Ardila et al., *Nat Med*, 2019

Clinical Research: Radiology





Wang et al., *Lancet Oncol*, 2022

Clinical Research: Radiology

Radiology

ORIGINAL RESEARCH • THORACIC IMAGING

Clinical Utility of a CT-based AI Prognostic Model for Segmentectomy in Non–Small Cell Lung Cancer

Kwon Joong Na, MD • Young Tae Kim, MD, PhD • Jin Mo Goo, MD, PhD • Hyungjin Kim, MD, PhD

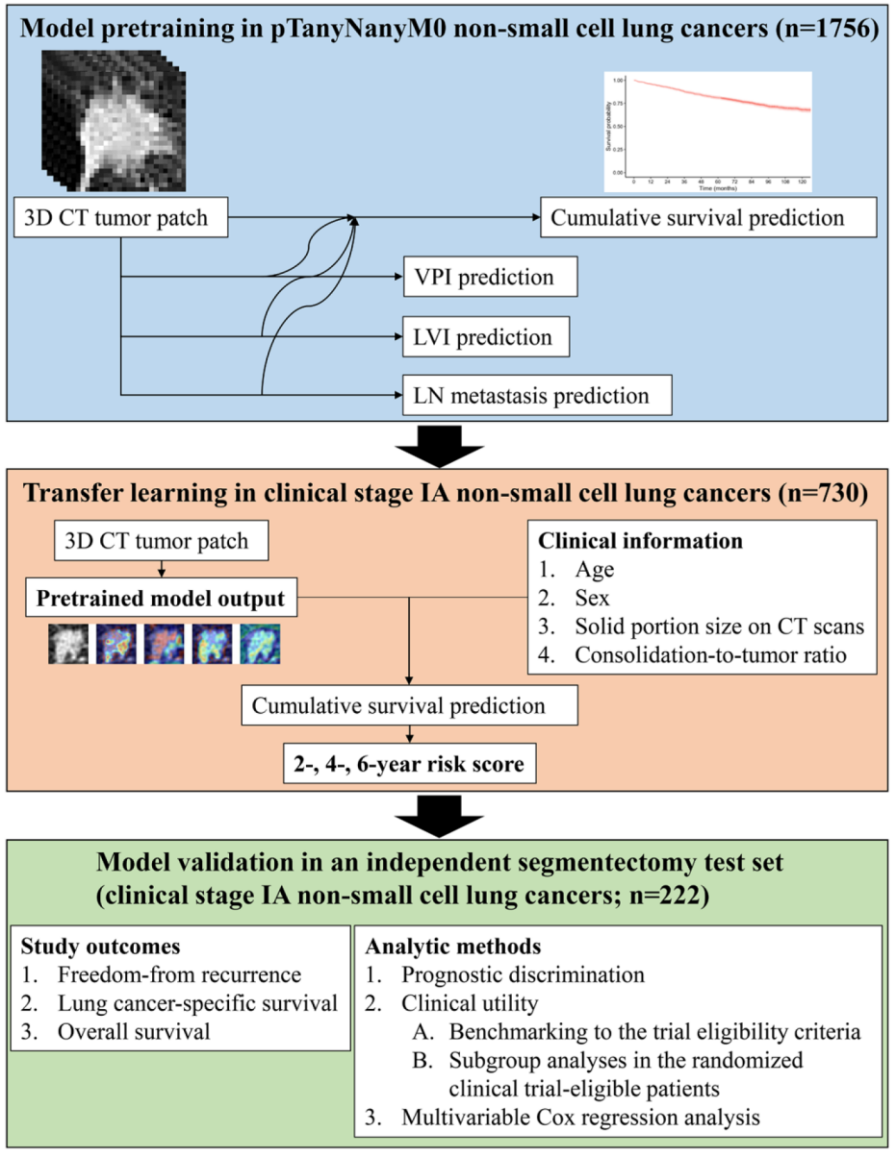
From the Department of Thoracic and Cardiovascular Surgery (K.J.N., Y.T.K.) and Department of Radiology (J.M.G., H.K.), Seoul National University Hospital and College of Medicine, 101 Daehak-ro, Jongno-gu, Seoul 03080, Korea; Seoul National University Cancer Research Institute, Seoul National University College of Medicine, Seoul, Korea (K.J.N., Y.T.K., J.M.G.); and Institute of Radiation Medicine, Seoul National University Medical Research Center, Seoul, Korea (J.M.G.). Received July 12, 2023; revision requested August 11; revision received February 27, 2024; accepted March 1. **Address correspondence to** H.K. (email: *khj.snuh@gmail.com*).

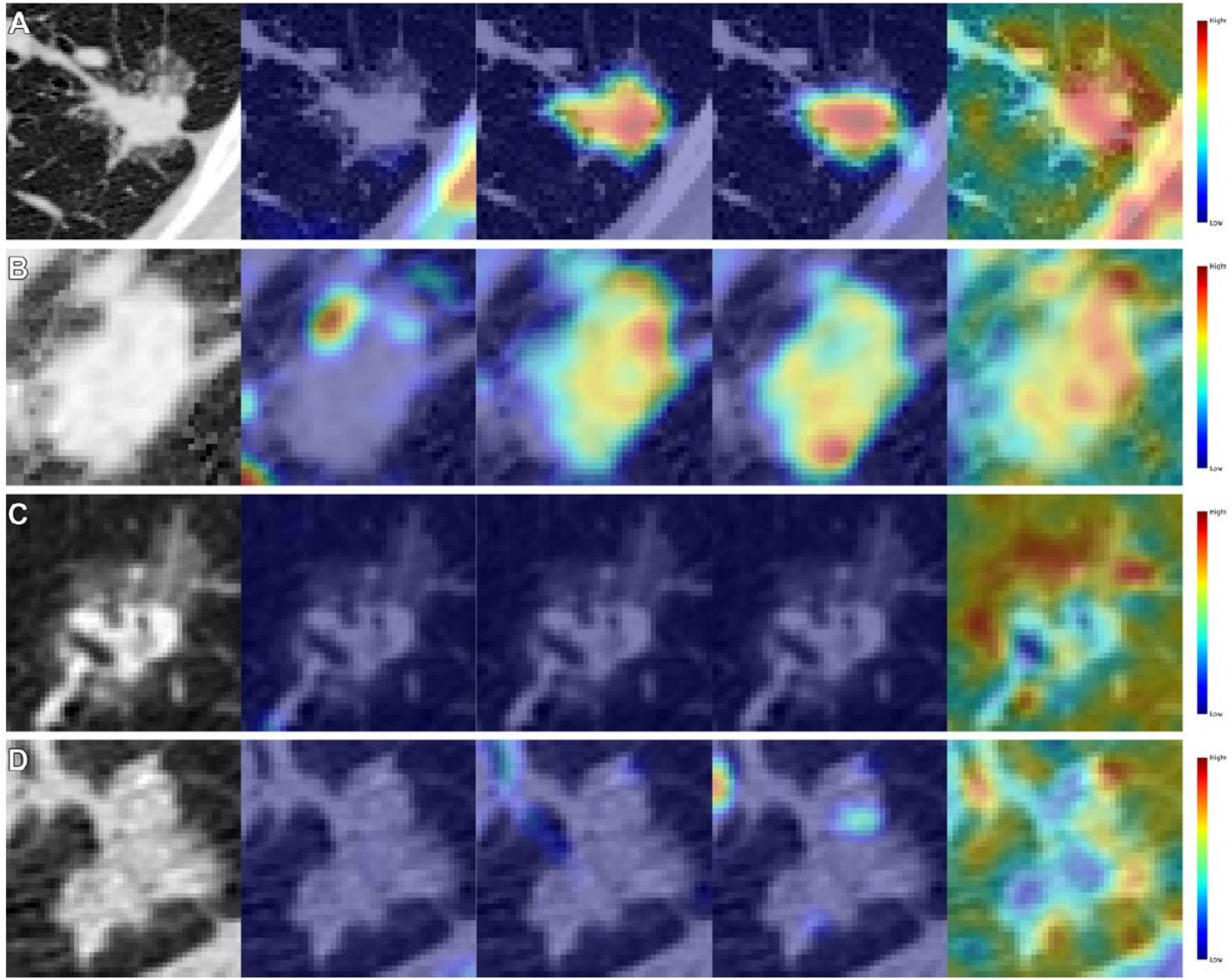
Supported by the Seoul National University Hospital Research Fund (03-2022-2170) and a National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2023-00207978). However, the funders had no role in the study design; in the collection, analysis, and interpretation of the data; in the writing of the report; and in the decision to submit the article for publication.

Conflicts of interest are listed at the end of this article.

Radiology 2024; 311(1):e231793 • https://doi.org/10.1148/radiol.231793 • Content codes: CH AI CT

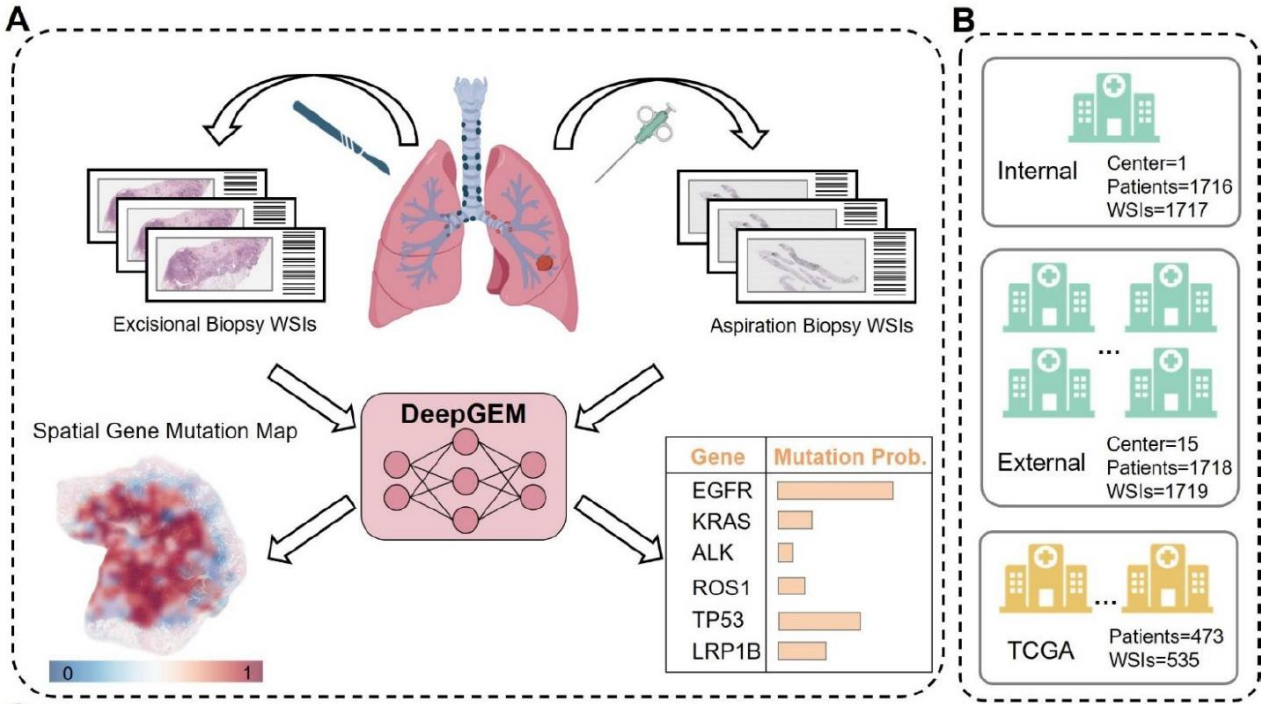
Clinical Research: Radiology



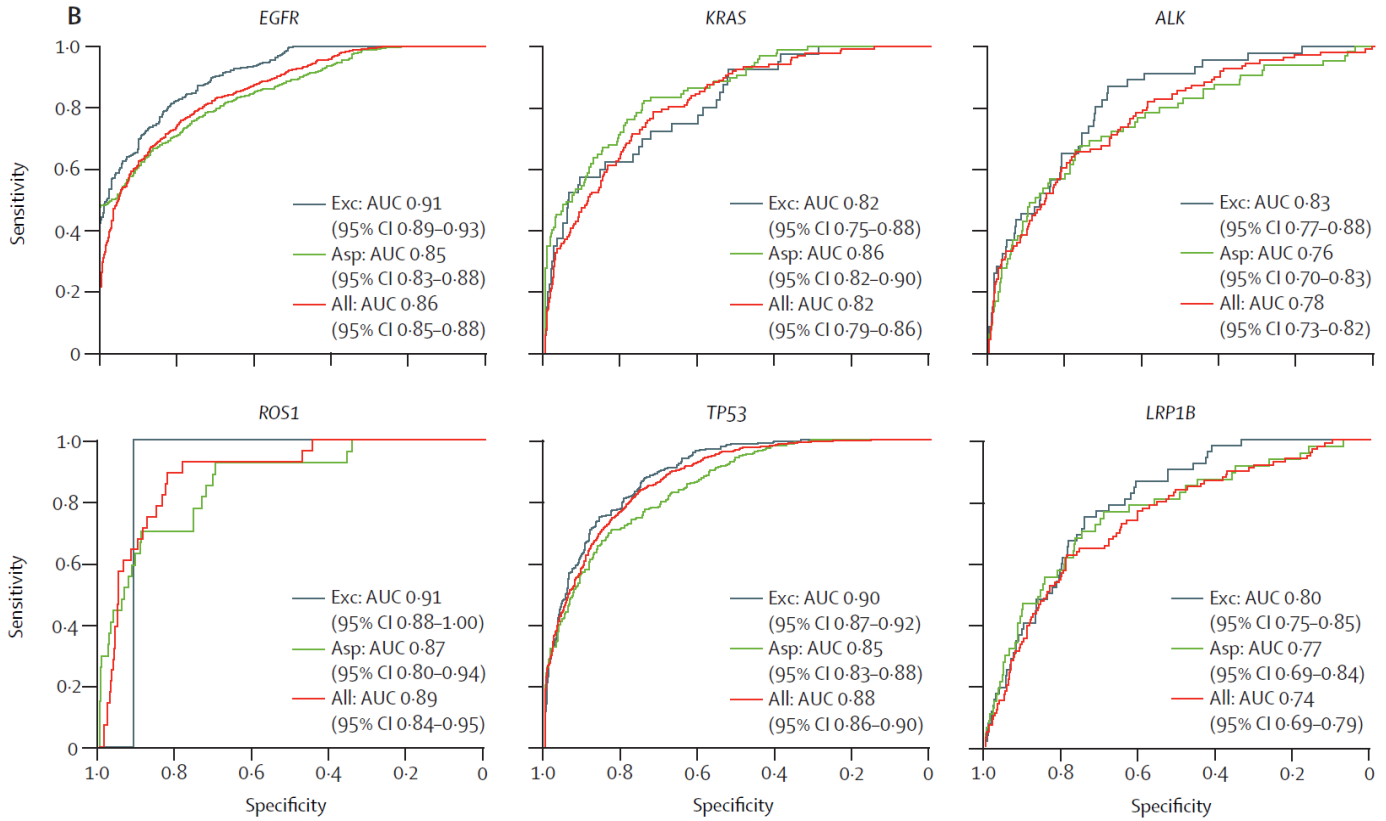


Na et al., *Radiology*, 2024

Clinical Research: Pathology

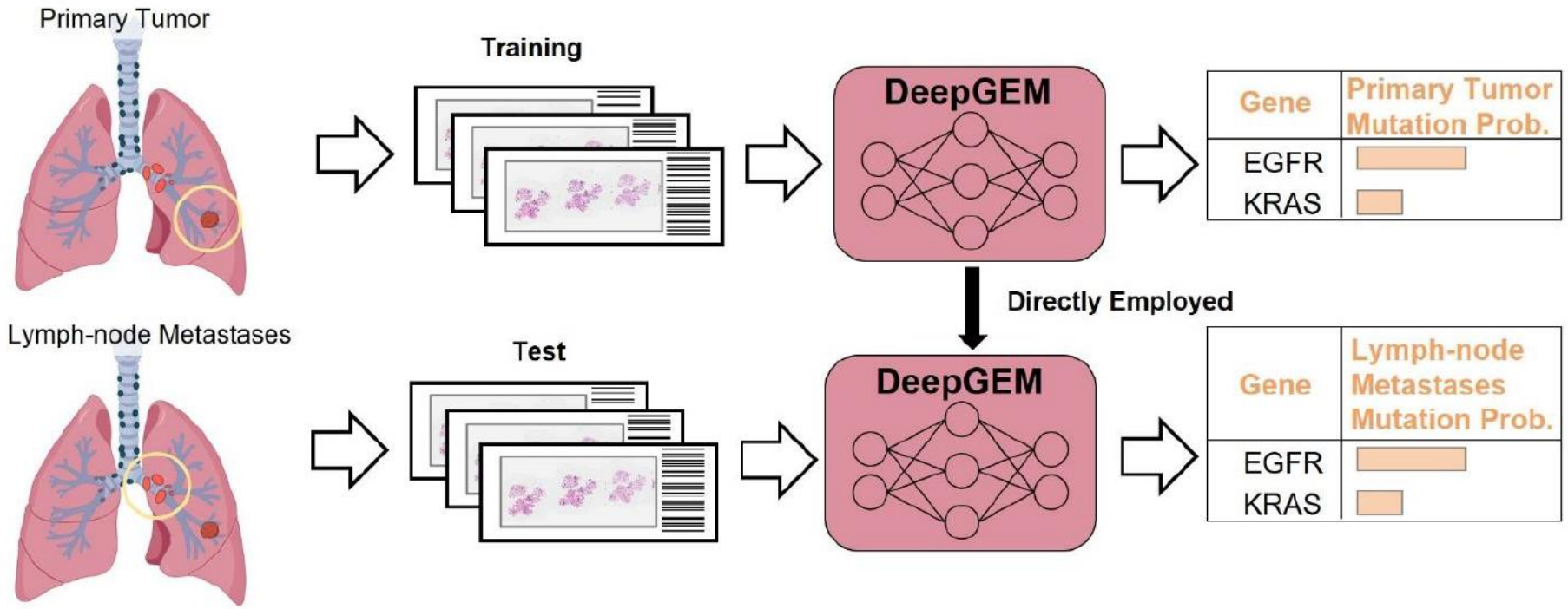


Clinical Research: Pathology



Zao et al., *Lancet Oncol*, 2025

Clinical Research: Pathology



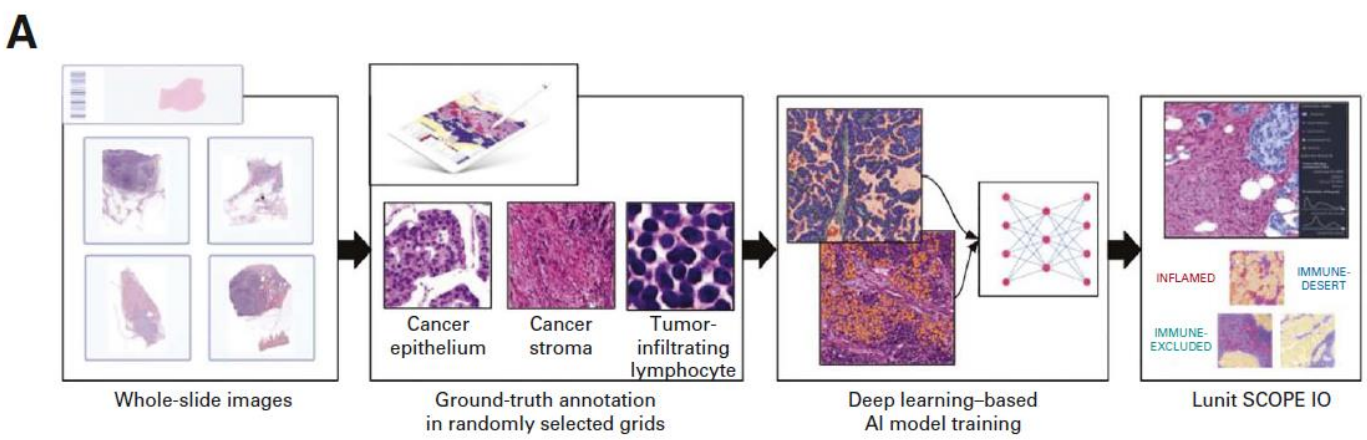
Zao et al., *Lancet Oncol*, 2025

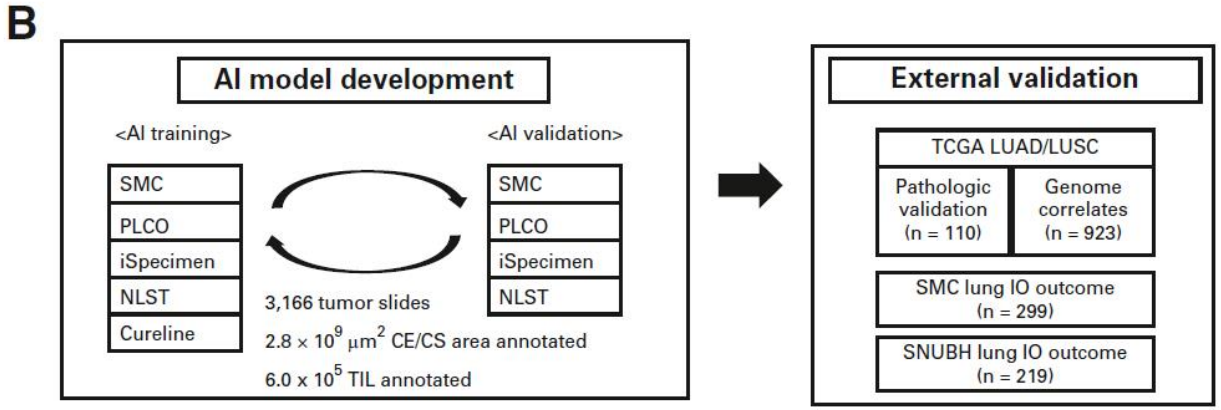
original reports

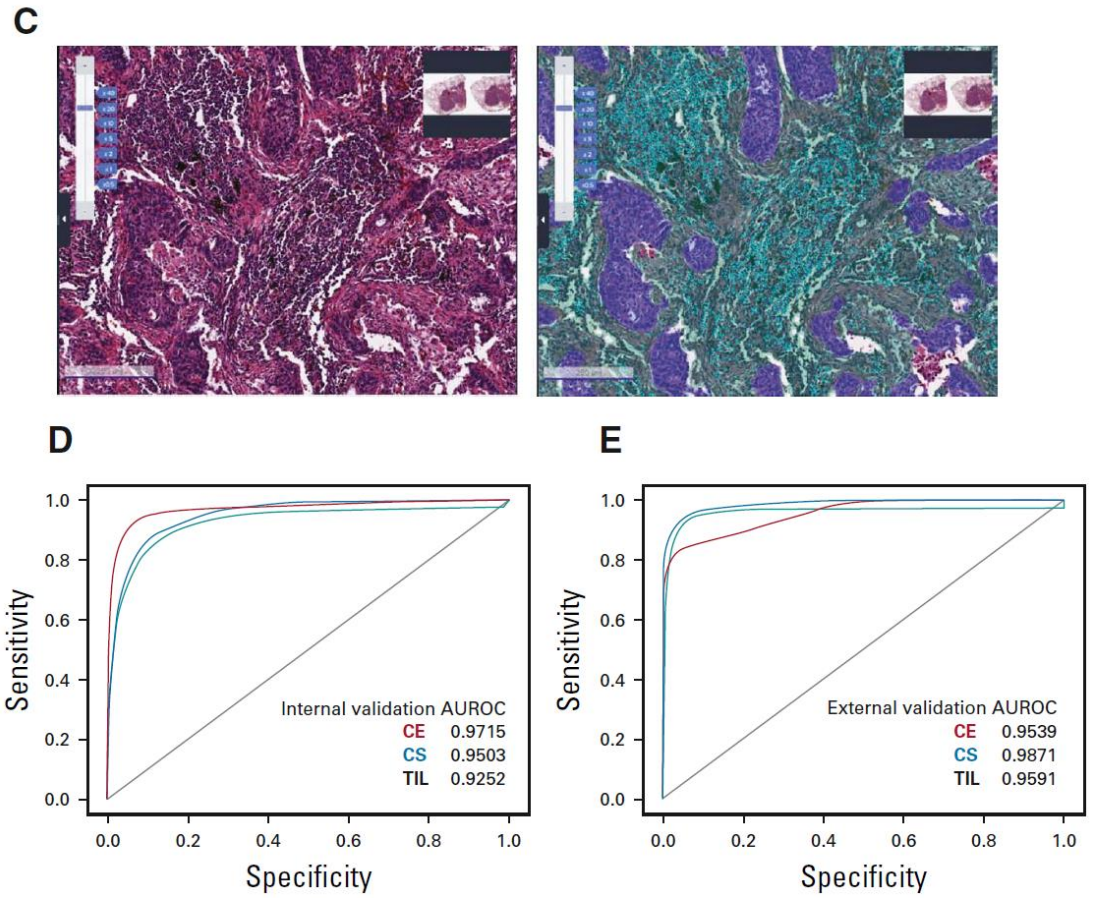
Artificial Intelligence–Powered Spatial Analysis of Tumor-Infiltrating Lymphocytes as Complementary Biomarker for Immune Checkpoint Inhibition in Non–Small-Cell Lung Cancer

Sehhoon Park, MD, PhD[1]; Chan-Young Ock, MD, PhD[2]; Hyojin Kim, MD, PhD[3]; Sergio Pereira, PhD[2]; Seonwook Park, PhD[2]; Minuk Ma, MS[2]; Sangjoon Choi, MD[4]; Seokhwi Kim, MD, PhD[5]; Seunghwan Shin, MD[2]; Brian Jaehong Aum, PhD[2]; Kyunghyun Paeng, MS[2]; Donggeun Yoo, PhD[2]; Hongui Cha, PhD[1]; Sunyoung Park, PhD[1]; Koung Jin Suh, MD[6]; Hyun Ae Jung, MD, PhD[1]; Se Hyun Kim, MD, PhD[6]; Yu Jung Kim, MD, PhD[6]; Jong-Mu Sun, MD, PhD[1]; Jin-Haeng Chung, MD, PhD[3]; Jin Seok Ahn, MD, PhD[1]; Myung-Ju Ahn, MD, PhD[1]; Jong Seok Lee, MD, PhD[6]; Keunchil Park, MD, PhD[1]; Sang Yong Song, MD, PhD[4]; Yung-Jue Bang, MD, PhD[7]; Yoon-La Choi, MD, PhD[4]; Tony S. Mok, MD[8]; and Se-Hoon Lee, MD, PhD[1,9]

Clinical Research: Pathology







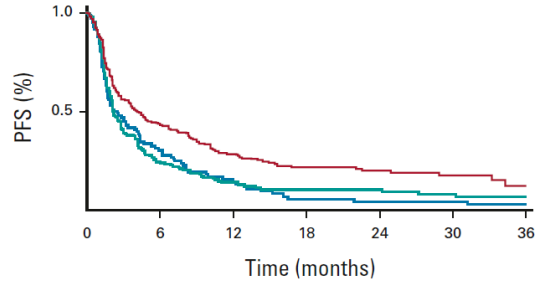
Park et al., *J Clin Oncol*, 2022

Clinical Research: Pathology

B

ICI Monotherapy

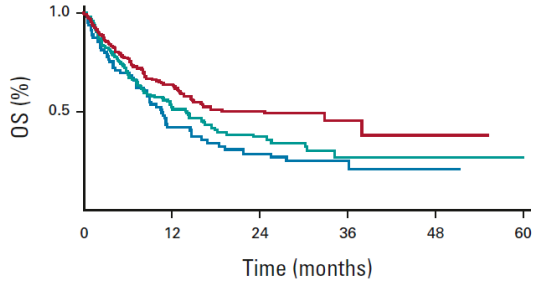
PFS	No.	ORR (%)	mPFS (95% CI)	HR (95% CI)	*P*
Inflamed	228	26.8	4.1 (2.8 to 6.2)	NA	NA
Immune-excluded	192	11.5	2.2 (2.0 to 2.8)	1.52 (1.23 to 1.88)	9.6×10^{-5}
Immune-desert	98	11.2	2.4 (1.7 to 4.2)	1.58 (1.23 to 2.03)	4.1×10^{-4}



No. at risk:							
Inflamed	228	95	53	34	23	12	1
Immune-excluded	192	46	22	15	11	6	2
Immune-desert	98	28	13	4	3	3	1

ICI Monotherapy

OS	mOS (95% CI)	HR (95% CI)	*P*
Inflamed	24.8 (14.6 to NR)	NA	NA
Immune-excluded	14.0 (10.8 to 17.9)	1.38 (1.05 to 1.83)	.023
Immune-desert	10.6 (8.6 to 16.1)	1.67 (1.21 to 2.31)	.002

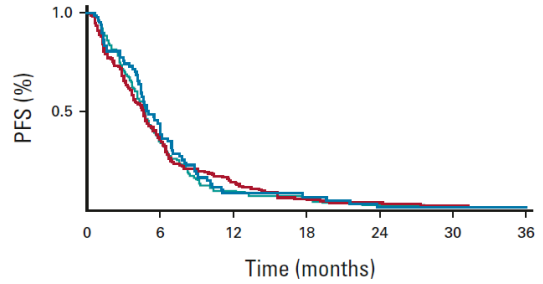


No. at risk:							
Inflamed	228	99	43	7	1	0	
Immune-excluded	192	66	28	8	2	1	
Immune-desert	98	29	16	6	2	0	

C

Platinum-Based Chemotherapy

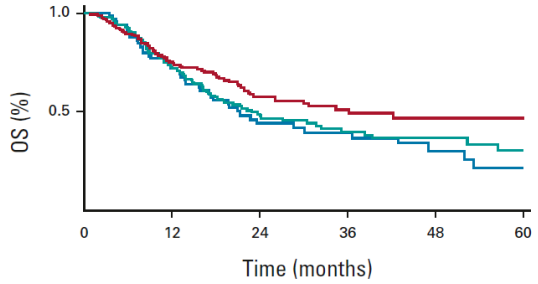
PFS	No.	ORR (%)	mPFS (95% CI)	HR (95% CI)	*P*
Inflamed	162	33.3	4.6 (3.8 to 5.4)	NA	NA
Immune-excluded	144	30.1	4.7 (4.1 to 5.3)	1.00 (0.79 to 1.26)	.998
Immune-desert	67	44.6	5.1 (4.5 to 6.6)	0.94 (0.70 to 1.26)	.670



No. at risk:							
Inflamed	162	56	22	8	6	4	3
Immune-excluded	144	49	14	7	3	2	2
Immune-desert	67	27	5	4	1	1	1

Platinum-Based Chemotherapy

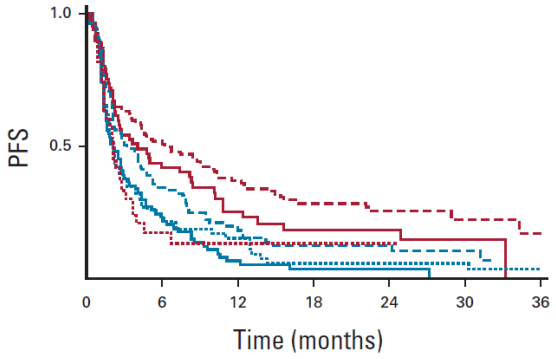
OS	mOS (95% CI)	HR (95% CI)	*P*
Inflamed	36.2 (23.1 to NR)	NA	NA
Immune-excluded	22.9 (17.9 to 38.4)	1.31 (0.94 to 1.81)	.108
Immune-desert	20.9 (15.8 to 47.0)	1.47 (1.00 to 2.17)	.049



No. at risk:							
Inflamed	162	107	55	28	10	7	
Immune-excluded	144	93	47	24	14	10	
Immune-desert	67	46	22	15	8	3	

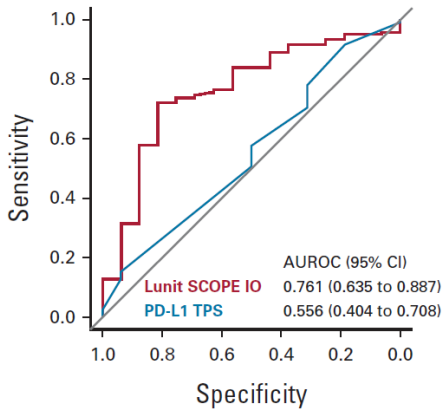
C

Subgroup by PD-L1 TPS	N	ORR (%)	mPFS (95% CI)	HR (95% CI)	*P*
TPS ≥ 50%, inflamed	117	36.8	6.2 (4.3 to 10.3)	0.63 (0.46 to 0.87)	.004
TPS ≥ 50%, noninflamed	89	20.2	3.2 (2.1 to 5.3)		
TPS 1%-49%, inflamed	57	22.8	4.0 (2.3 to 8.4)	0.54 (0.37 to 0.79)	.001
TPS 1%-49%, noninflamed	77	3.9	2.1 (1.6 to 3.3)		
TPS 0%, inflamed	33	3.0	2.1 (1.5 to 3.7)	1.03 (0.66 to 1.61)	.891
TPS 0%, noninflamed	71	5.6	2.1 (1.7 to 3.2)		



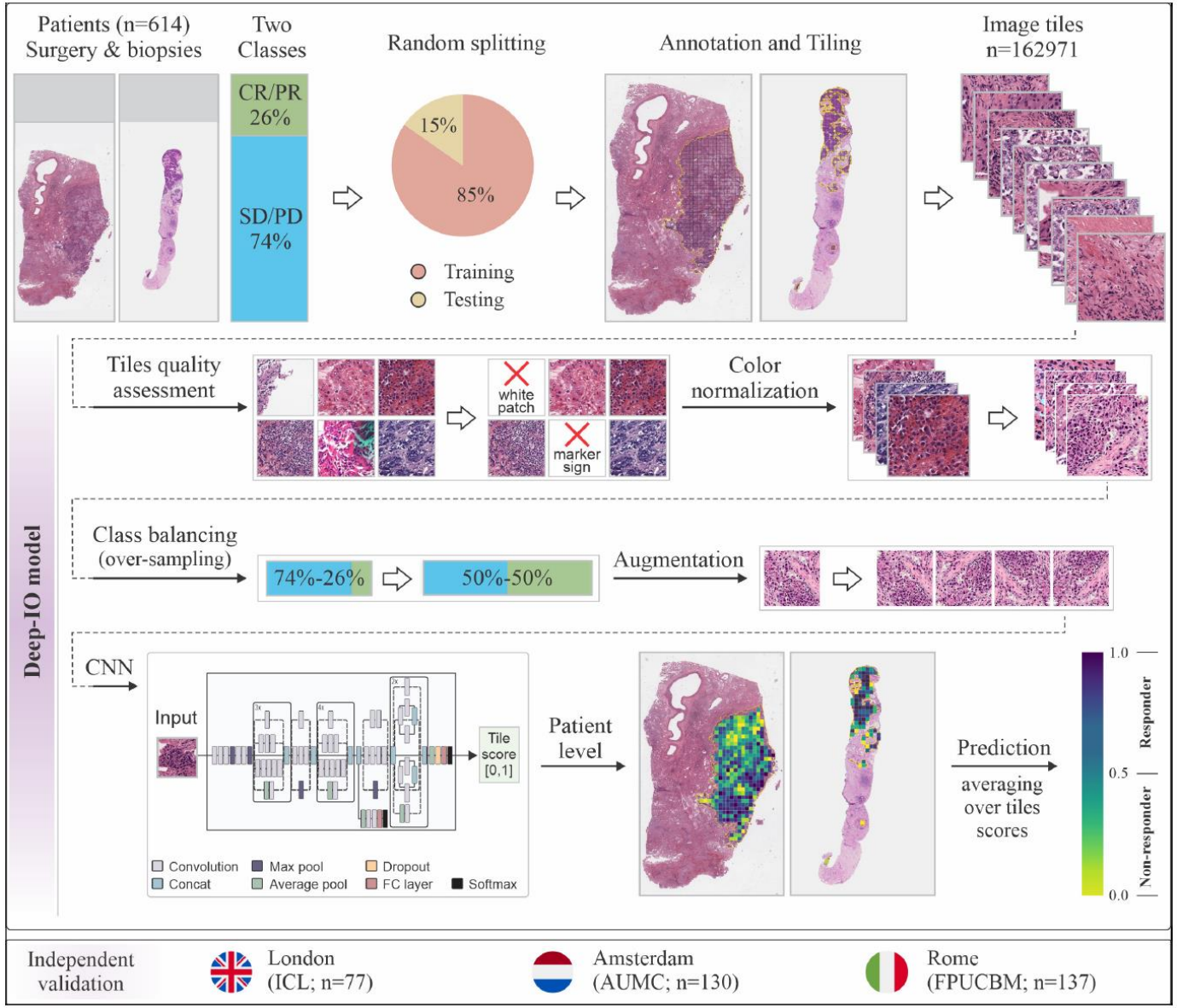
No. at risk:							
TPS ≥ 50%, inflamed	117	58	37	23	16	7	1
TPS ≥ 50%, noninflamed	89	30	14	9	6	3	0
TPS1%-49%, inflamed	57	24	11	8	5	4	0
TPS1%-49%, noninflamed	77	17	5	2	1	0	0
TPS 0%, inflamed	33	4	3	2	1	0	0
TPS 0%, noninflamed	71	16	8	3	3	3	1

D



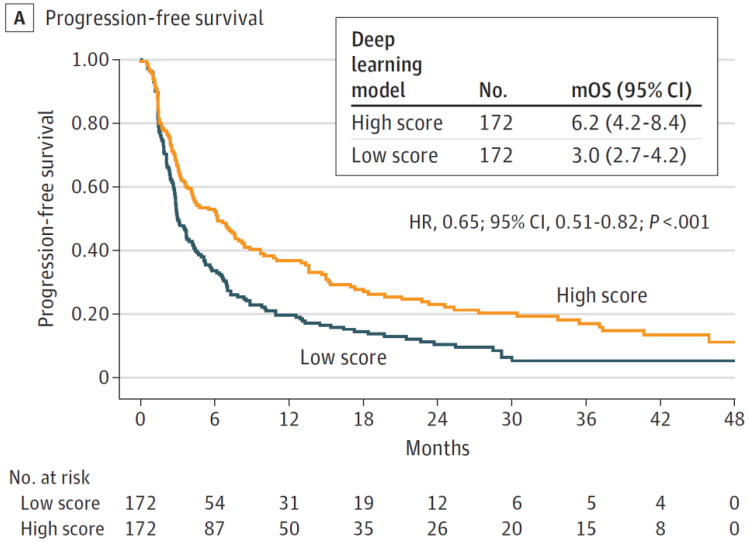
Park et al., *J Clin Oncol*, 2022

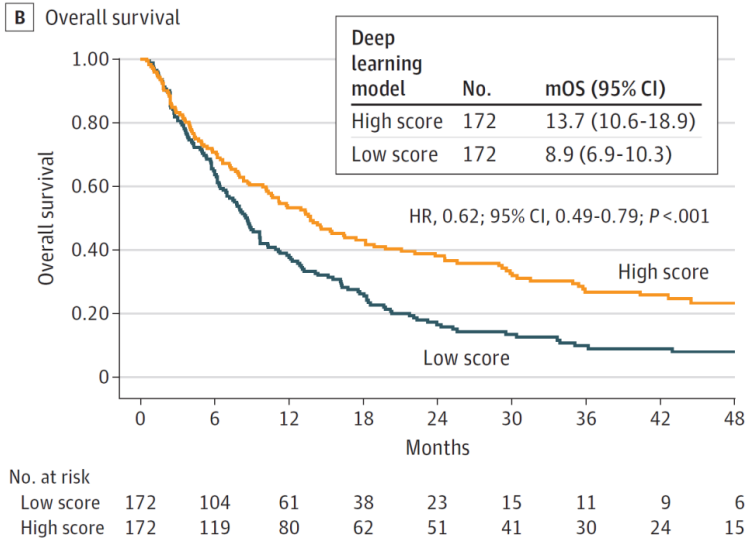
Clinical Research: Pathology

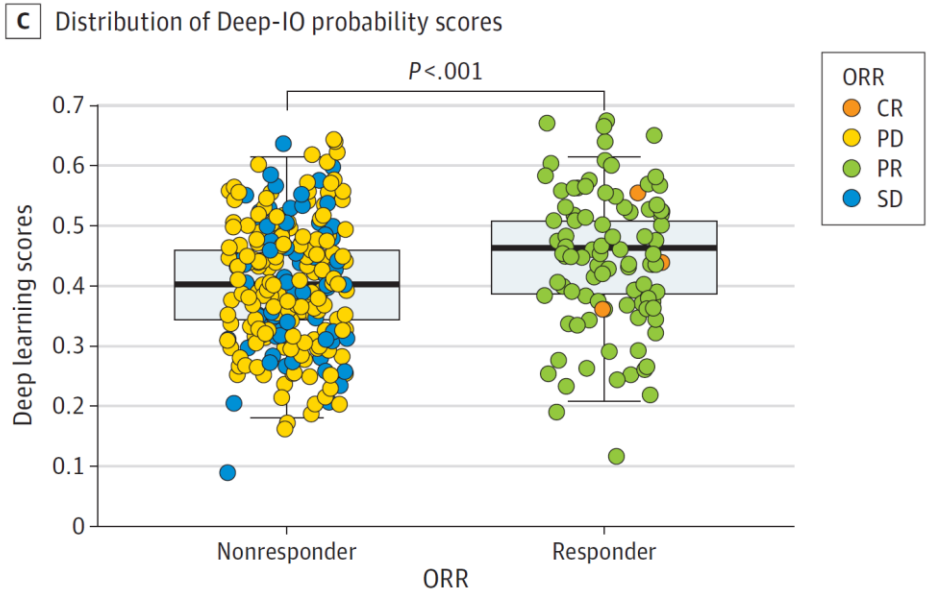


Rakaee et al., *JAMA Oncol*, 2024

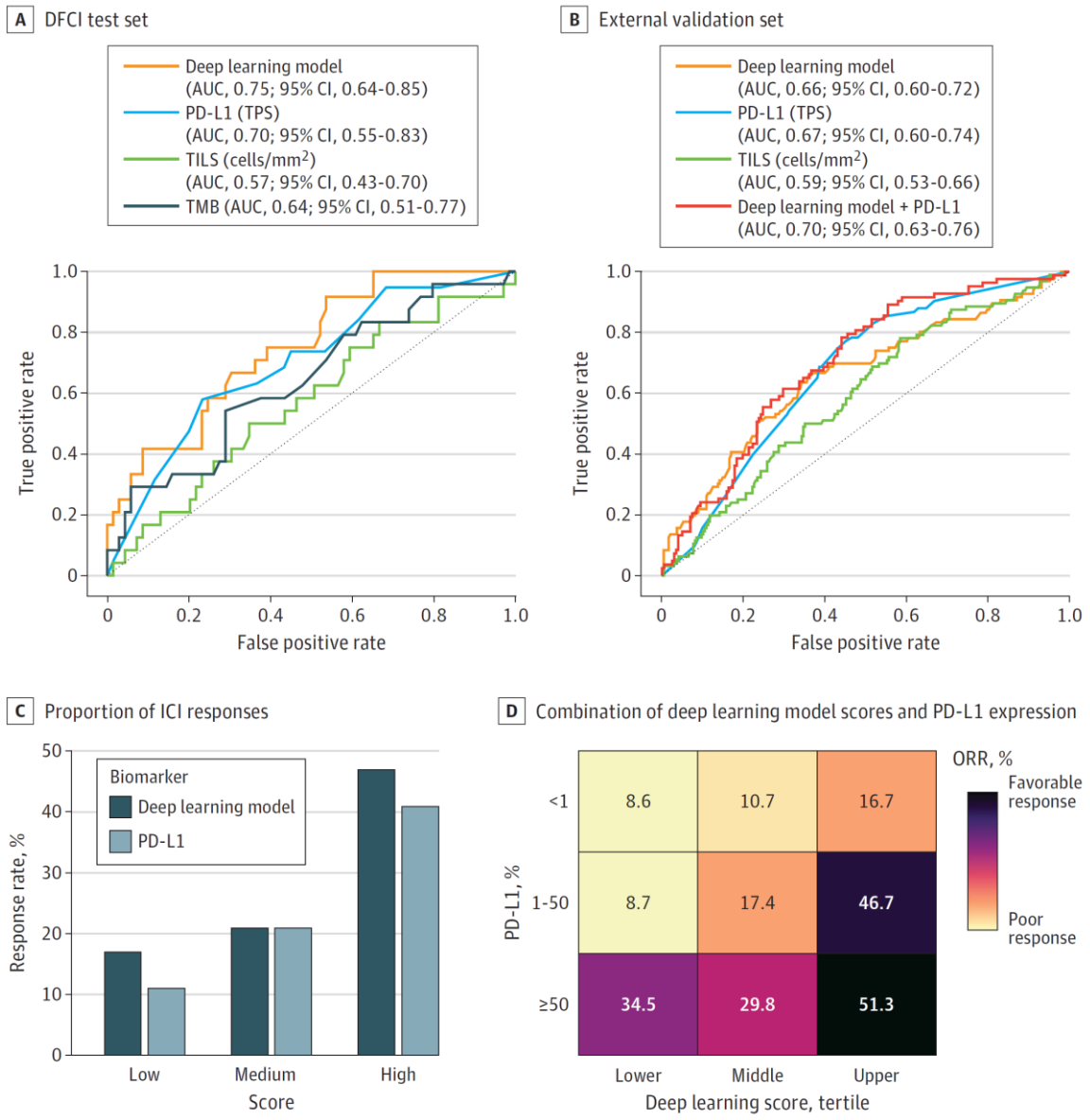
Clinical Research: Pathology





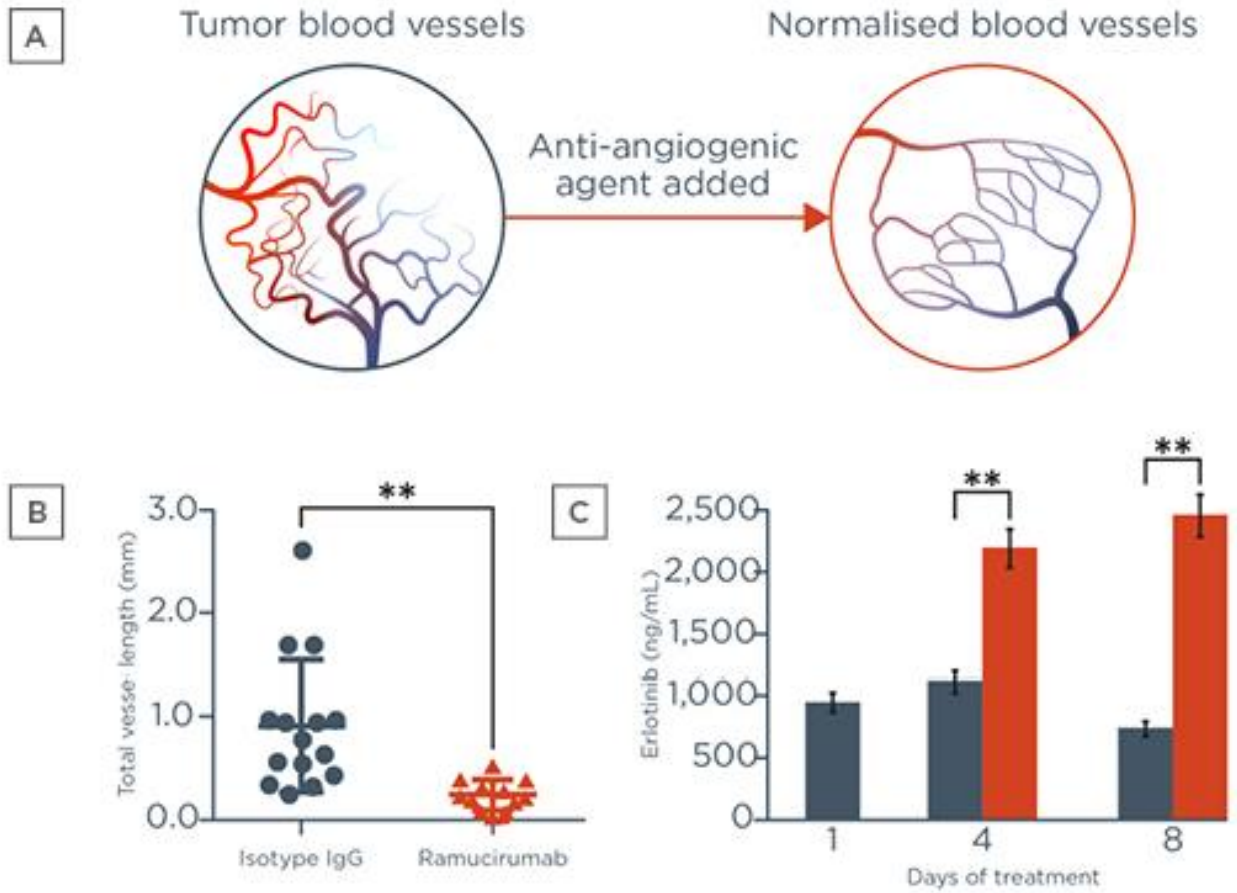


Clinical Research: Pathology



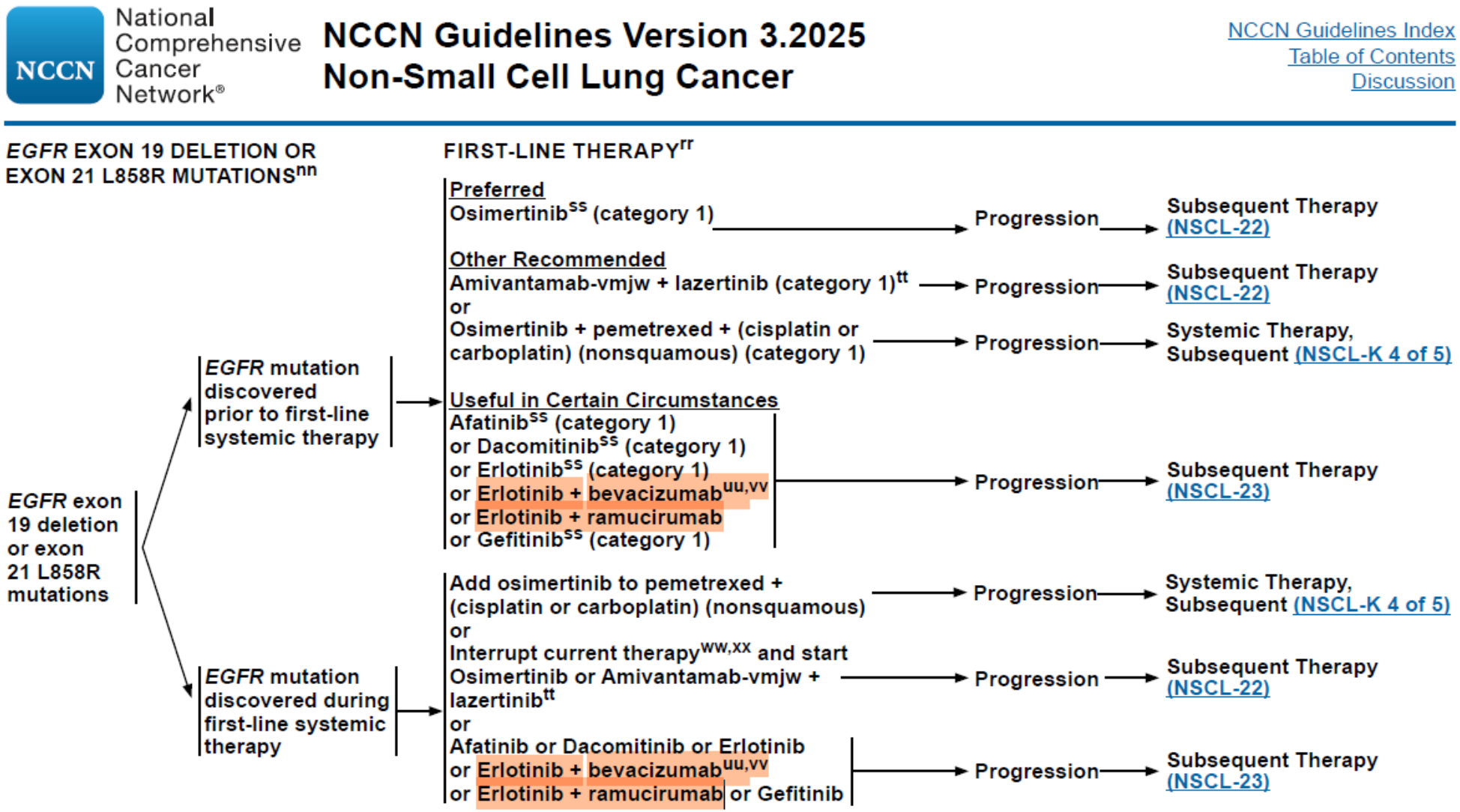
Rakaee et al., *JAMA Oncol*, 2024

Tumor vasculature

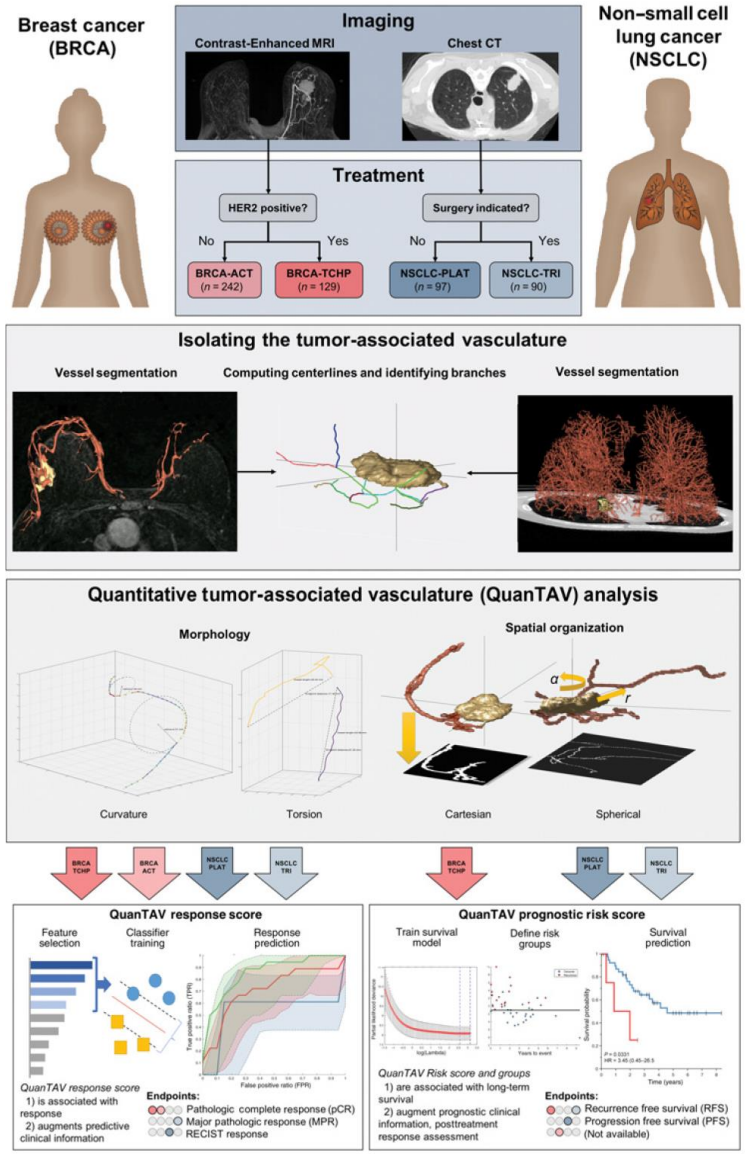


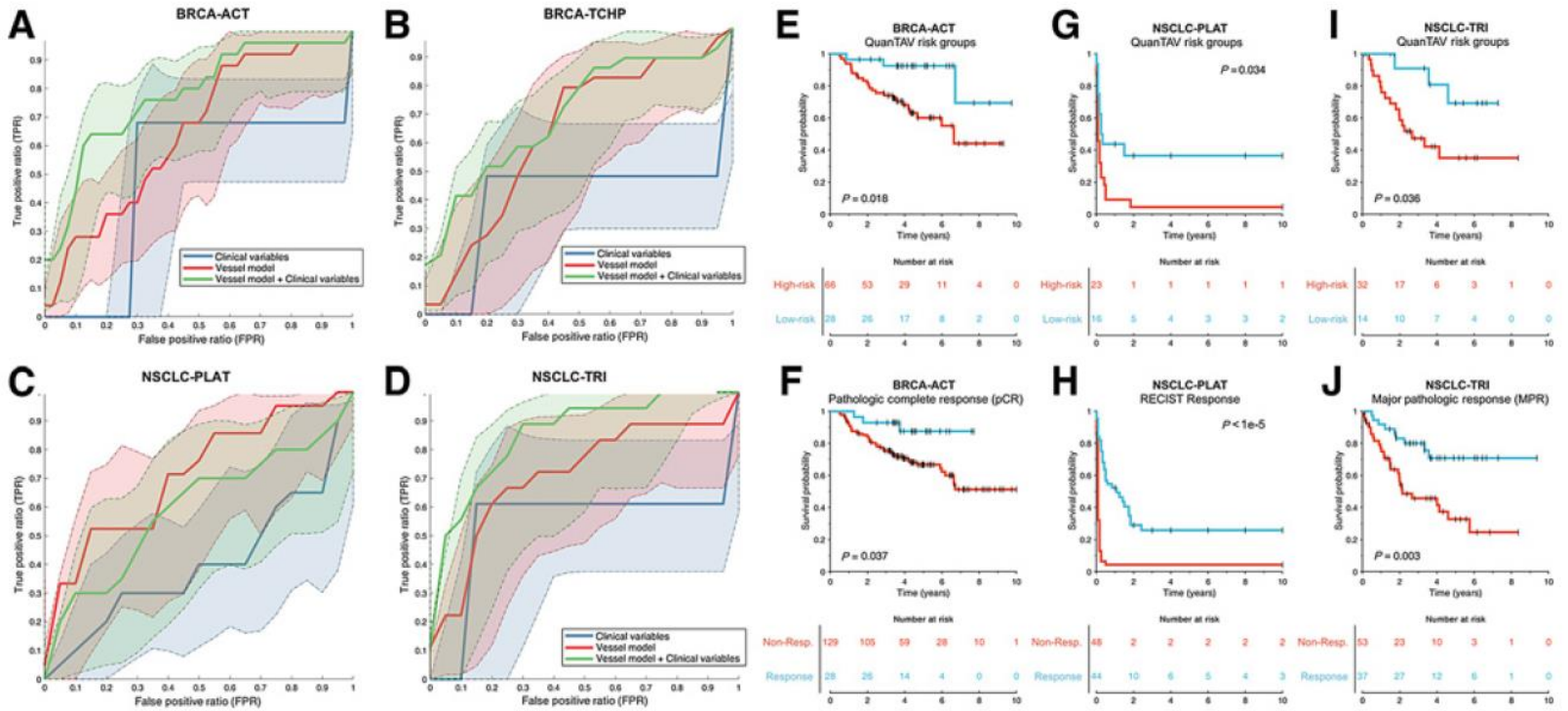
Carmeliet P et al., *Nat Rev Drug Discov*, 2011; Goel HL et al., *Nat Rev Cancer*, 2013; Chatterjee S et al., *Cancer Res*, 2014; Stylianopoulos T et al., *Trends Cancer*, 2018

Tumor vasculature

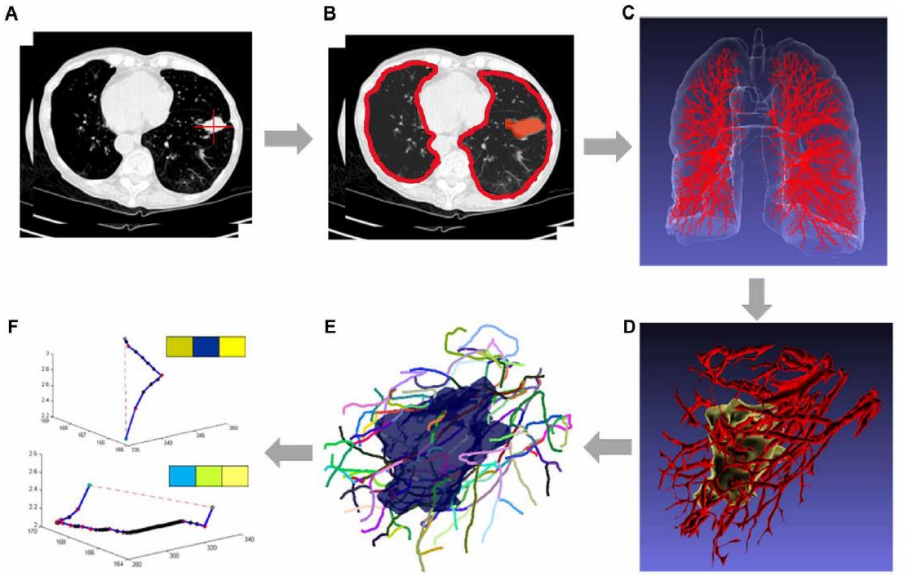


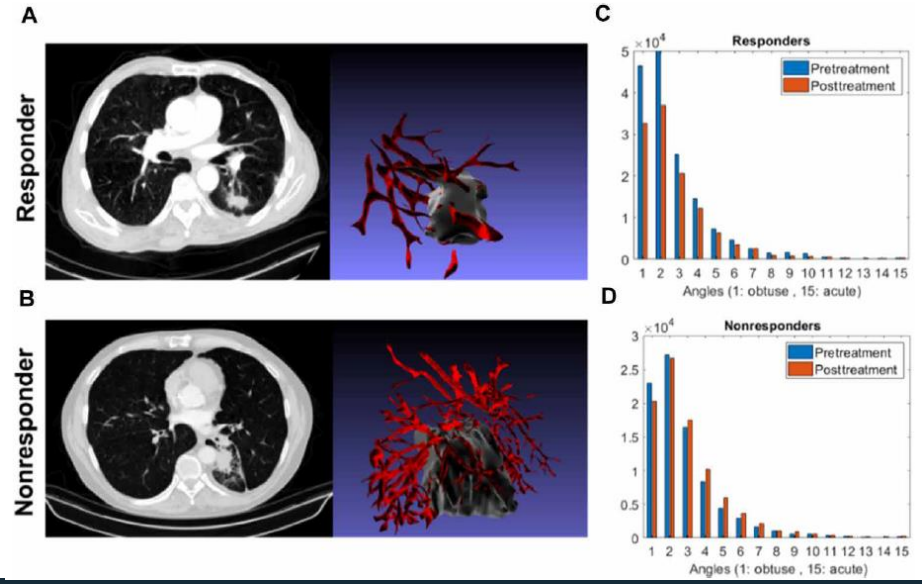
Tumor vasculature

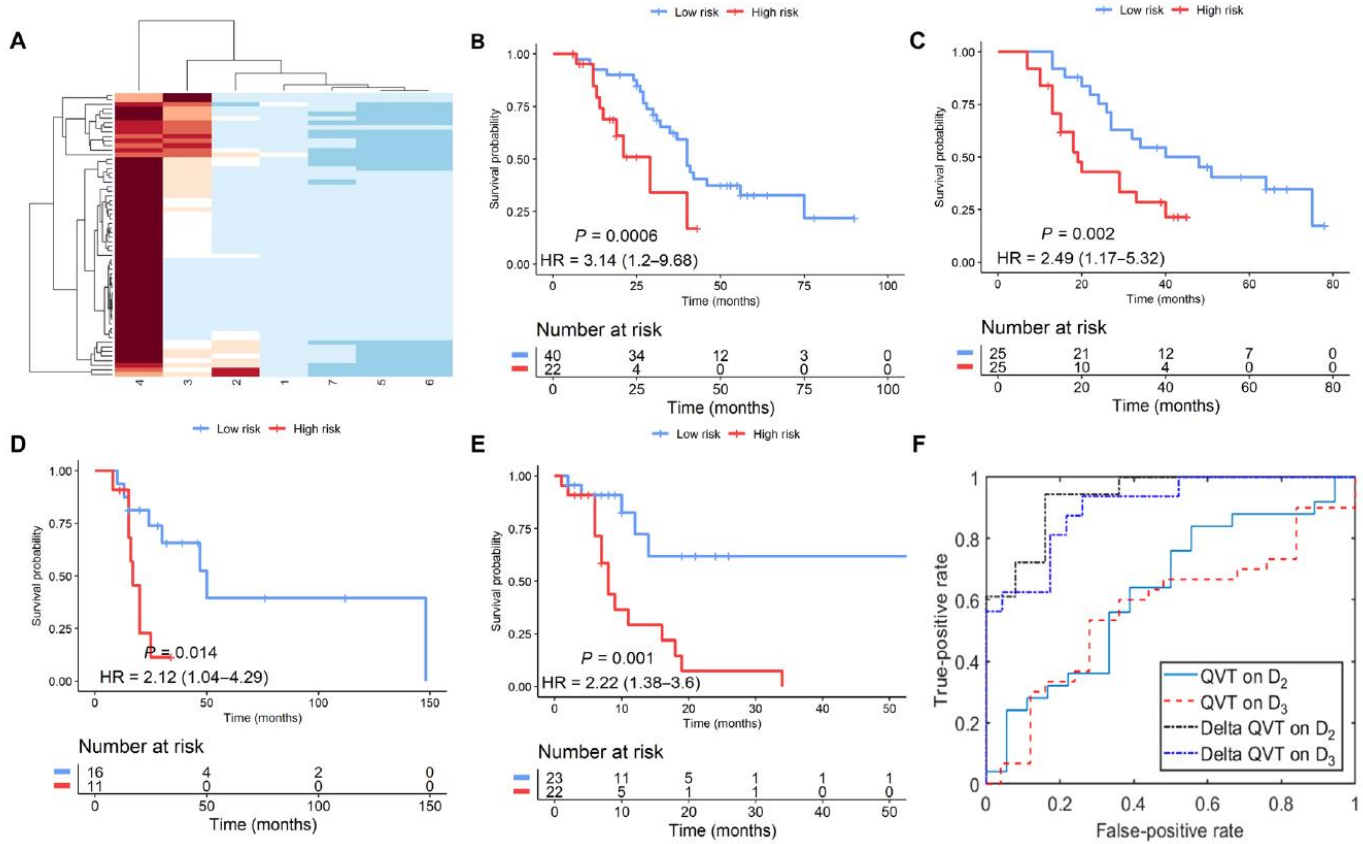




Tumor vasculature

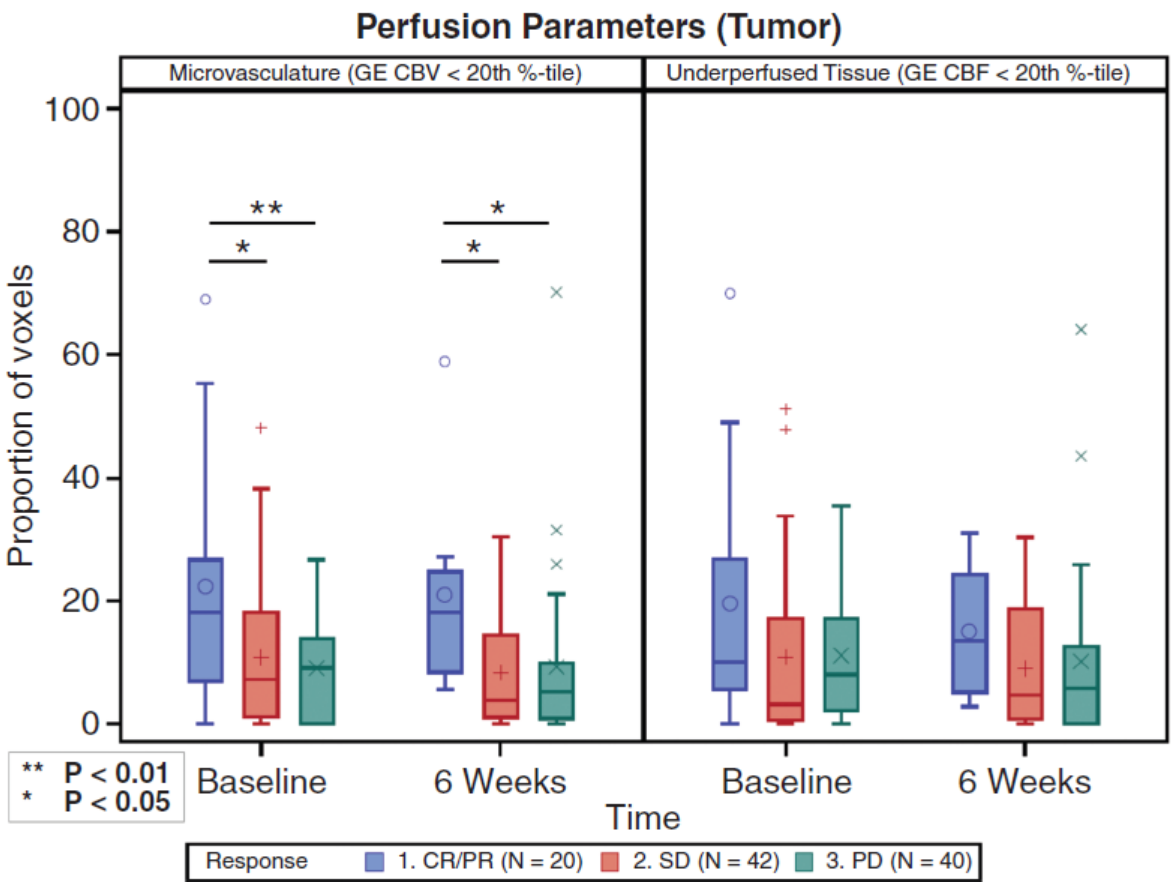


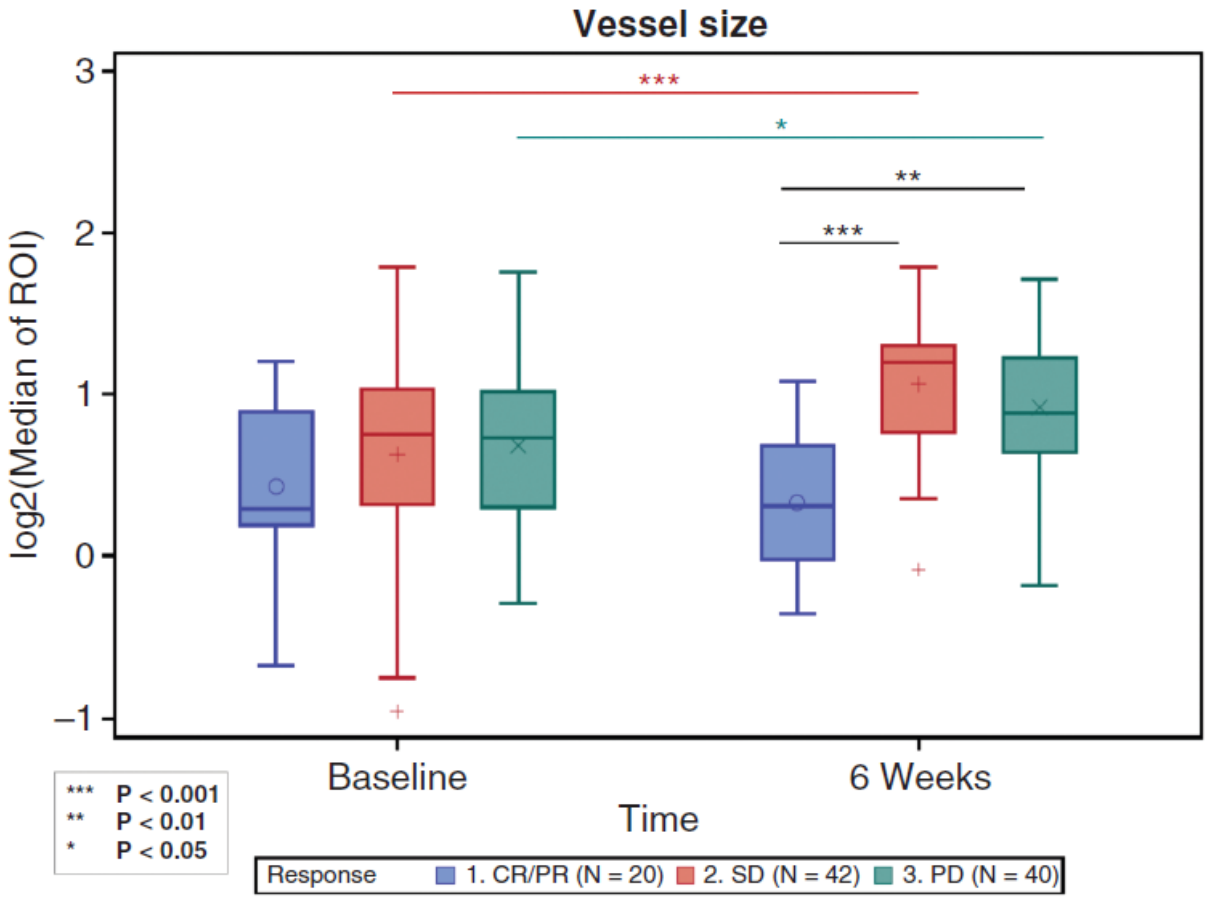




Alilou et al., *Sci Adv*, 2022

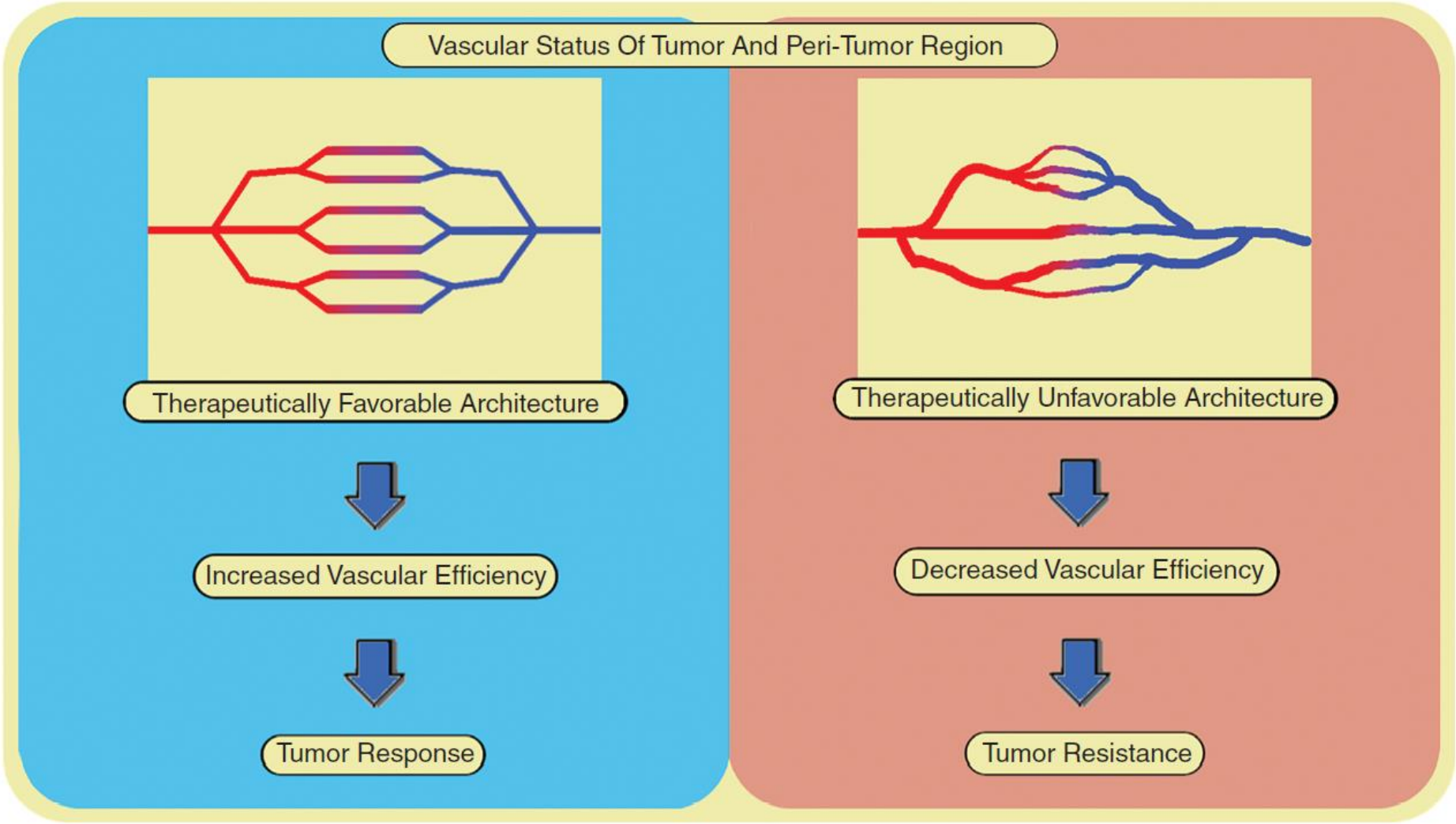
Tumor vasculature





Kim et al., *Neuro Oncol*, 2024

Tumor vasculature



Kim et al., *Neuro Oncol*, 2024

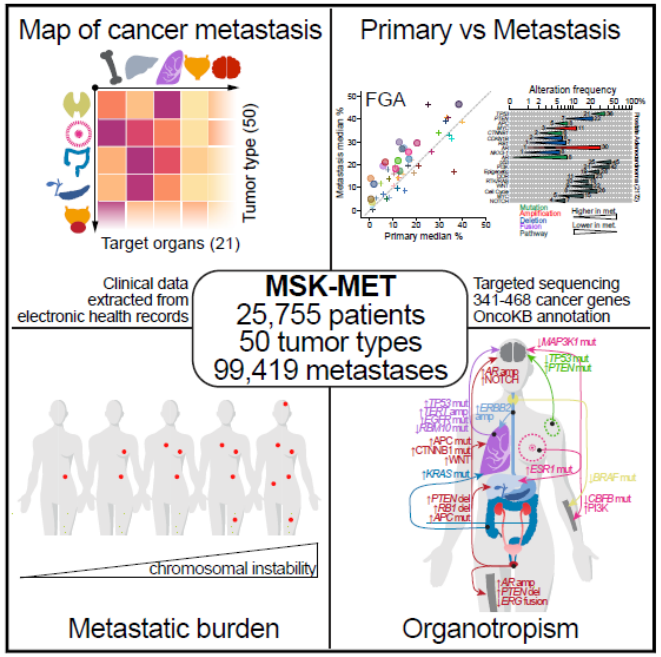
Genomic data

Cell

Resource

Genomic characterization of metastatic patterns from prospective clinical sequencing of 25,000 patients

Graphical abstract



Authors

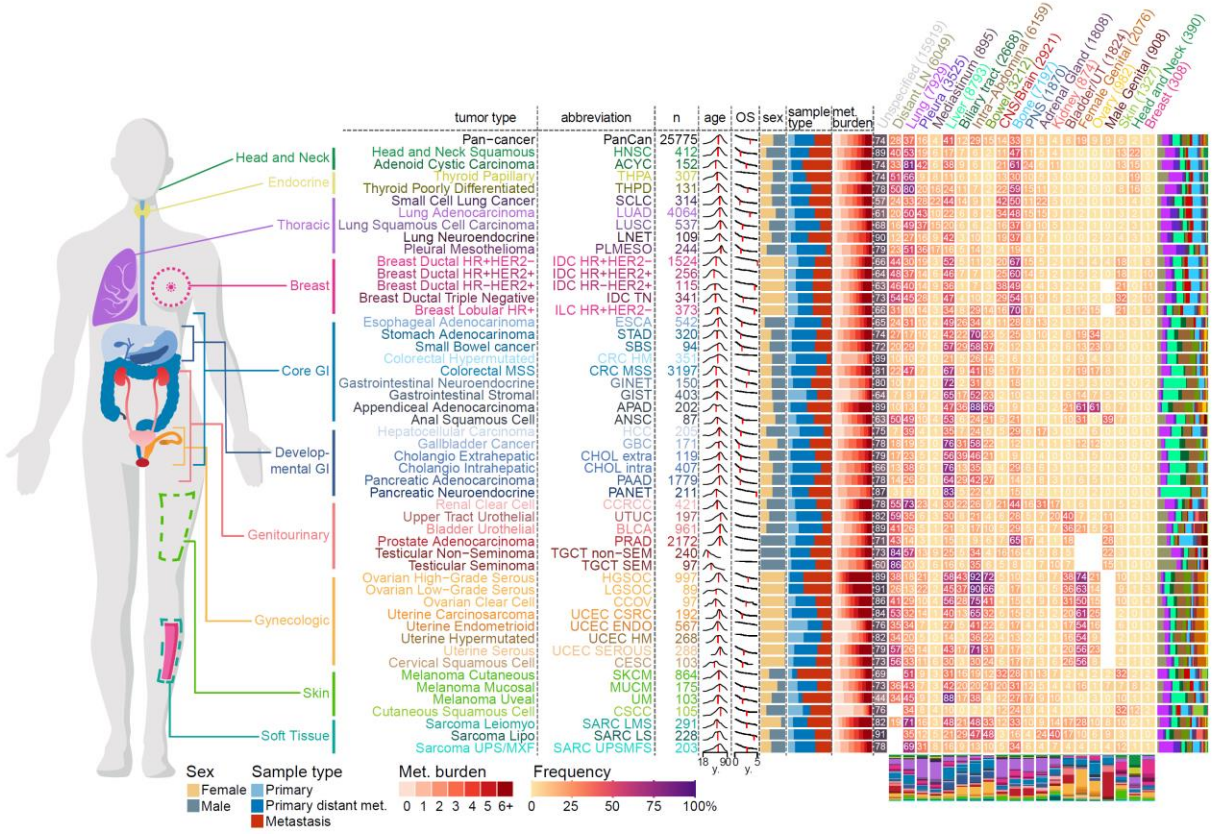
Bastien Nguyen, Christopher Fong, Anisha Luthra, ..., Samuel F. Bakhoum, Francisco Sanchez-Vega, Nikolaus Schultz

Correspondence

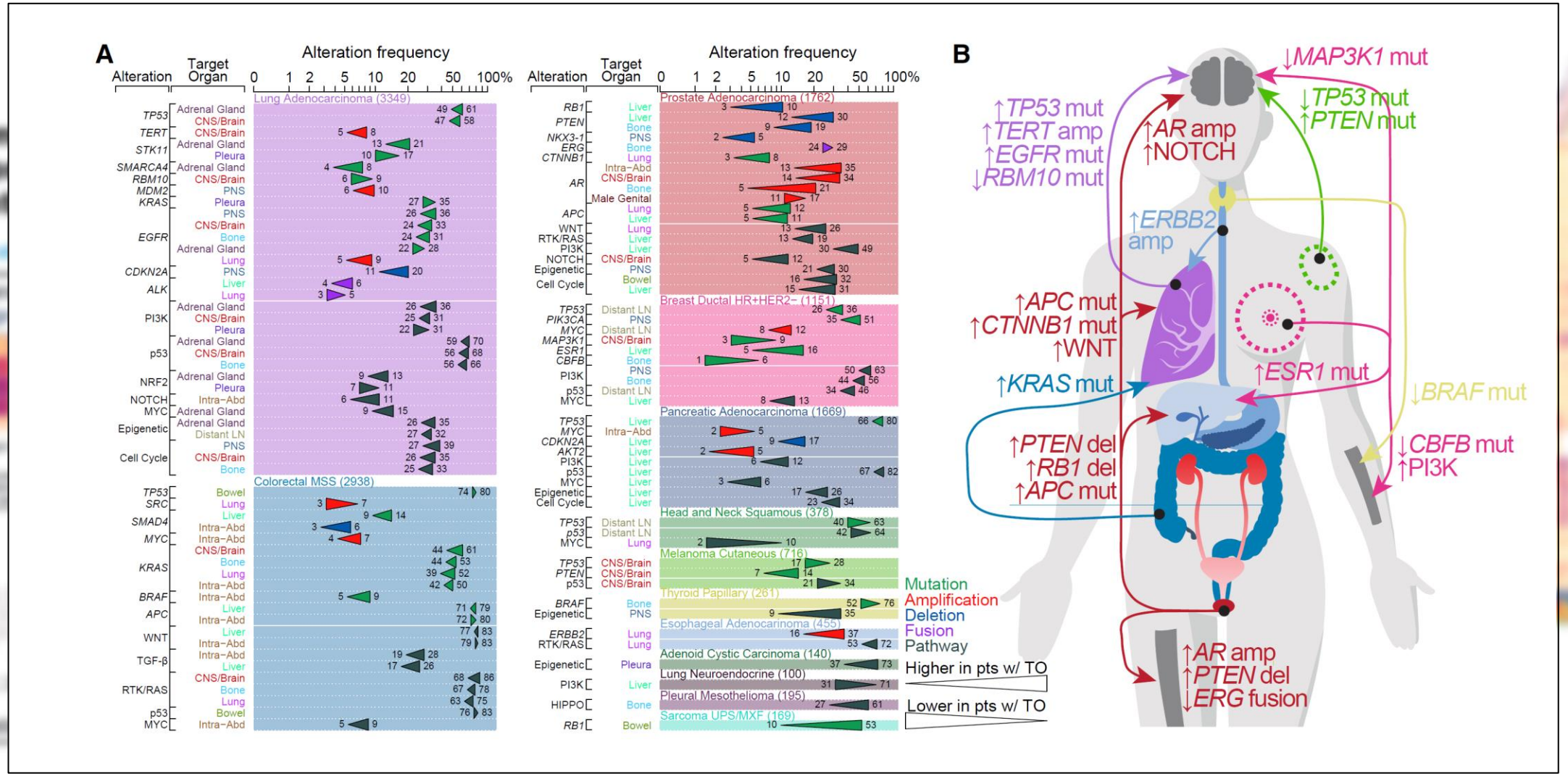
schultzn@mskcc.org (N.S.), sanchezf@mskcc.org (F.S.-V.)

In brief

Clinico-genomic analysis of MSK-MET, a cohort of over 25,000 patients with metastasis across 50 cancer types, identifies somatic alterations associated with organ-specific metastasis and highlights that chromosomal instability correlates with metastatic burden in a cancer type-dependent manner.

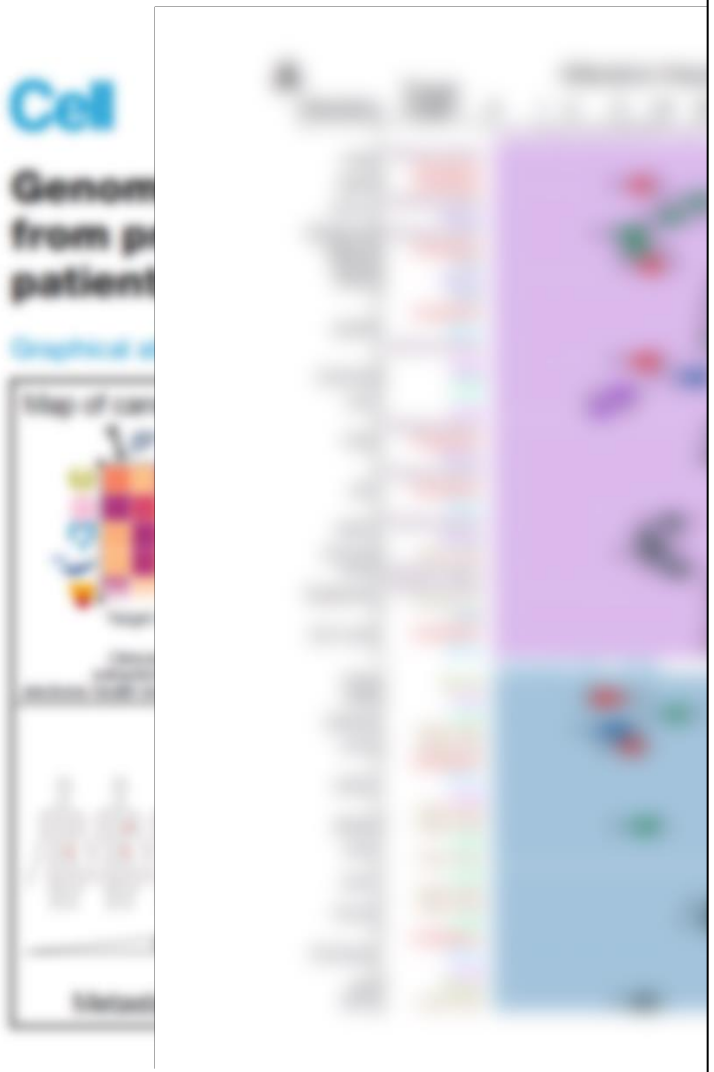


Genomic data



Nguyen et al., *Cell*, 2022

Genomic data



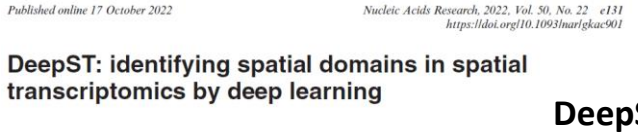
Limitations of the study

Our study has several limitations. First, while the overall cohort is large, sample size varied significantly between tumor types, which prevented us from drawing robust conclusions in less common tumor types. Therefore, the lack of significant differences in those tumor types might be due to a lack of statistical power and should be interpreted with caution. Also, our definition of tumor types could be further refined in some cases, to account, for example, for different predominant histologic subtypes in lung adenocarcinomas (Caso et al., 2020). This might provide additional valuable insights but would also result in decreased sample sizes and lower statistical power for those refined groups. Second, the ICD billing codes used in our study likely do not fully capture all metastatic events and may be affected by inter-physician variability. Future improvements to the clinical data extraction process could come from the use of natural-language processing and machine learning approaches, which will be required to mine the wealth of data contained in EHR systems at scale. Third, because of our use of a targeted sequencing panel, we may be missing biologically or clinically relevant signals that could be discovered using alternative approaches such as whole-exome or whole-genome sequencing. Finally, all analyses presented here have been performed using a single representative sample for each patient. In the future, longitudinal sampling of multiple anatomical locations at different time points from the same patient will allow us to investigate additional questions about the timing of genomic

Nguyen et al., *Cell*, 2022

Spatial transcriptomics

BayesSpace



DeepST

PASTE

SpatialGlue

stLearn

CellCharter

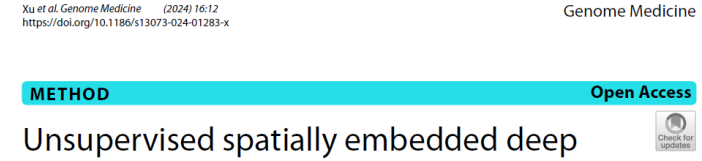
MISO

SPACEL

STAGATE

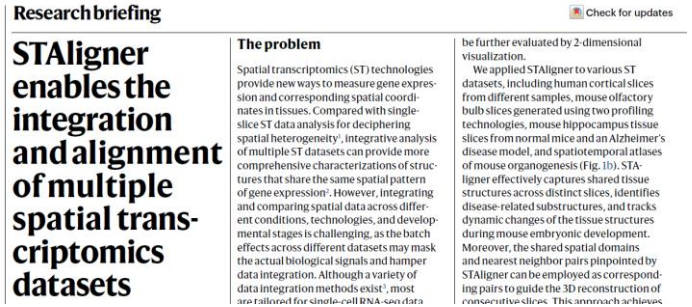
TRACERx-PHLEX

CellTrek



SEDR

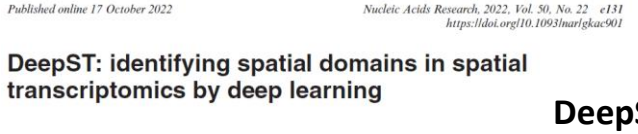
SpaGCN



STAligner

Spatial transcriptomics

BayesSpace



DeepST

PASTE

SpatialGlue

stLearn

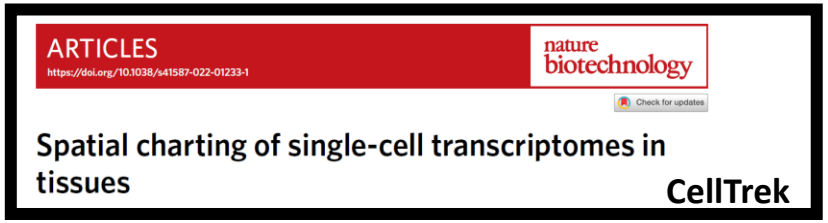
CellCharter

MISO

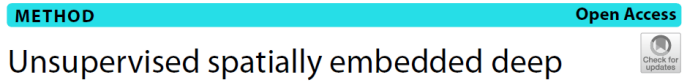
SPACEL

STAGATE

TRACERx-PHLEX

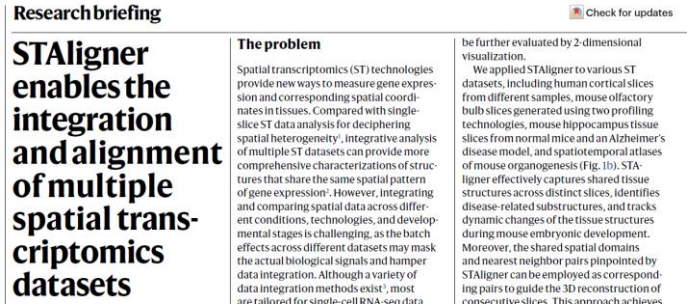


CellTrek



SEDR

SpaGCN



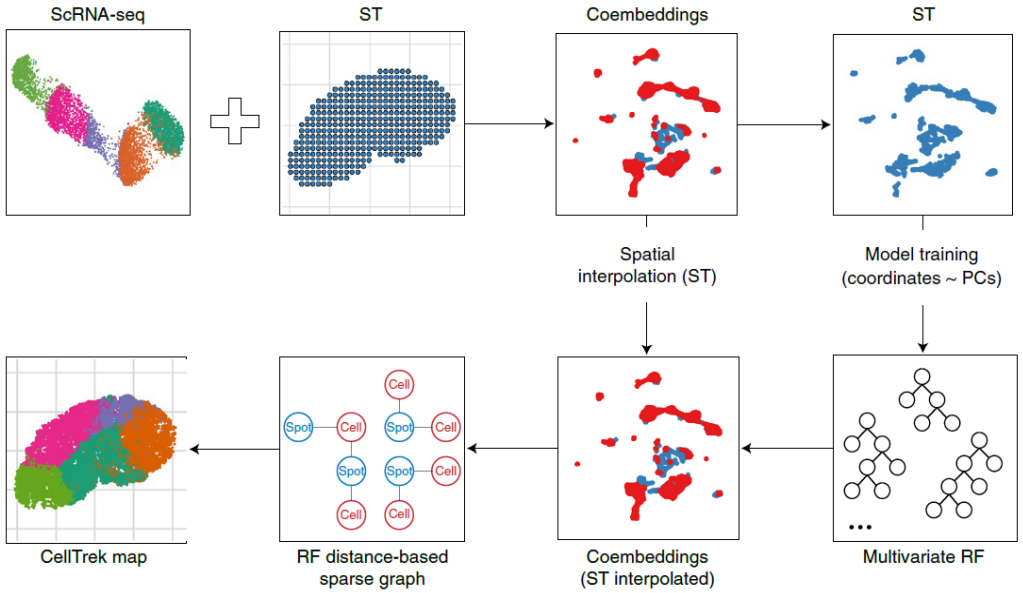
STAligner

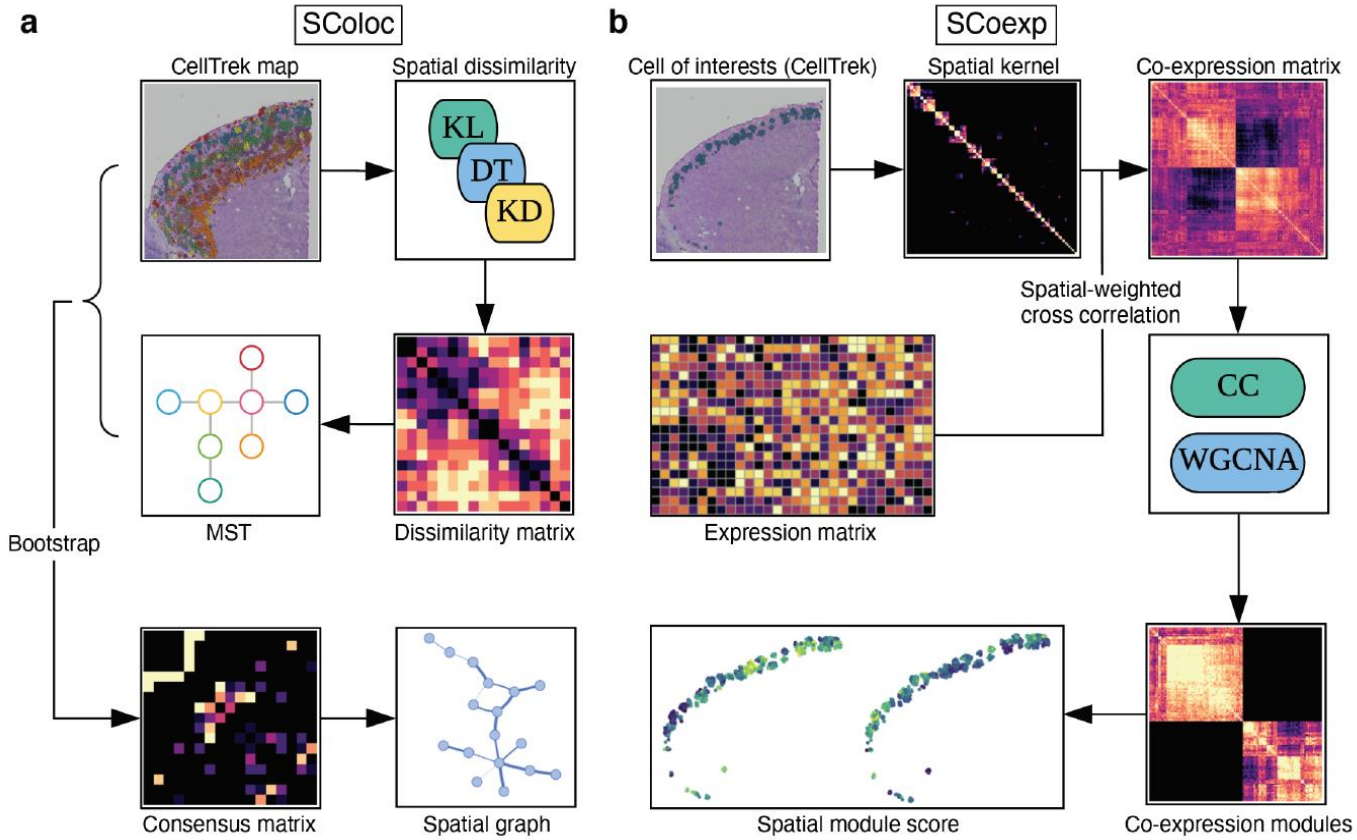
Spatial transcriptomics

ARTICLES
https://doi.org/10.1038/s41587-022-01233-1
nature biotechnology
Check for updates

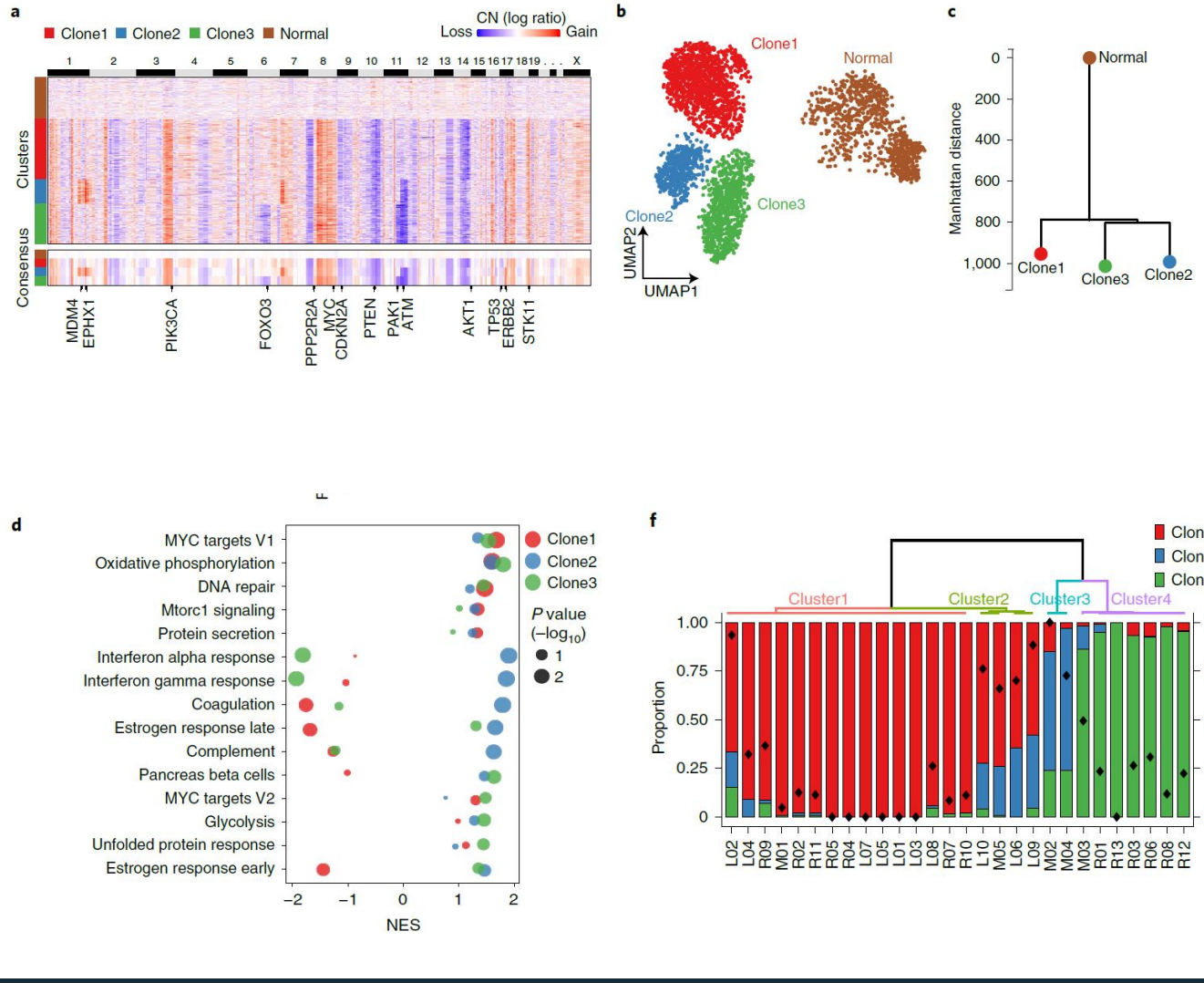
Spatial charting of single-cell transcriptomes in tissues

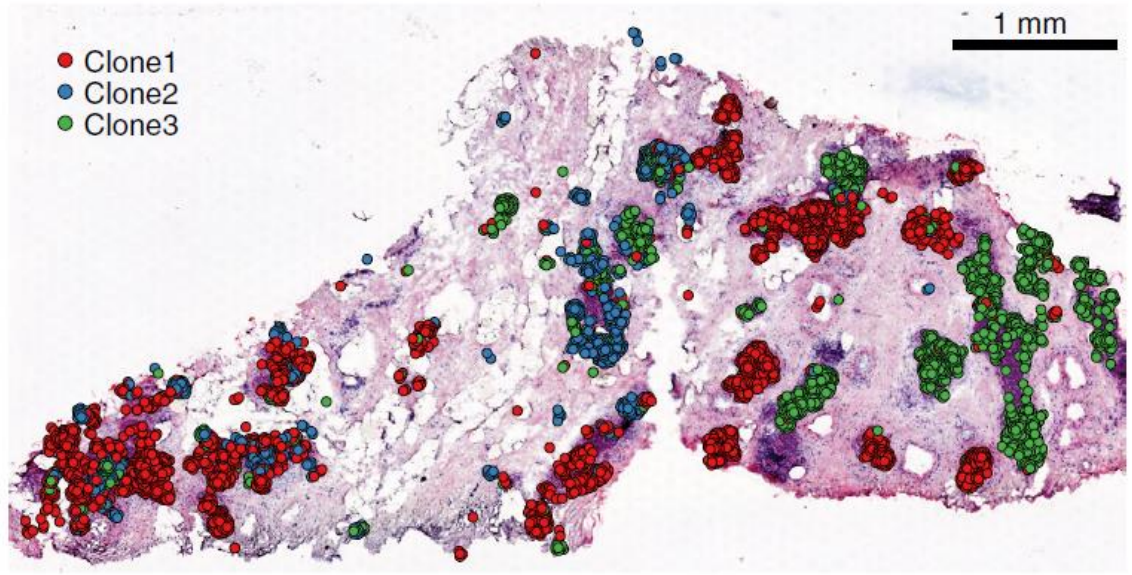
CellTrek

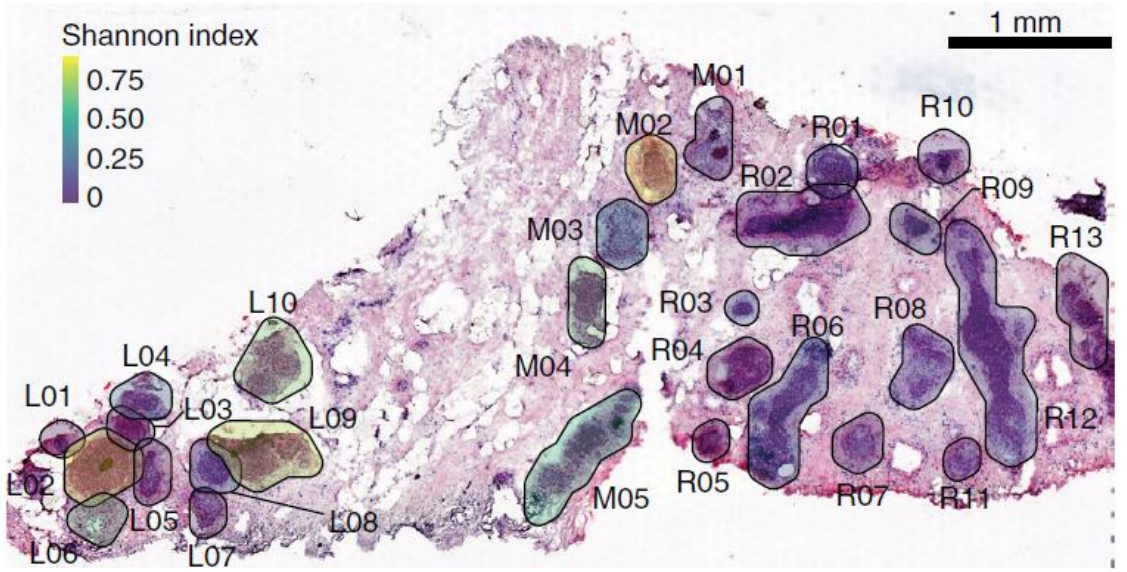




Spatial transcriptomics

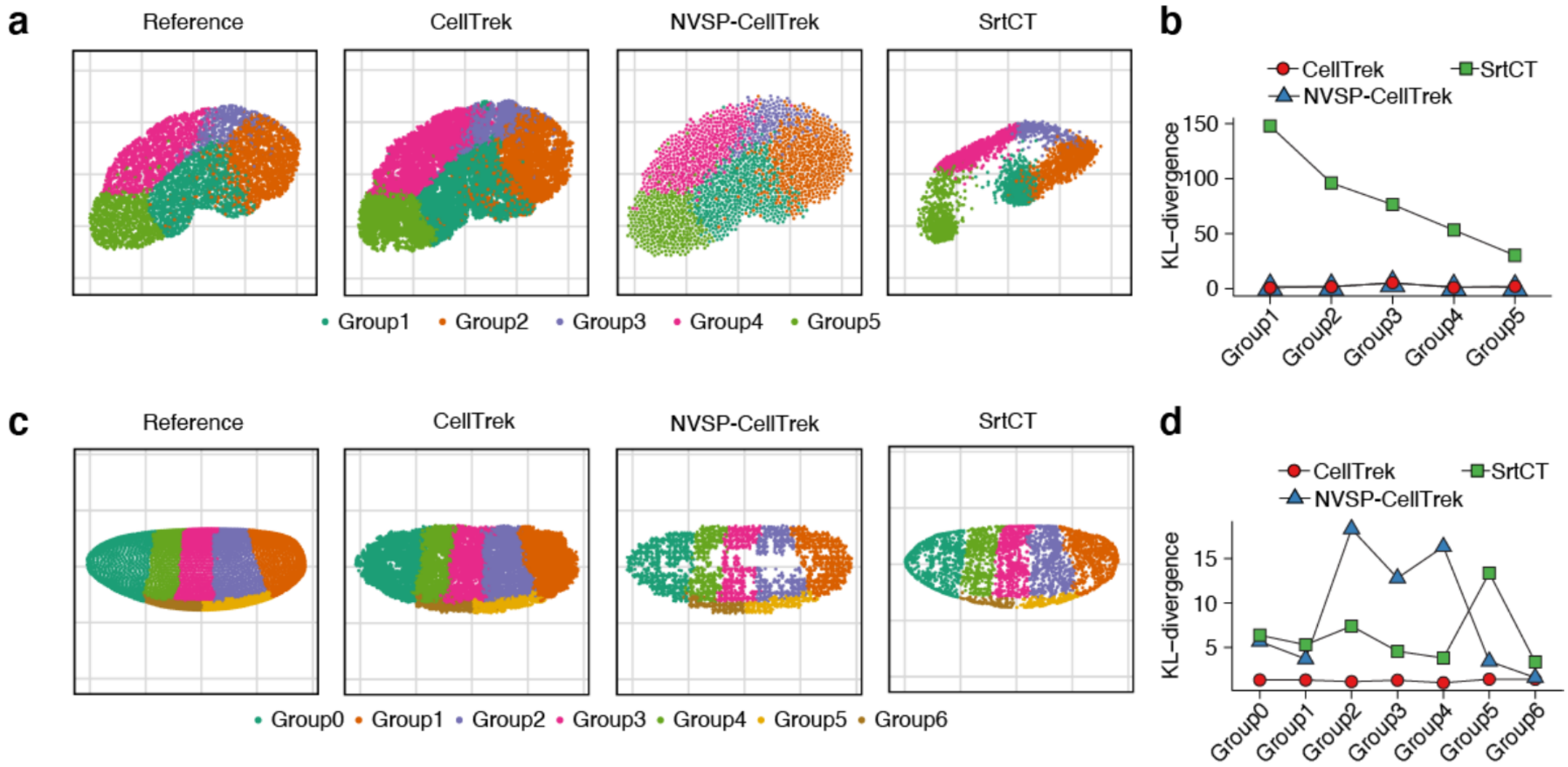






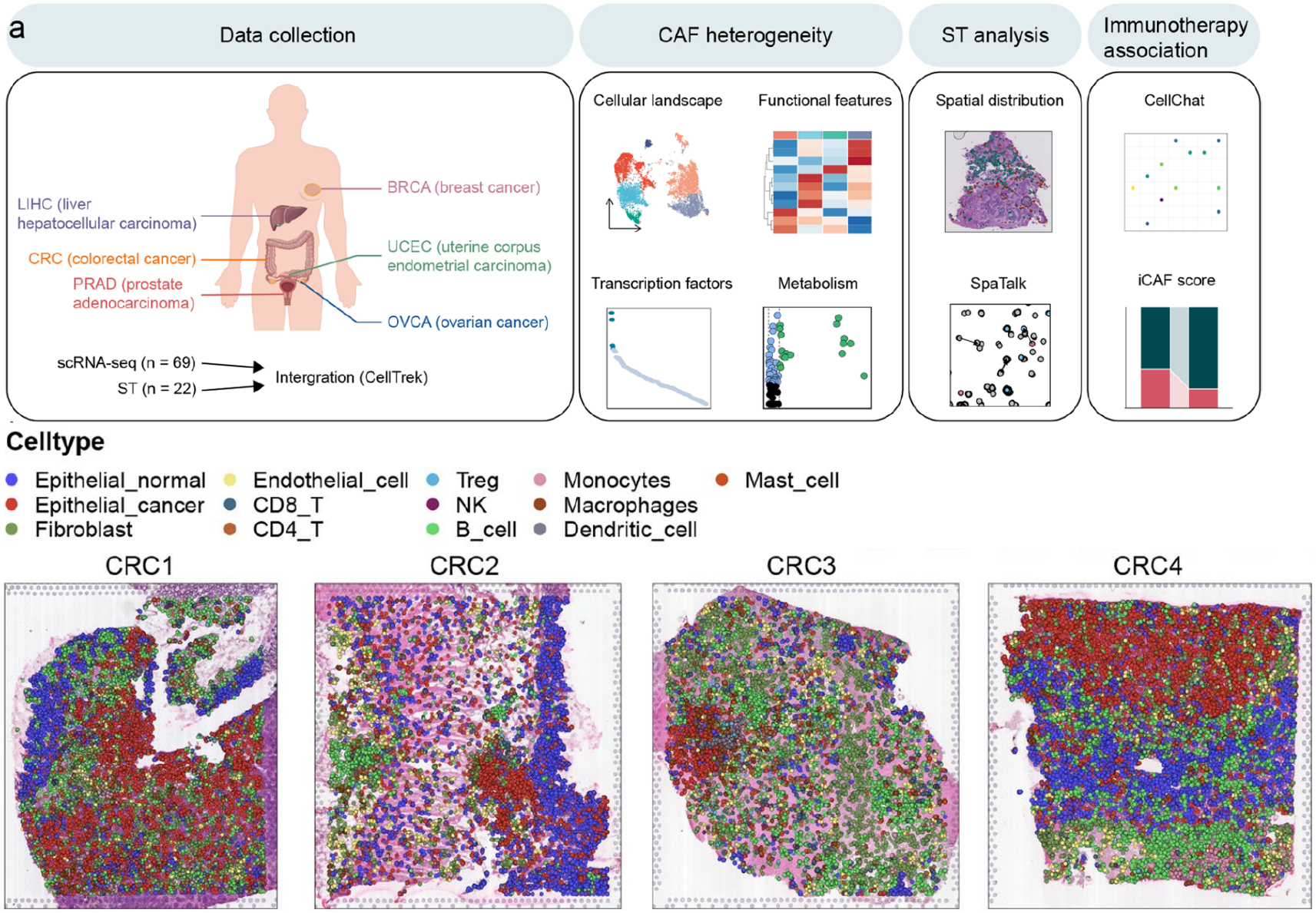
Wei et al., *Nat Biotechnol*, 2022

Spatial transcriptomics



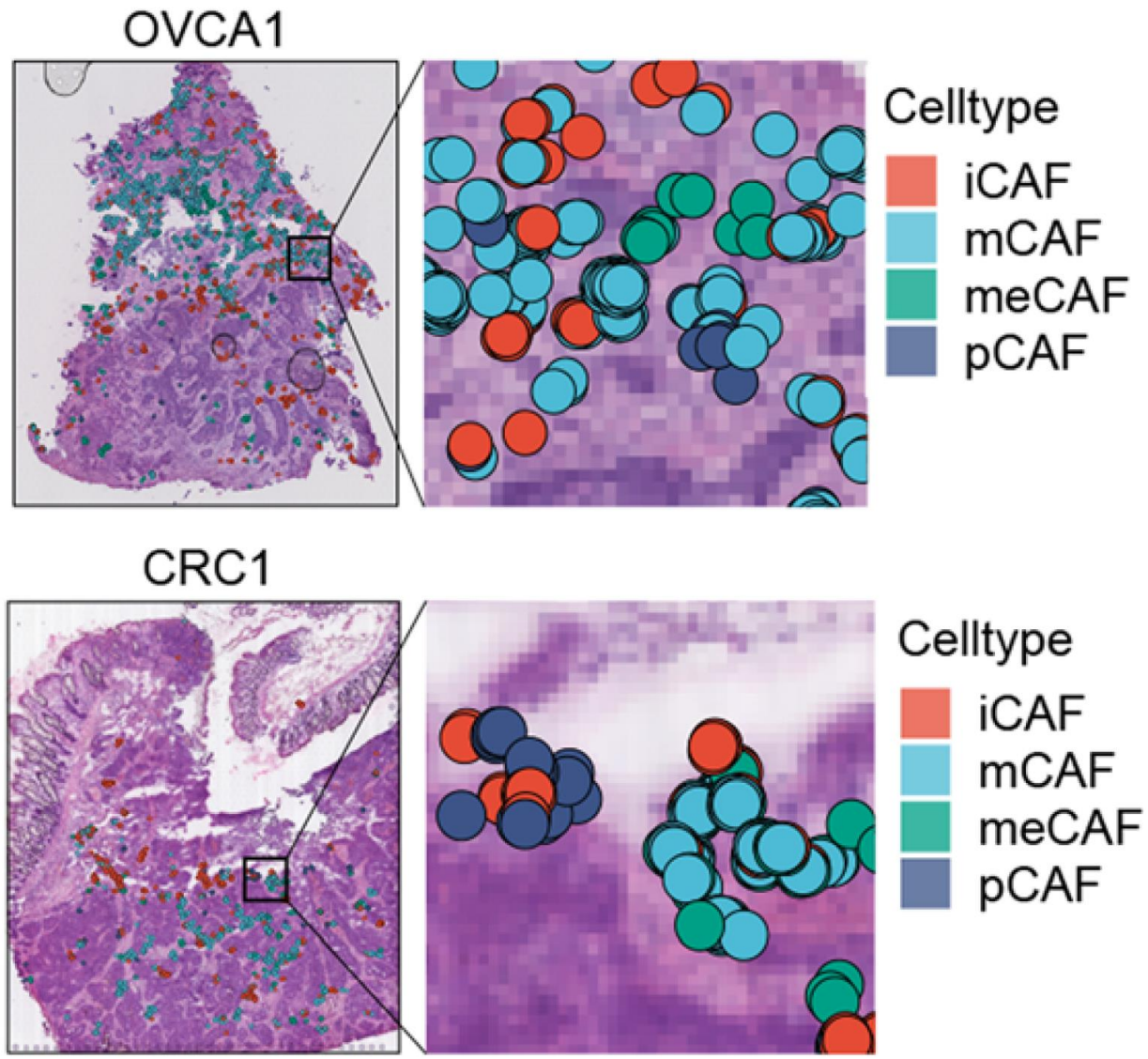
Wei et al., *Nat Biotechnol*, 2022

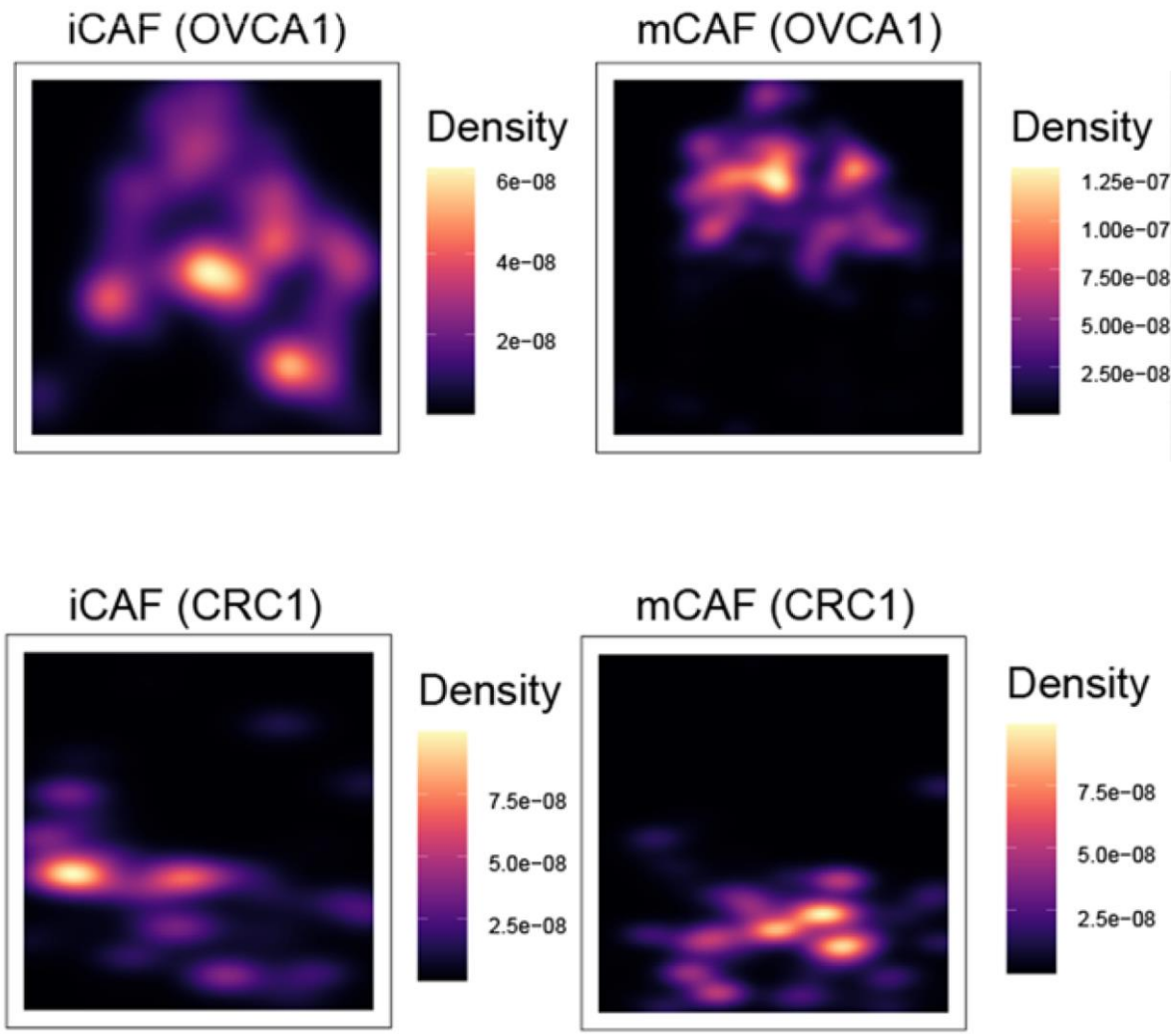
Spatial transcriptomics



Ma et al., *Mol Cancer*, 2023

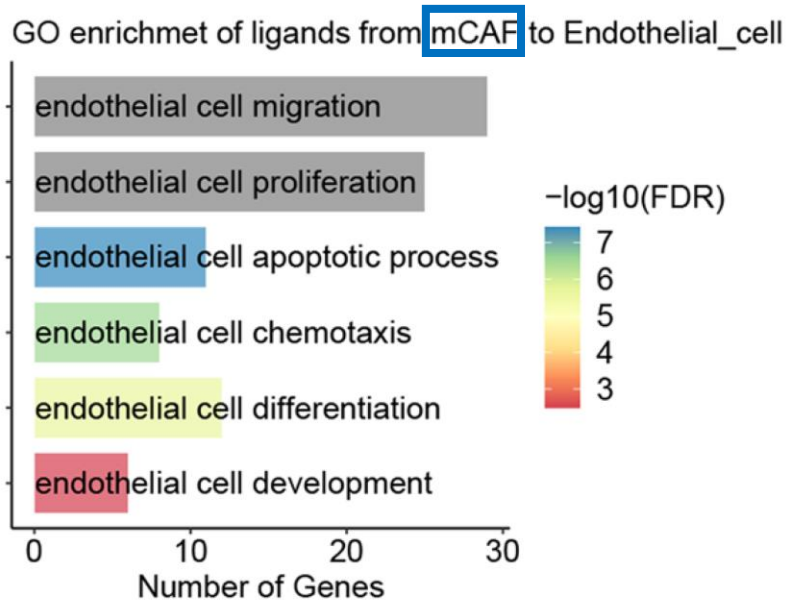
Spatial transcriptomics



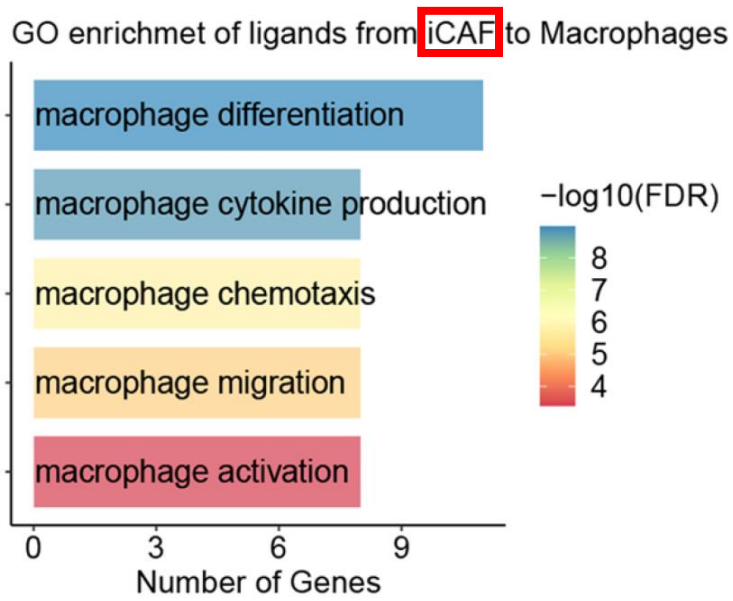


Ma et al., *Mol Cancer*, 2023

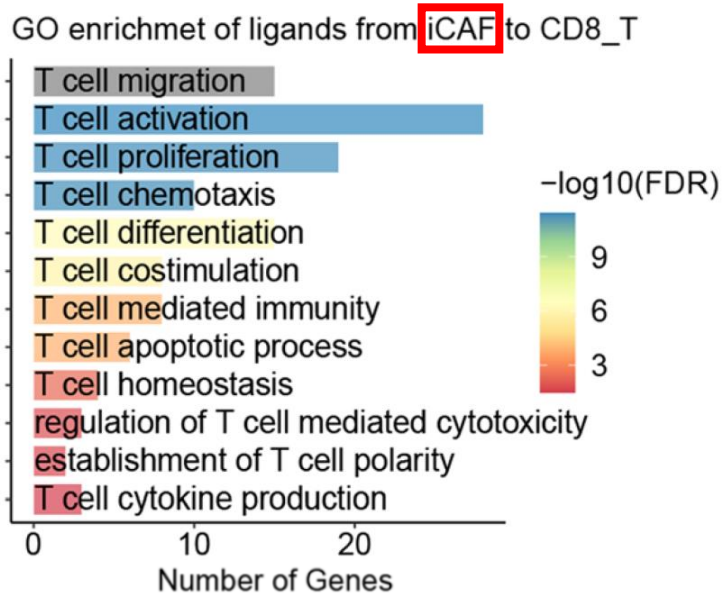
Spatial transcriptomics



Promotes tumor angiogenesis through interactions with endothelial cells



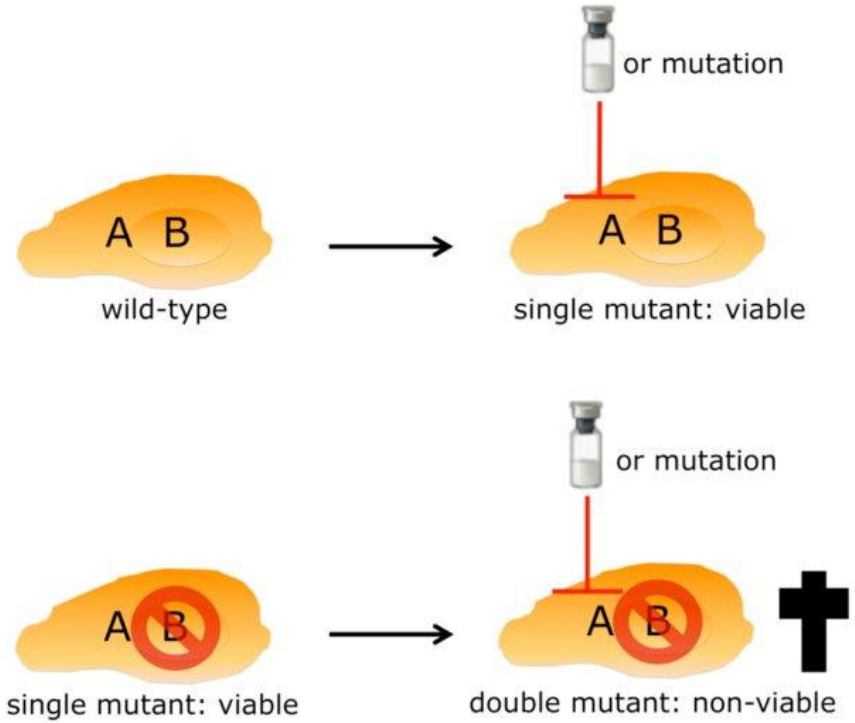
Induces an immunosuppressive environment through interactions with M2 macrophages

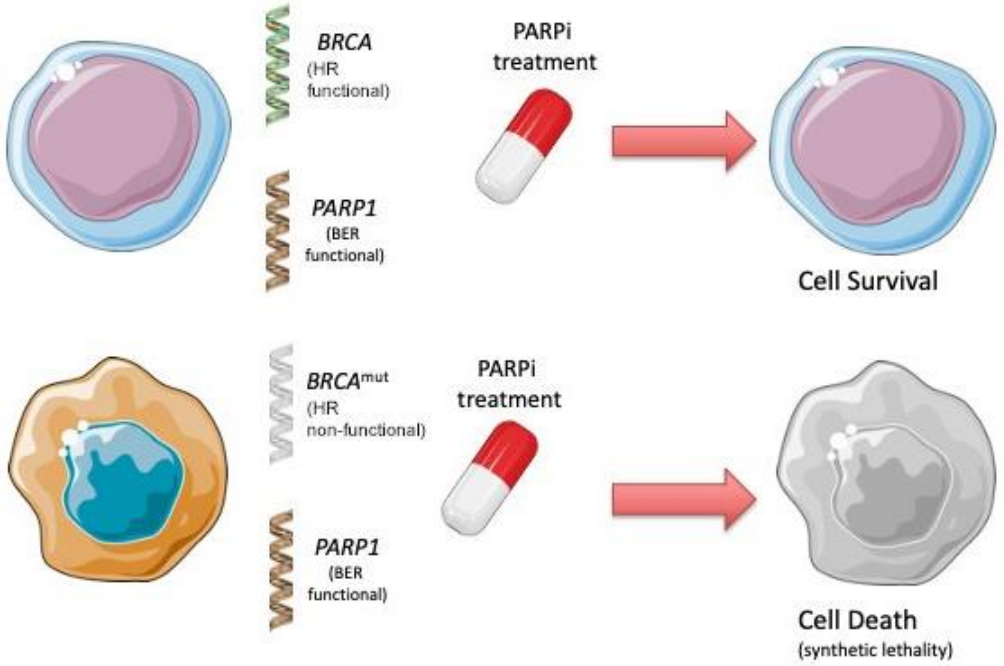


Suppresses CD8+ T cell function

Treatment target

Synthetic Lethality





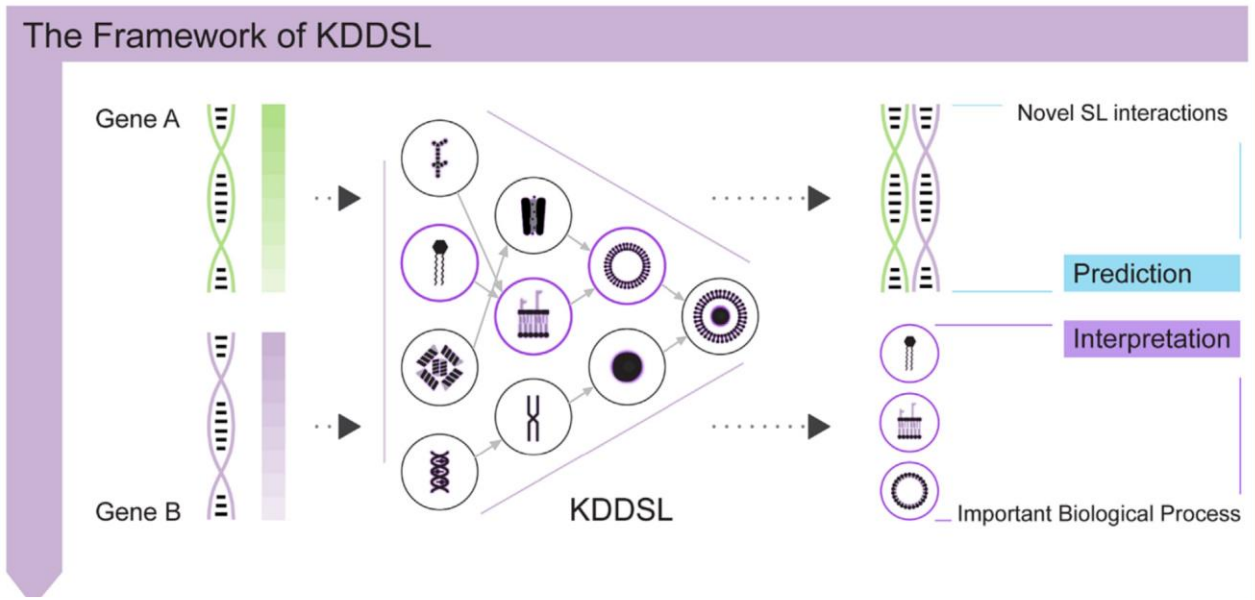
Treatment target

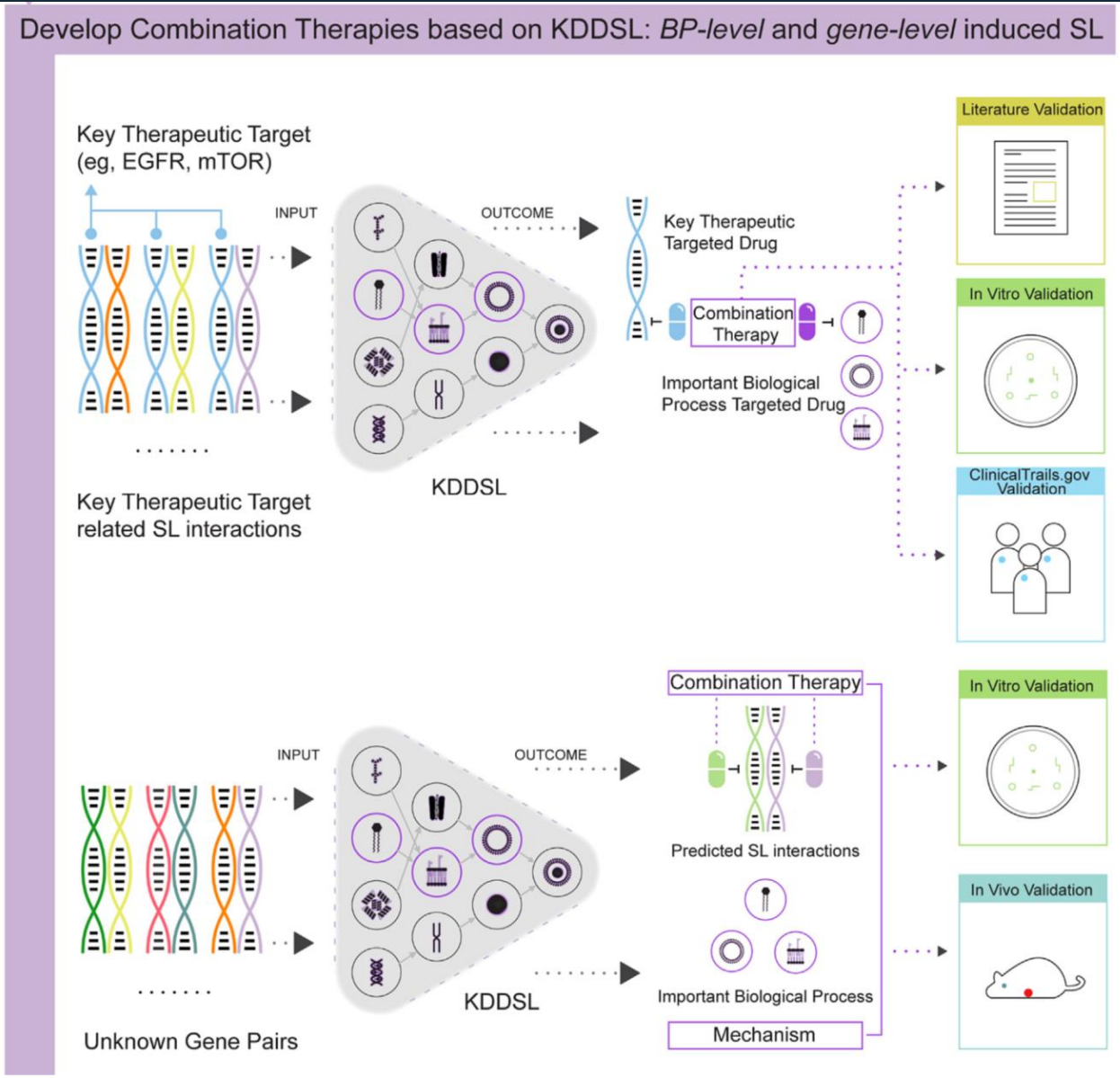
Table 1 | List of supervised machine learning methods for SL prediction

Model & Ref.	Year	Description
SL^2MF[31]	2018	SL^2MF uses logistic matrix factorization to learn gene representations, which are then used to identify potential SL interactions. The authors design an importance weighting scheme to distinguish known and unknown SL pairs and combine PPI and GO information for the prediction.
GRSMF[33]	2019	GRSMF is a method based on graph regularized self-representation matrix factorization (MF). It learns self-representation from known SL interactions and further integrates GO information to predict potential SL interactions.
CMFW[32]	2020	CMFW is a collective matrix factorization-based method that integrates multiple heterogeneous data sources for SL prediction.
DDGCN[34]	2020	DDGCN is the first graph neural network (GNN)-based method for SL prediction. It uses graph convolutional network (GCN) and known SL interaction matrix as features. The authors use coarse-grained node dropout and fine-grained edge dropout to address the issue of overfitting of GCNs on sparse graphs.
GCATSL[35]	2021	GCATSL proposes a graph contextualized attention network to learn gene representations for SL prediction. The authors use data of GO and PPI to generate a set of feature graphs as model inputs and introduce attention mechanisms at the node and feature levels to capture the influence of neighbors and learn gene expression from different feature graphs.
SLMGAE[36]	2021	SLMGAE is a method for predicting SL interactions by leveraging a multi-view graph autoencoder. The authors incorporate data from PPI and GO as supporting views, while utilizing the SL graph as the main view, and apply a graph autoencoder (GAE) to reconstruct these views.
MGE4SL[37]	2021	MGE4SL is a method based on Multi-Graph Ensemble (MGE) to integrate biological knowledge from PPI, GO, and Pathway. It combines the embeddings of features with different neural networks.
KG4SL[39]	2021	KG4SL is a novel model based on graphical neural networks (GNN), and the first method that utilizes knowledge graph (KG) for SL prediction. The integration of KG helps the model obtain more information.
PTGNN[38]	2021	PTGNN is a pre-training method based on graph neural networks that can integrate various data sources and leverage the features obtained from graph-based reconstruction tasks to initialize models for downstream link prediction tasks.
PiLSL[41]	2022	PiLSL is a graph neural network (GNN)-based method that predicts SL by learning the representation of pairwise interaction between two genes.
NSF4SL[42]	2022	NSF4SL is a contrastive learning-based model for SL prediction that eliminates the need for negative samples. It frames the SL prediction task as a gene ranking problem and utilizes two interacting neural network branches to learn representations of SL-related genes, thereby capturing the characteristics of positive SL samples.
SLGNN[40]	2023	SLGNN is a knowledge graph neural networks-based method for synthetic lethality prediction that models gene preferences in distinct relationships in a knowledge graph, providing better interpretability.

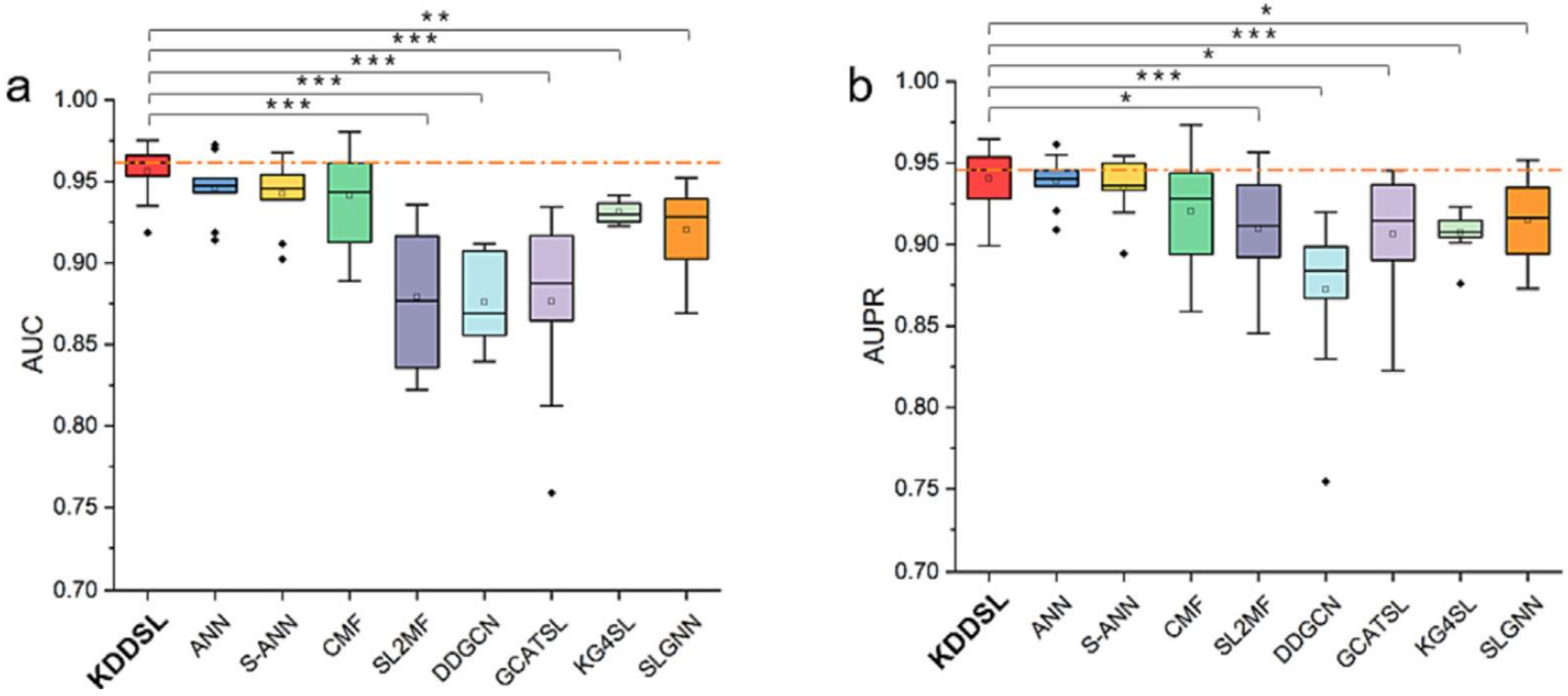
Feng et al., *Nat Commun*, 2024

Treatment target





Treatment target



Wang et al., *J Adv Res*, 2024

Treatment target

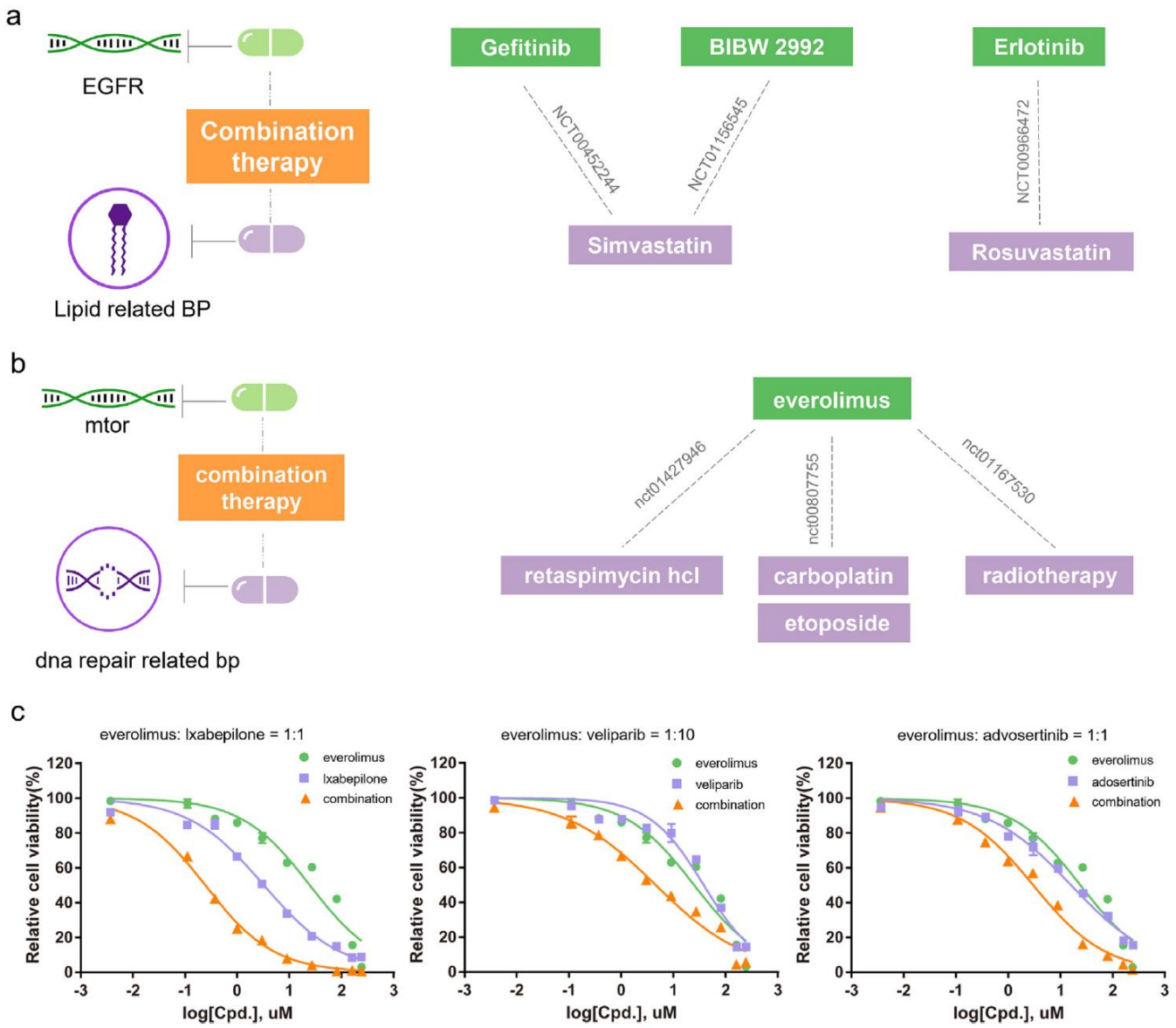
Rank	Layer1	Layer2	Layer3
1	Neutrophil degranulation	Peptidyl-tyrosine phosphorylation	Positive regulation of gene expression
2	Positive regulation of phosphatidylinositol 3-kinase signaling	Cholesterol metabolic process	Negative regulation of multicellular organismal process
3	Nucleobase-containing small molecule metabolic process	Negative regulation of cell population proliferation	Regulation of transcription by RNA polymerase II
4	Bone development	Leukocyte degranulation	Positive regulation of transcription, DNA-templated
5	Regulation of signal transduction by p53 class mediator	Cytokine-mediated signaling pathway	Cellular response to cytokine stimulus
6	Negative regulation of epidermal growth factor receptor signaling pathway	Alcohol biosynthetic process	Peptidyl-amino acid modification
7	Cholesterol biosynthetic process	Skin development	Regulation of cell growth
8	Synapse organization	Positive regulation of transcription by RNA polymerase II	Regulation of lipid metabolic process
9	Negative regulation of NF-kappaB transcription factor activity	Protein-DNA complex assembly	Regulation of muscle system process
10	Lamellipodium assembly	Regulation of response to DNA damage stimulus	Regulation of inflammatory response

Table S4. Important subsystems of EGFR-related SL pairs. (Layer represents the position where the subsystems located in the neural network. As the layer number increases, subsystems in this layer tend to be relatively broader)

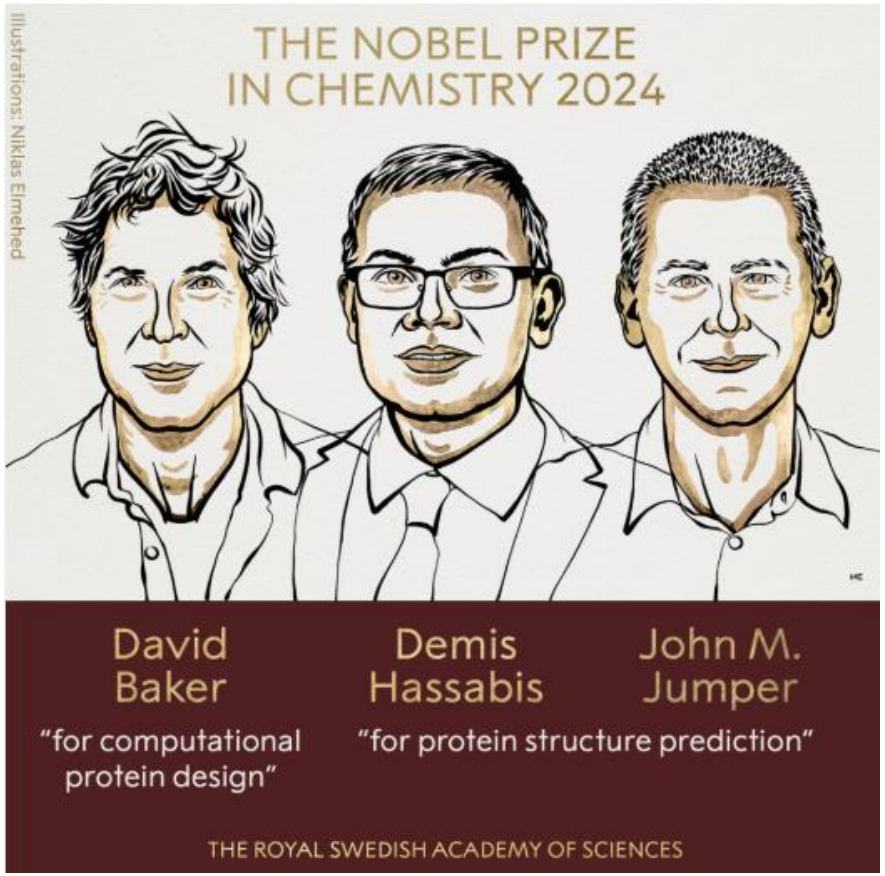
Rank	Layer1	Layer2	Layer3
1	Regulation of neurotransmitter levels	Mitotic DNA damage checkpoint signaling	Intracellular protein transport
2	Chemical synaptic transmission	Synaptic signaling	Supramolecular fiber organization
3	Regulation of regulated secretory pathway	DNA biosynthetic process	Negative regulation of gene expression
4	Exocytic process	Regulation of transcription by RNA polymerase II	Cellular protein-containing complex assembly
5	Signal release	Double-strand break repair	Positive regulation of gene expression
6	Regulation of signal transduction by p53 class mediator	Histone modification	Regulation of cell growth
7	Synapse organization	Positive regulation of cell projection organization	Anatomical structure homeostasis
8	Axon guidance	negative regulation of G1/S transition of mitotic cell cycle	Negative regulation of multicellular organismal process
9	DNA damage response, signal transduction by p53 class mediator resulting in cell cycle arrest	Regulation of response to DNA damage stimulus	Animal organ morphogenesis
10	Positive regulation of ATP-dependent activity	Cytokine-mediated signaling pathway	Organelle localization

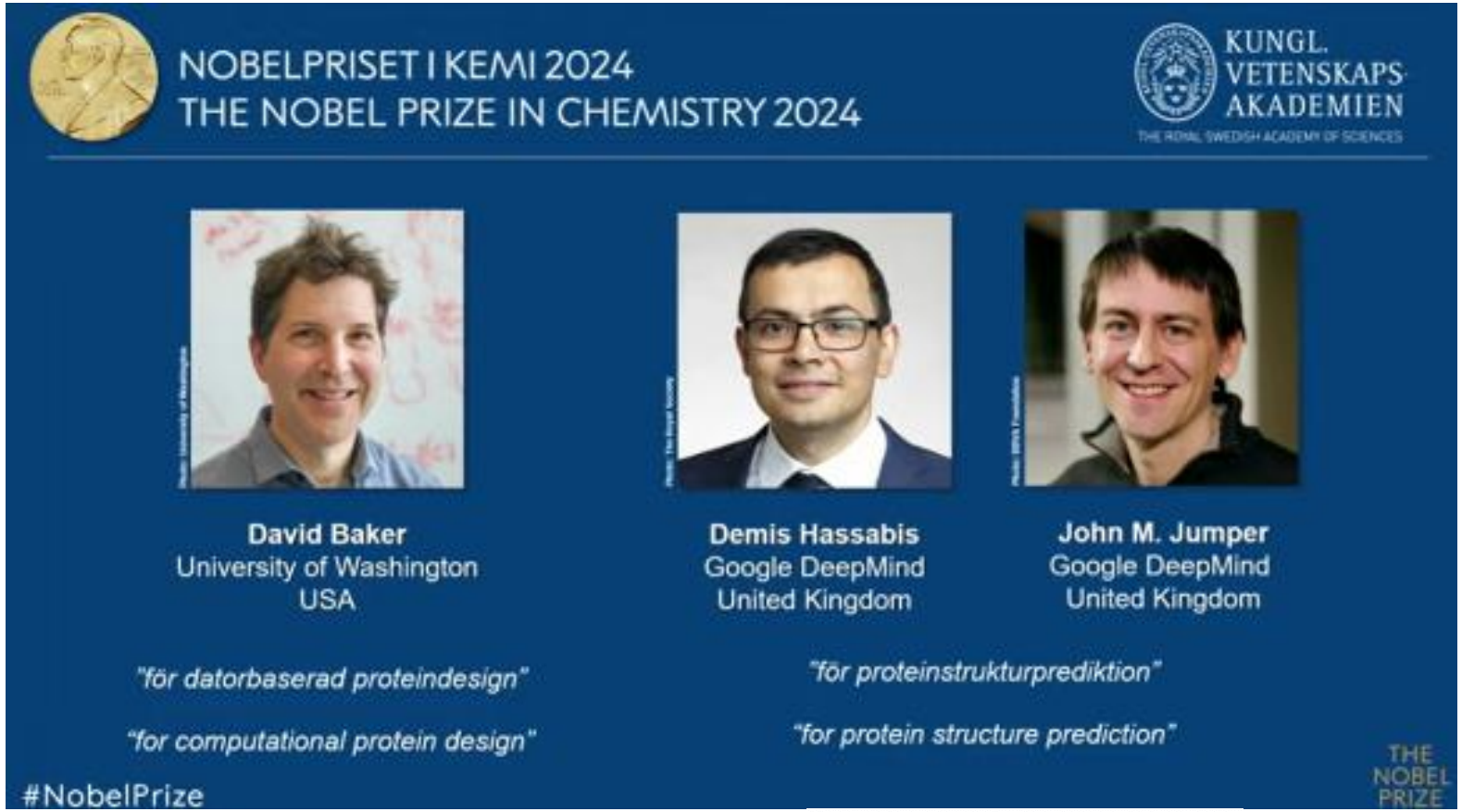
Table S5. Important subsystems of mTOR-related SL pairs.

Treatment target



AlphaFold





RoseTTAFold

AlphaFold

Accelerating breakthroughs in biology with AI

AlphaFold

RESEARCH ARTICLE

PROTEIN FOLDING

Accurate prediction of protein structures and interactions using a three-track neural network

Minkyung Baek[1,2], Frank DiMaio[1,2], Ivan Anishchenko[1,2], Justas Dauparas[1,2], Sergey Ovchinnikov[3,4], Gyu Rie Lee[1,2], Jue Wang[1,2], Qian Cong[5,6], Lisa N. Kinch[7], R. Dustin Schaeffer[6], Claudia Millán[8], Hahnbeom Park[1,2], Carson Adams[1,2], Caleb R. Glassman[9,10,11], Andy DeGiovanni[12], Jose H. Pereira[12], Andria V. Rodrigues[12], Alberdina A. van Dijk[13], Ana C. Ebrecht[13], Diederik J. Opperman[14], Theo Sagmeister[15], Christoph Buhlheller[15,16], Tea Pavkov-Keller[15,17], Manoj K. Rathinaswamy[18], Udit Dalwadi[19], Calvin K. Yip[19], John E. Burke[18], K. Christopher Garcia[9,10,11,20], Nick V. Grishin[6,7,21], Paul D. Adams[12,22], Randy J. Read[8], David Baker[1,2,23]*

Baek et al., *Science*, 2021

AlphaFold

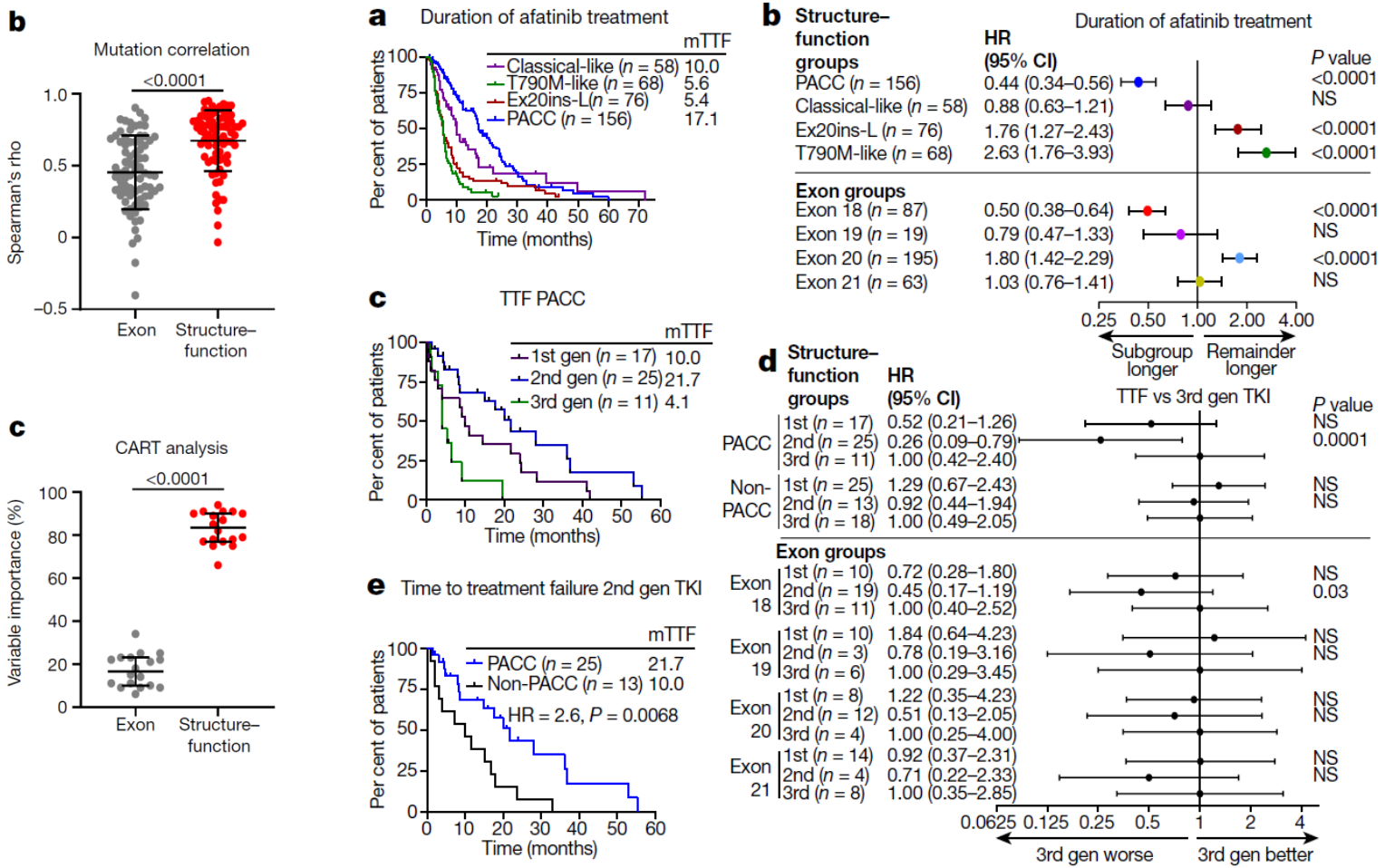


Fig. 4 | Structure–function groups better predict patient outcomes than exon-based groups.

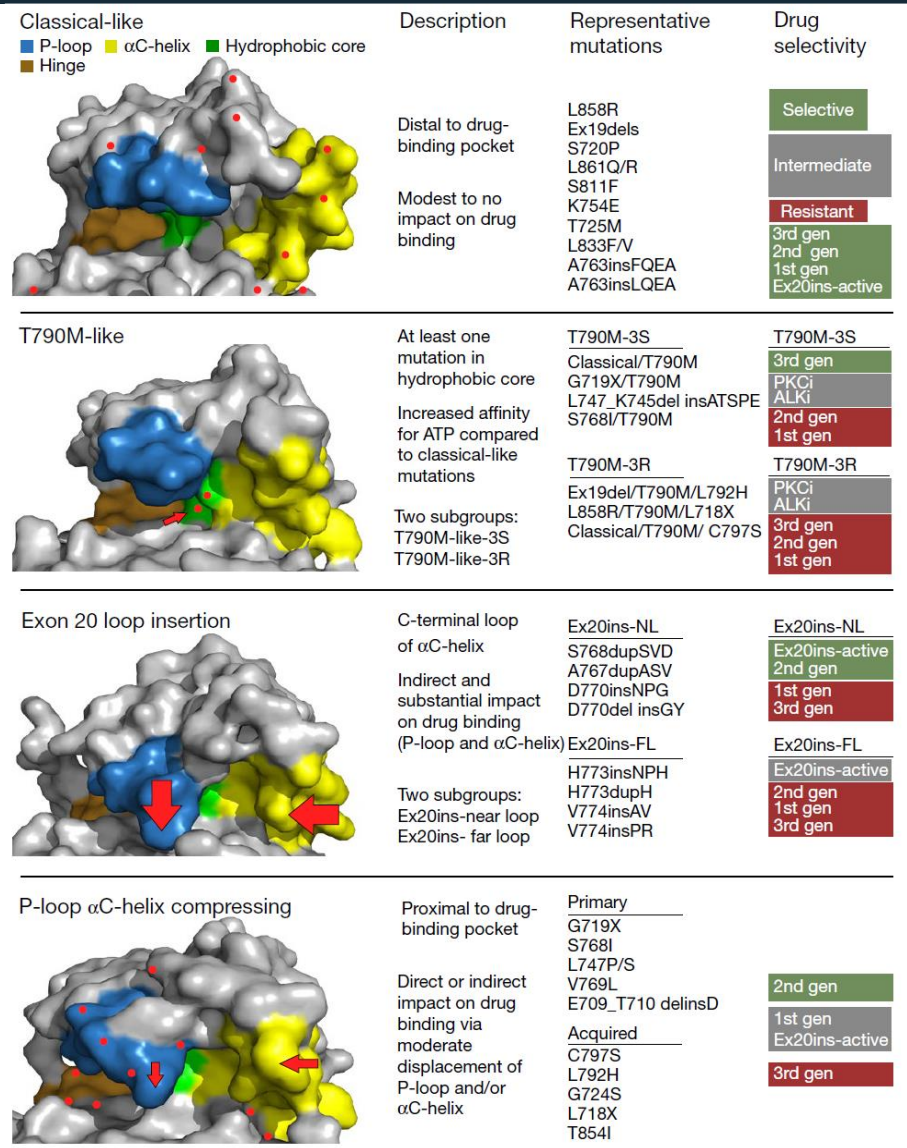
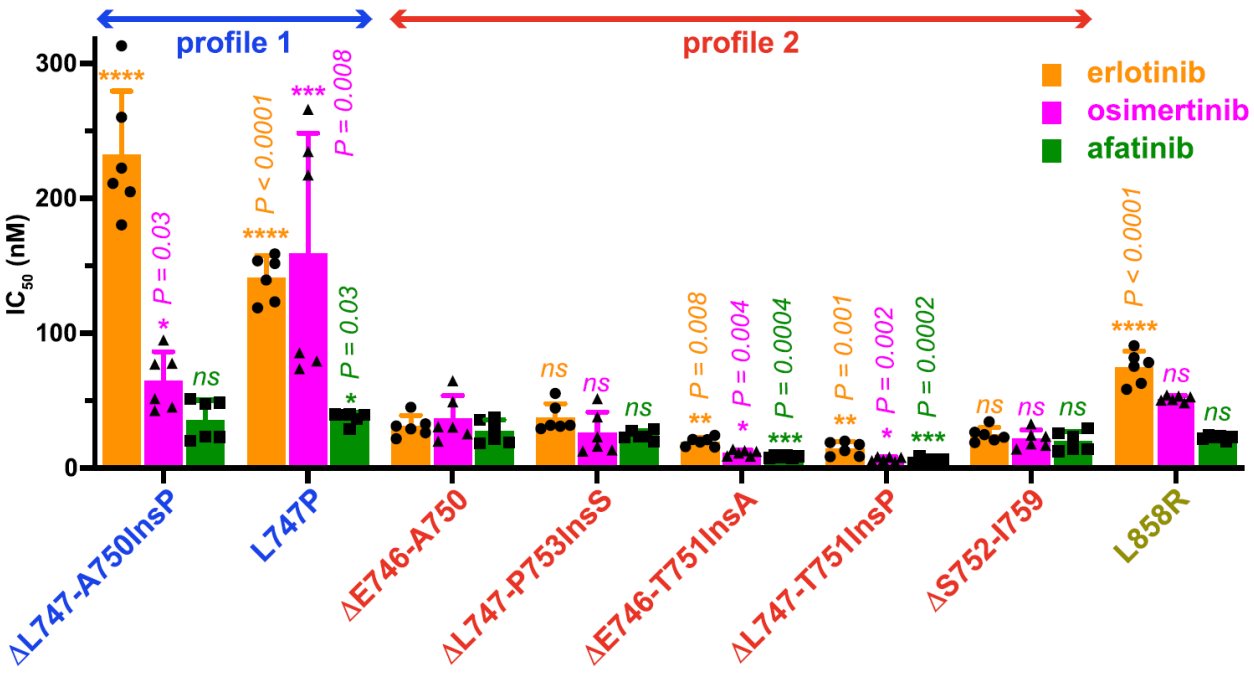
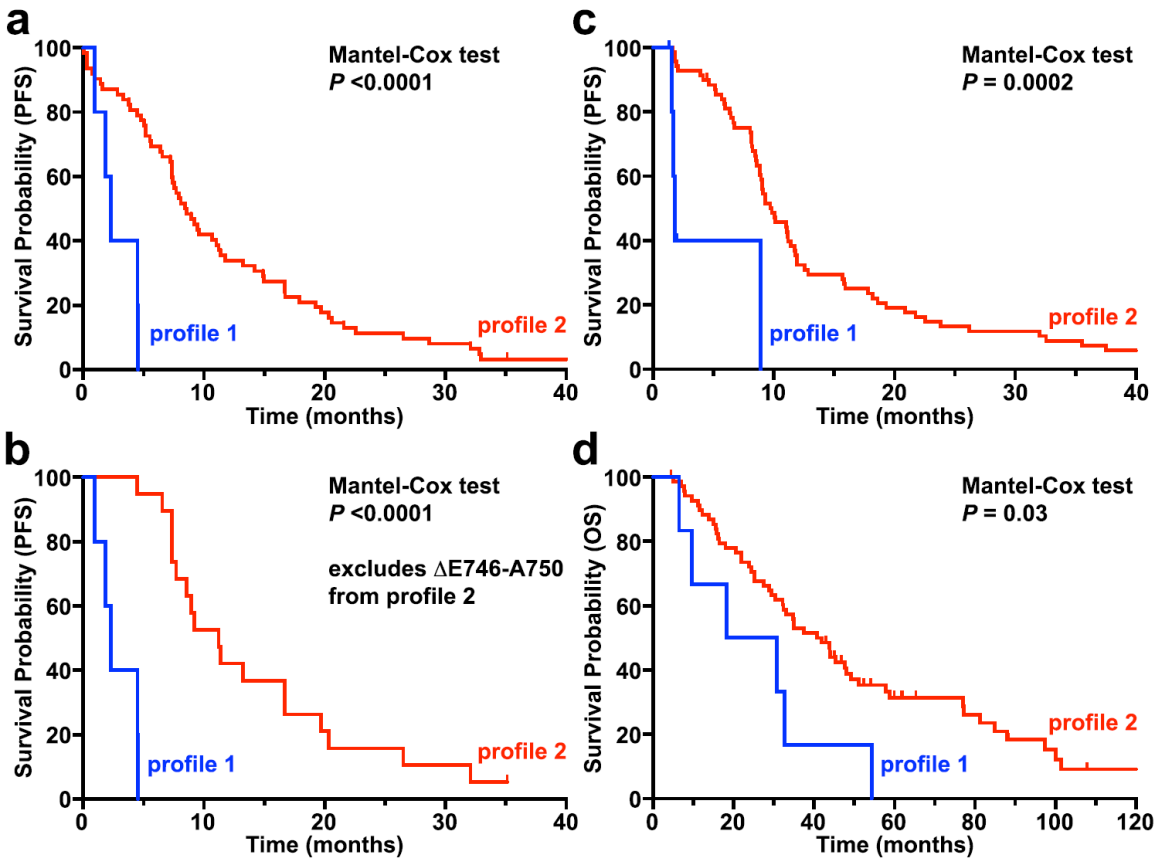


Fig. 5 | EGFR mutations can be divided into four distinct subgroups.

Robichaux et al., *Nature*, 2021

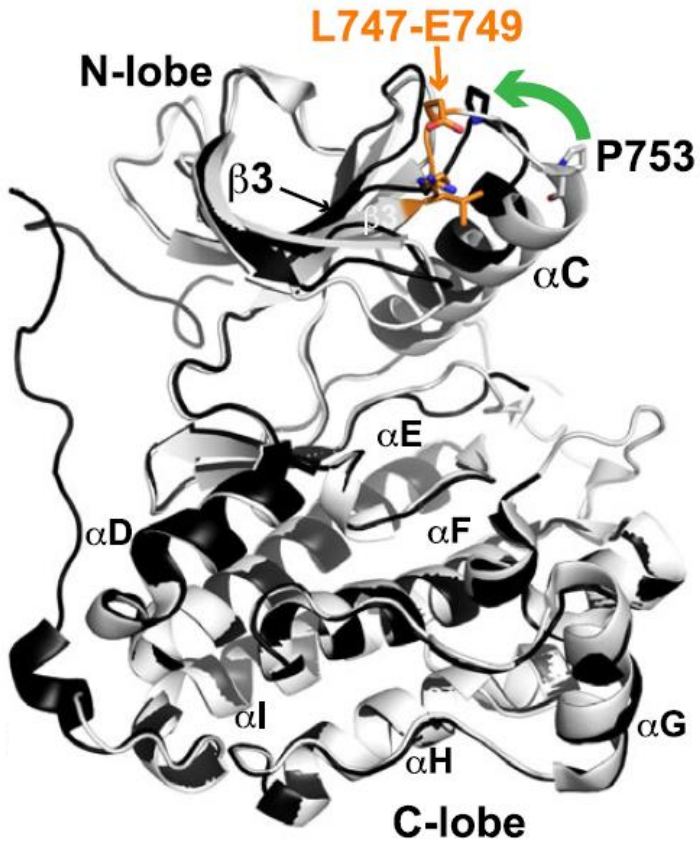
AlphaFold



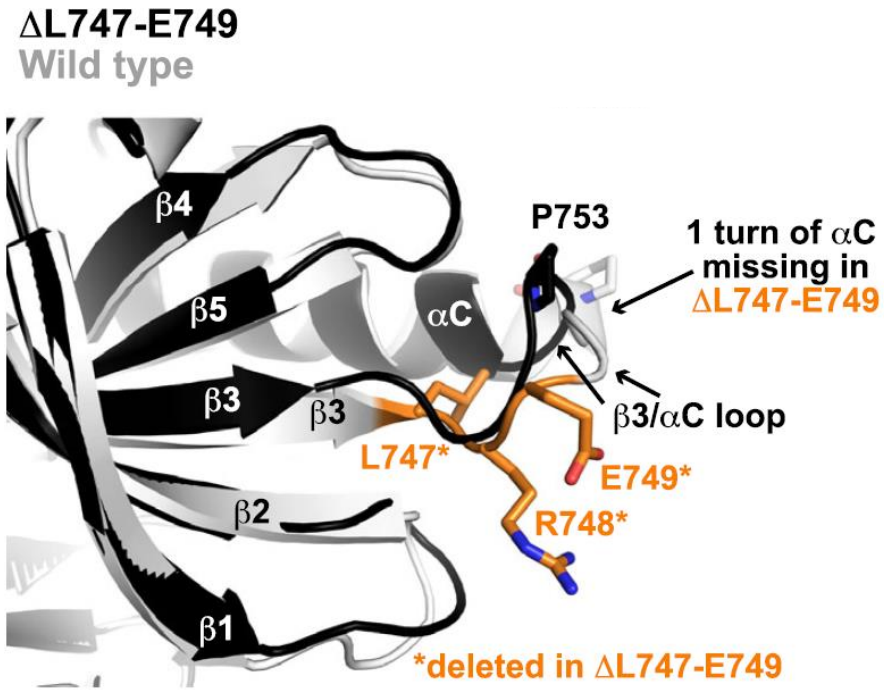


Rosenburgh et al., *Nat Commun*, 2022

AlphaFold

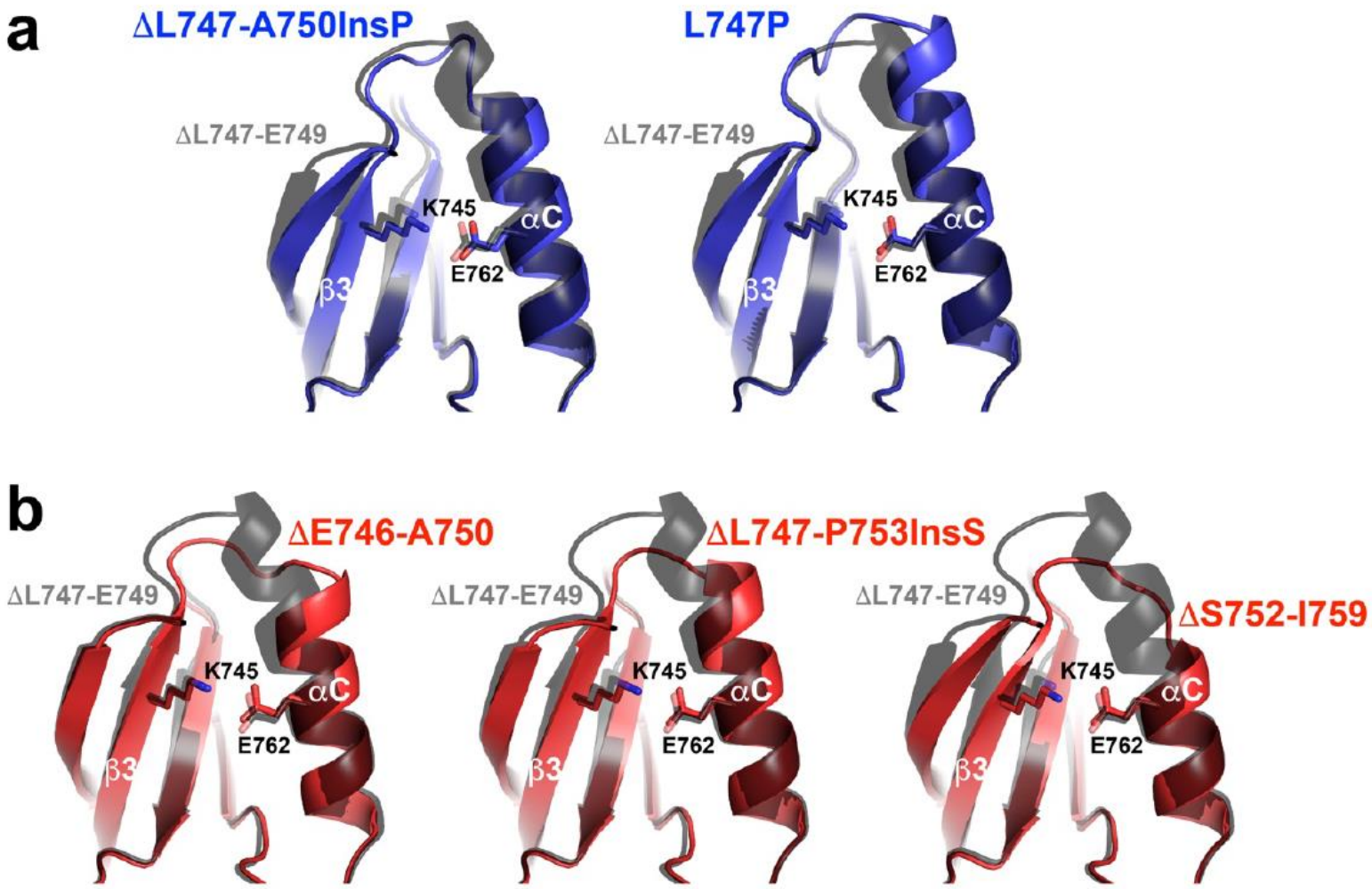


ΔL747-E749 exon 19 variant



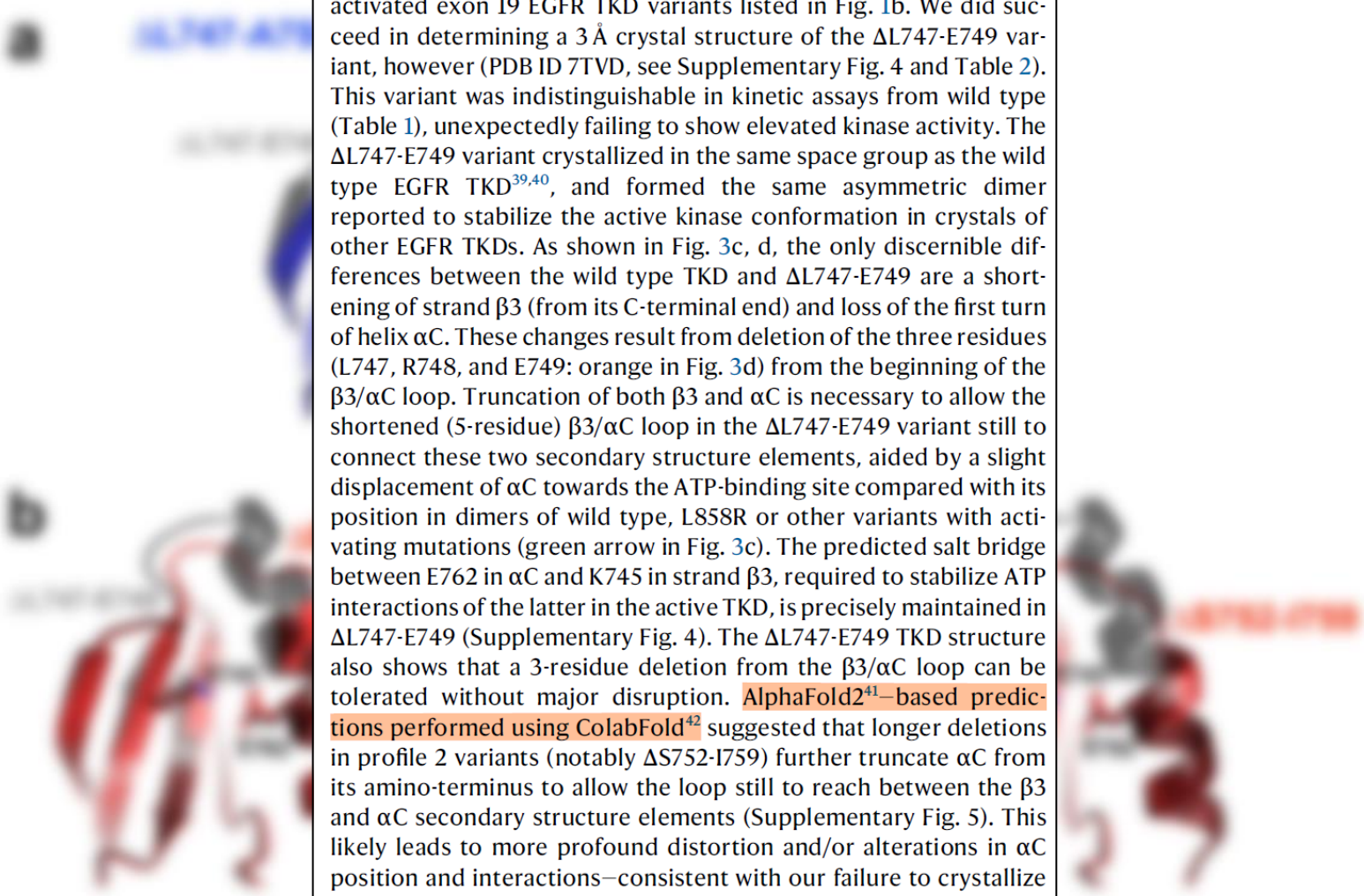
Wild type

AlphaFold



Rosenburgh et al., *Nat Commun*, 2022

AlphaFold

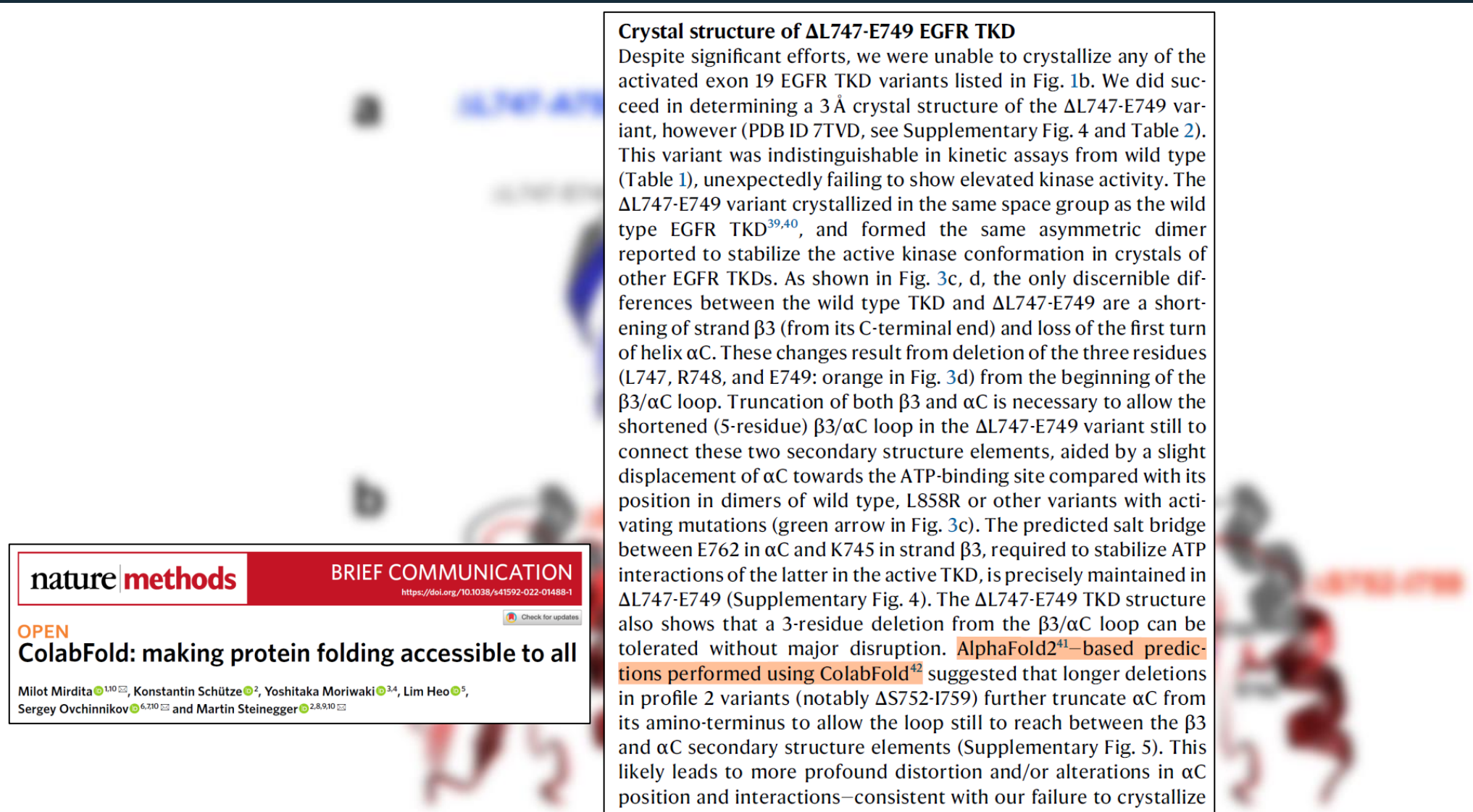


Crystal structure of ΔL747-E749 EGFR TKD

Despite significant efforts, we were unable to crystallize any of the activated exon 19 EGFR TKD variants listed in Fig. 1b. We did succeed in determining a 3 Å crystal structure of the ΔL747-E749 variant, however (PDB ID 7TVD, see Supplementary Fig. 4 and Table 2). This variant was indistinguishable in kinetic assays from wild type (Table 1), unexpectedly failing to show elevated kinase activity. The ΔL747-E749 variant crystallized in the same space group as the wild type EGFR TKD[39,40], and formed the same asymmetric dimer reported to stabilize the active kinase conformation in crystals of other EGFR TKDs. As shown in Fig. 3c, d, the only discernible differences between the wild type TKD and ΔL747-E749 are a shortening of strand β3 (from its C-terminal end) and loss of the first turn of helix αC. These changes result from deletion of the three residues (L747, R748, and E749: orange in Fig. 3d) from the beginning of the β3/αC loop. Truncation of both β3 and αC is necessary to allow the shortened (5-residue) β3/αC loop in the ΔL747-E749 variant still to connect these two secondary structure elements, aided by a slight displacement of αC towards the ATP-binding site compared with its position in dimers of wild type, L858R or other variants with activating mutations (green arrow in Fig. 3c). The predicted salt bridge between E762 in αC and K745 in strand β3, required to stabilize ATP interactions of the latter in the active TKD, is precisely maintained in ΔL747-E749 (Supplementary Fig. 4). The ΔL747-E749 TKD structure also shows that a 3-residue deletion from the β3/αC loop can be tolerated without major disruption. AlphaFold2[41]–based predictions performed using ColabFold[42] suggested that longer deletions in profile 2 variants (notably ΔS752-I759) further truncate αC from its amino-terminus to allow the loop still to reach between the β3 and αC secondary structure elements (Supplementary Fig. 5). This likely leads to more profound distortion and/or alterations in αC position and interactions—consistent with our failure to crystallize such exon 19 variants. The profile 1 variants have less truncated αC helices.

Rosenburgh et al., *Nat Commun*, 2022

AlphaFold



nature methods
BRIEF COMMUNICATION
https://doi.org/10.1038/s41592-022-01488-1
Check for updates
OPEN

ColabFold: making protein folding accessible to all

Milot Mirdita[1,10] ✉, Konstantin Schütze[2], Yoshitaka Moriwaki[3,4], Lim Heo[5], Sergey Ovchinnikov[6,7,10] ✉ and Martin Steinegger[2,8,9,10] ✉

Crystal structure of ΔL747-E749 EGFR TKD

Despite significant efforts, we were unable to crystallize any of the activated exon 19 EGFR TKD variants listed in Fig. 1b. We did succeed in determining a 3 Å crystal structure of the ΔL747-E749 variant, however (PDB ID 7TVD, see Supplementary Fig. 4 and Table 2). This variant was indistinguishable in kinetic assays from wild type (Table 1), unexpectedly failing to show elevated kinase activity. The ΔL747-E749 variant crystallized in the same space group as the wild type EGFR TKD[39,40], and formed the same asymmetric dimer reported to stabilize the active kinase conformation in crystals of other EGFR TKDs. As shown in Fig. 3c, d, the only discernible differences between the wild type TKD and ΔL747-E749 are a shortening of strand β3 (from its C-terminal end) and loss of the first turn of helix αC. These changes result from deletion of the three residues (L747, R748, and E749: orange in Fig. 3d) from the beginning of the β3/αC loop. Truncation of both β3 and αC is necessary to allow the shortened (5-residue) β3/αC loop in the ΔL747-E749 variant still to connect these two secondary structure elements, aided by a slight displacement of αC towards the ATP-binding site compared with its position in dimers of wild type, L858R or other variants with activating mutations (green arrow in Fig. 3c). The predicted salt bridge between E762 in αC and K745 in strand β3, required to stabilize ATP interactions of the latter in the active TKD, is precisely maintained in ΔL747-E749 (Supplementary Fig. 4). The ΔL747-E749 TKD structure also shows that a 3-residue deletion from the β3/αC loop can be tolerated without major disruption. AlphaFold2[41]–based predictions performed using ColabFold[42] suggested that longer deletions in profile 2 variants (notably ΔS752-I759) further truncate αC from its amino-terminus to allow the loop still to reach between the β3 and αC secondary structure elements (Supplementary Fig. 5). This likely leads to more profound distortion and/or alterations in αC position and interactions—consistent with our failure to crystallize such exon 19 variants. The profile 1 variants have less truncated αC helices.

Rosenburgh et al., *Nat Commun*, 2022

AlphaFold

AlphaFold2.ipynb

파일 수정 보기 삽입 런타임 도구 도움말

+ 코드 + 텍스트 Drive로 복사

ColabFold v1.5.5: AlphaFold2 using MMseqs2

Easy to use protein structure and complex prediction using AlphaFold2 and Alphafold2-multimer. Sequence alignments/templates are generated through MMseqs2 and HHsearch. For more details, see bottom of the notebook, checkout the ColabFold GitHub and Nature Protocols.

Old versions: v1.4, v1.5.1, v1.5.2, v1.5.3-patch

Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: Making protein folding accessible to all. *Nature Methods*, 2022

AlphaFold

AlphaFold2.ipynb

파일 수정 보기 삽입 런타임 도구 도움말

+ 코드 + 텍스트 Drive로 복사

ColabFold v1.5.5: AlphaFold2 using MMseqs2

Easy to use protein structure and complex prediction using AlphaFold2 and Alphafold2-multimer. Sequence alignments/templates are generated through MMseqs2 and HHsearch. For more details, see bottom of the notebook, checkout the ColabFold GitHub and Nature Protocols.

Old versions: v1.4, v1.5.1, v1.5.2, v1.5.3-patch

Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: Making protein folding accessible to all. *Nature Methods*, 2022

National Library of Medicine
National Center for Biotechnology Information

Protein | Protein | Advanced

FASTA | Send to:

surface glycoprotein [Severe acute respiratory syndrome coronavirus 2]

NCBI Reference Sequence: YP_009724390.1

GenPept | Identical Proteins | Graphics

```
>YP_009724390.1 surface glycoprotein [Severe acute respiratory syndrome coronavirus 2]
MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV
SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPF
LGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTP
NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYN
ENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASV
YAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIAD
YNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYF
PLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL
PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLT
PTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLG
AENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGI
AVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDC
LGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIG
VTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLND
LSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLM
SFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNT
FVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVA
KNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDD
SEPVLKGVKLHYT
```

AlphaFold

AlphaFold2.ipynb

파일 수정 보기 삽입 런타임 도구 도움말

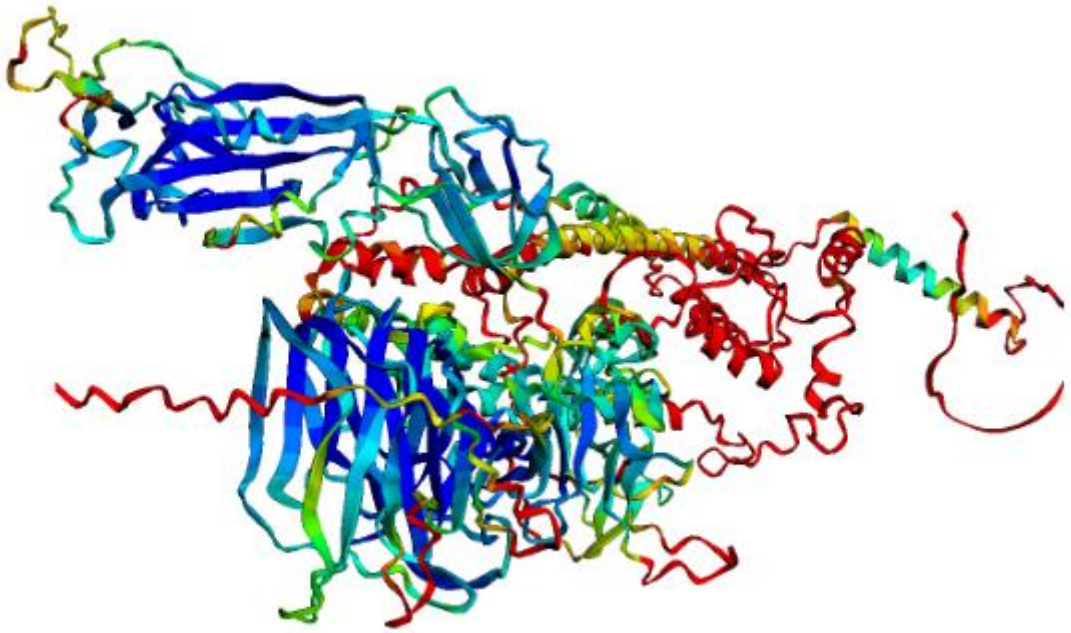
+ 코드 + 텍스트 Drive로 복사

ColabFold v1.5.5: AlphaFold2 using MMseqs2

Easy to use protein structure and complex prediction using AlphaFold2 and Alphafold2-multimer. Sequence alignments/templates are generated through MMseqs2 and HHsearch. For more details, see bottom of the notebook, checkout the ColabFold GitHub and Nature Protocols.

Old versions: v1.4, v1.5.1, v1.5.2, v1.5.3-patch

Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: Making protein folding accessible to all. *Nature Methods*, 2022



NIH National Library of Medicine
National Center for Biotechnology Information

Protein | Protein | Advanced

FASTA | Send to:

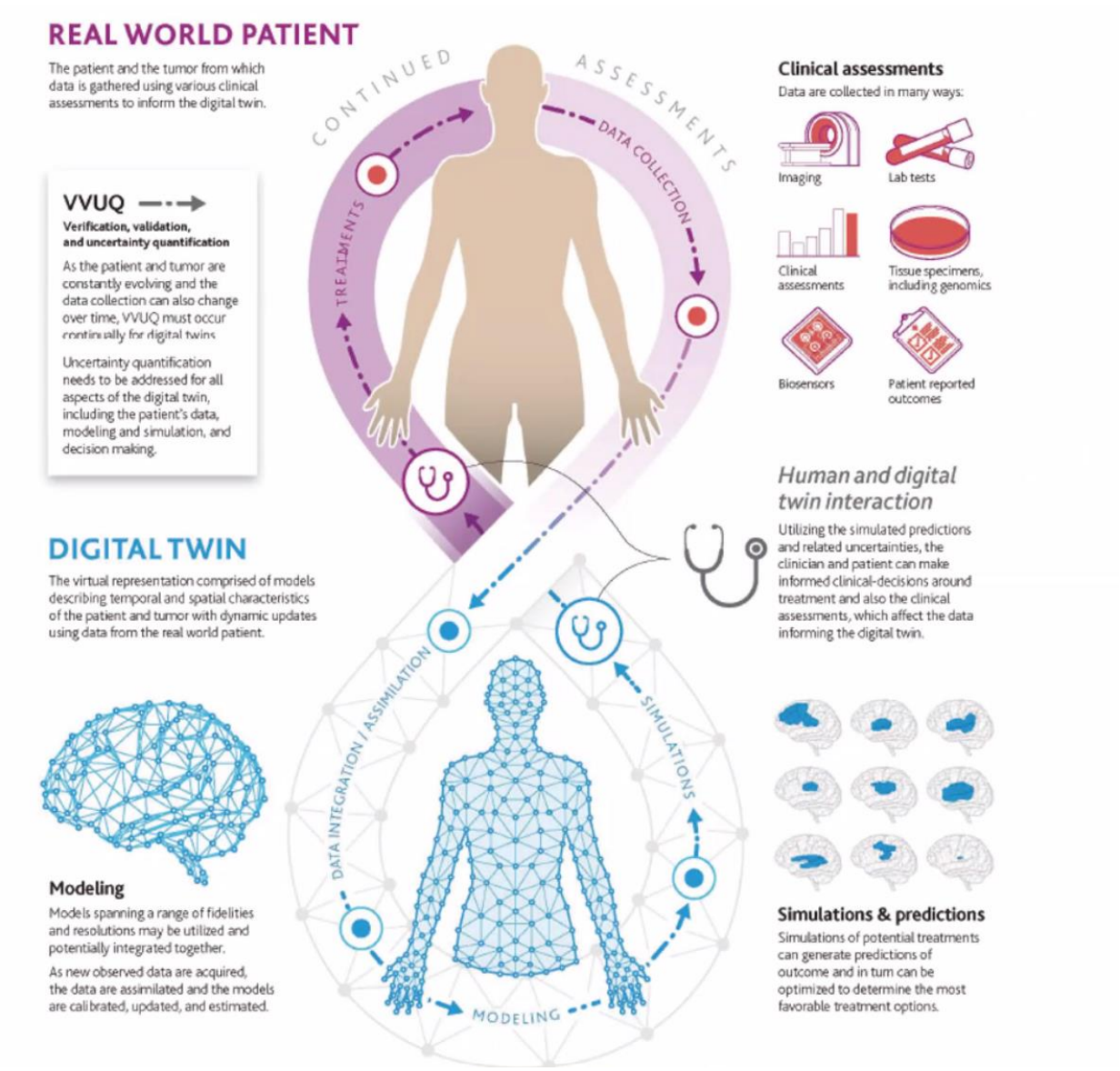
surface glycoprotein [Severe acute respiratory syndrome coronavirus 2]

NCBI Reference Sequence: YP_009724390.1

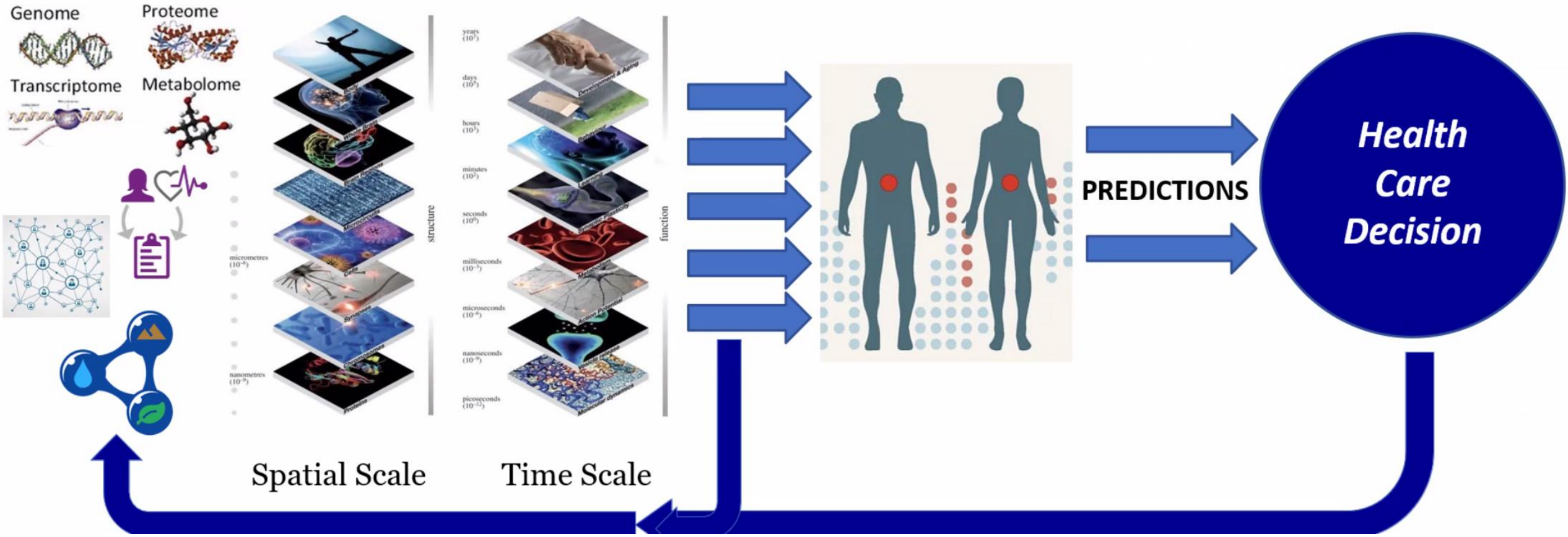
GenPept | Identical Proteins | Graphics

```
>YP_009724390.1 surface glycoprotein [Severe acute respiratory syndrome coronavirus 2]
MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHV
SGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPF
LGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPI
NLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYN
ENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASV
YAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIAD
YNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYF
PLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFL
PFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLT
PTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLG
AENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGI
AVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDC
LGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIG
VTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLND
LSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLM
SFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNT
FVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVA
KNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDD
SEPVLKGVKLHYT
```

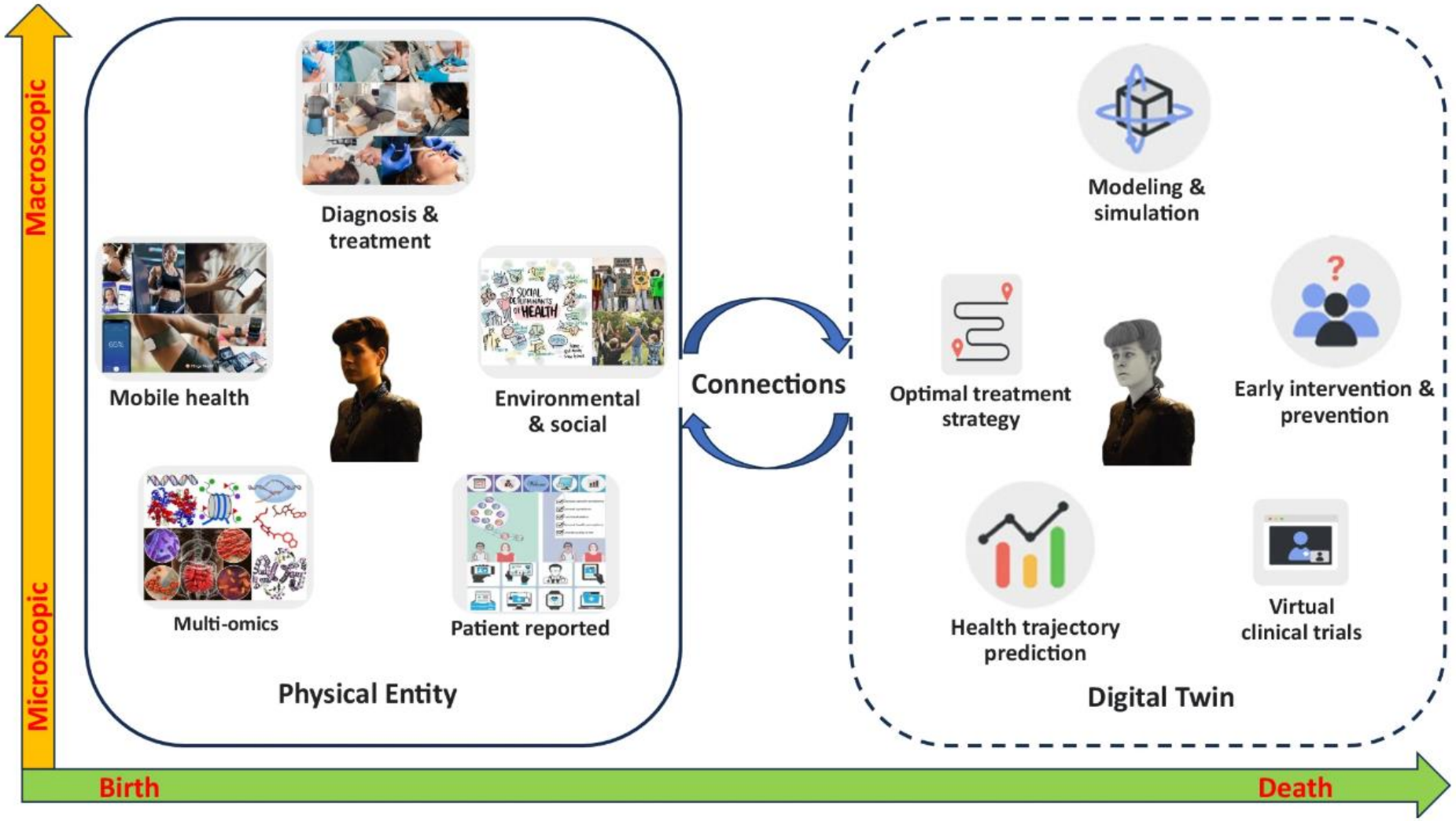
Digital twin



Digital twin

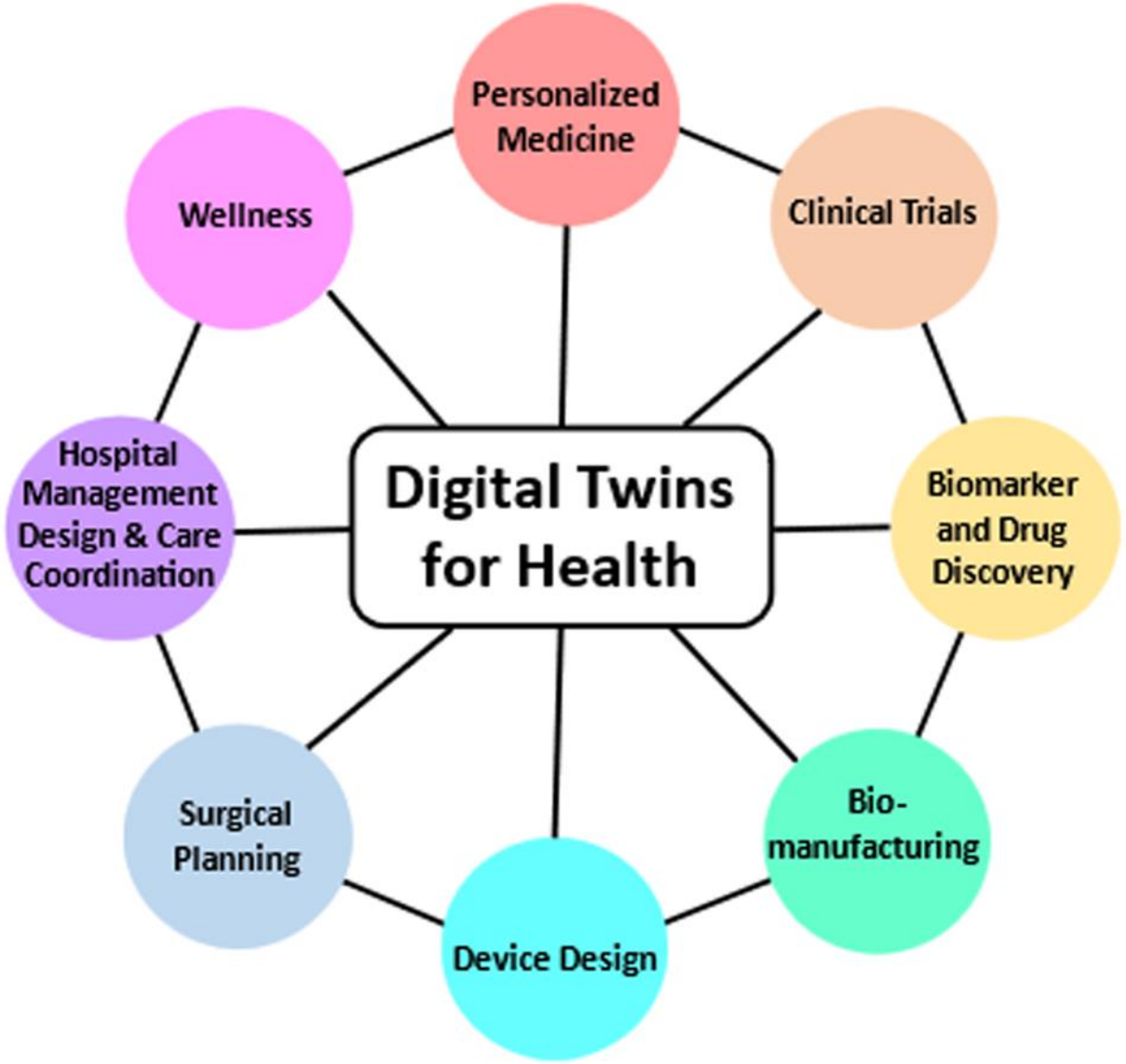


Digital twin



Katsoulakis et al., *npj Digit Med*, 2024

Digital twin



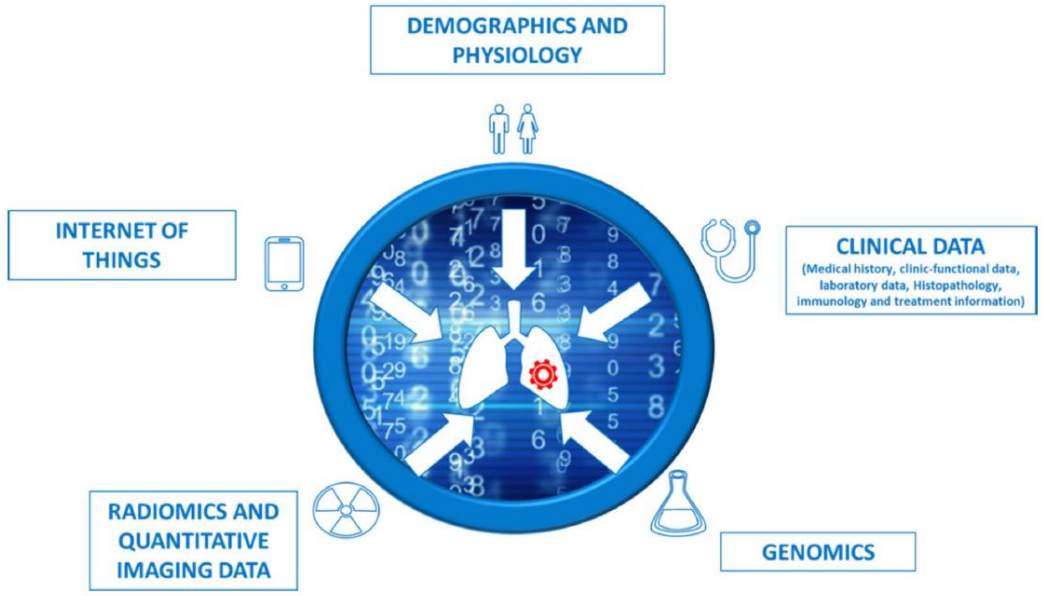
Katsoulakis et al., *npj Digit Med*, 2024

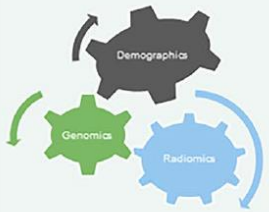
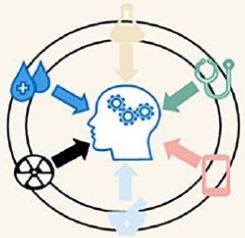
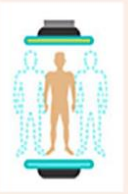
Digital twin

STUDY PROTOCOL Open Access

Lung cancer multi-omics digital human avatars for integrating precision medicine into clinical practice: the LANTERN study

Filippo Lococo[1,2*†], Luca Boldrini[1,3†], Charles-Davies Diepriye[3], Jessica Evangelista[1,2], Camilla Nero[1,4], Sara Flamini[2], Angelo Minucci[1,5], Elisa De Paolis[5,6], Emanuele Vita[1,7], Alfredo Cesario[1,8,9], Salvatore Annunziata[1,10], Maria Lucia Calcagni[1,10], Marco Chiappetta[1,2], Alessandra Cancellieri[1,11], Anna Rita Larici[1,12], Giuseppe Cicchetti[1,12], Esther G.C. Troost[13,14,15,16,17], Róza Ádány[18], Núria Farré[19], Ece Öztürk[20], Dominique Van Doorne[21], Fausto Leoncini[1,22], Andrea Urbani[1,6,23], Rocco Trisolini[1,22], Emilio Bria[1,7], Alessandro Giordano[1,10], Guido Rindi[1,11], Evis Sala[1,12], Giampaolo Tortora[1,7], Vincenzo Valentini[1,3], Stefania Boccia[1,24], Stefano Margaritora[1,2†] and Giovanni Scambia[1,4†]



WP1	WP2	WP3	WP4
Omics data collection	Omics data archiving and inter-actionability	Omics data modelling, Digital Human avatar (DHA) creation and Validation	DHA Implementation in healthcare
Coordinating Team: FPG Supporting Teams: DEB, HSCSP, TUD, KU	Coordinating Team: TUD Supporting Teams: FPG, HSCSP, KU	Coordinating Team: TUD Supporting Teams: FPG, HSCSP, KU	Coordinating Team: AdPEE supporting Teams: DEB, FPG
T1.1: Ethical Protocol preparation and approval in all the recruiting centres [1st Month] T1.2: Clinical study implementation [Month 2-30]	T2.1: Data Archiving QA [Month 2-30] T2.2: Inter-actionability QA: [Month 2-30]	Task 3.1. Omics features identification and selection: [Month12-23] Task 3.2. Predictive model development DHA creation: [Month 15-26] Task 3.3. Predictive model and DHA validation [Month 21-32]	T4.1: Stakeholder engagement, with particular reference to Patients Involvement [Month 1-36] T4.2: ELSI and HTA assessment [Month 12-36] T4.3: Dissemination of project results [Month 18-36] T4.4: Educational and knowledge transfer events [Month 18-33]
			
Month 1-30	Month 2-30	Month 12-32	Month 1-36

Lococo et al., *BMC Cancer*, 2023

Digital twin

ORIGINAL ARTICLE

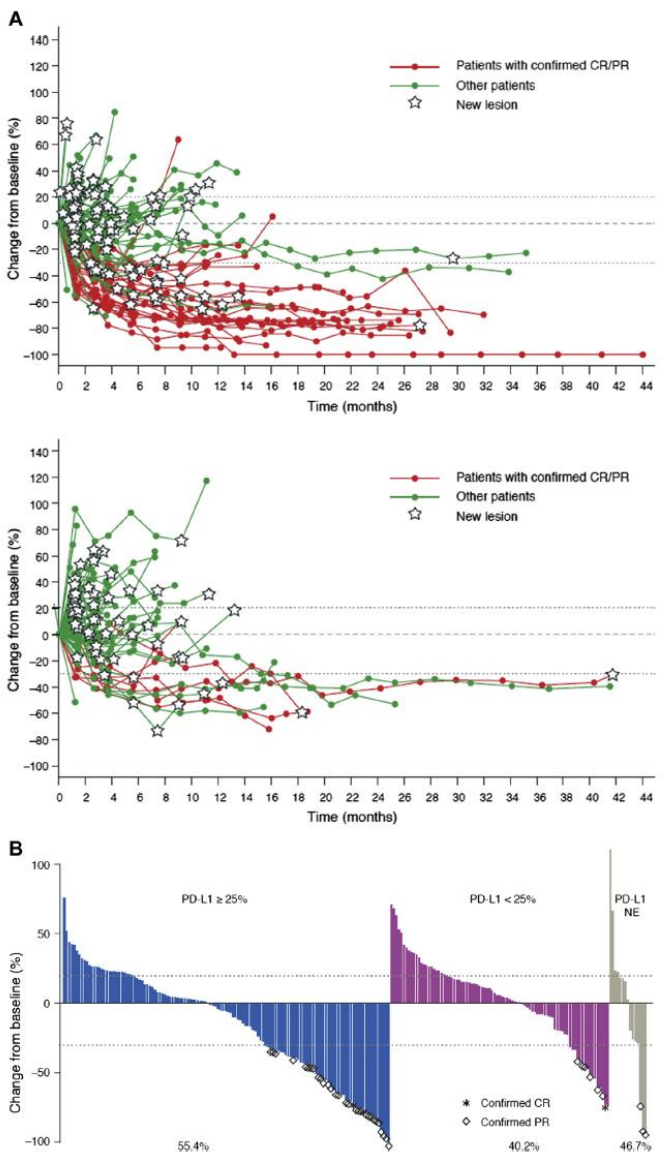
IASLC

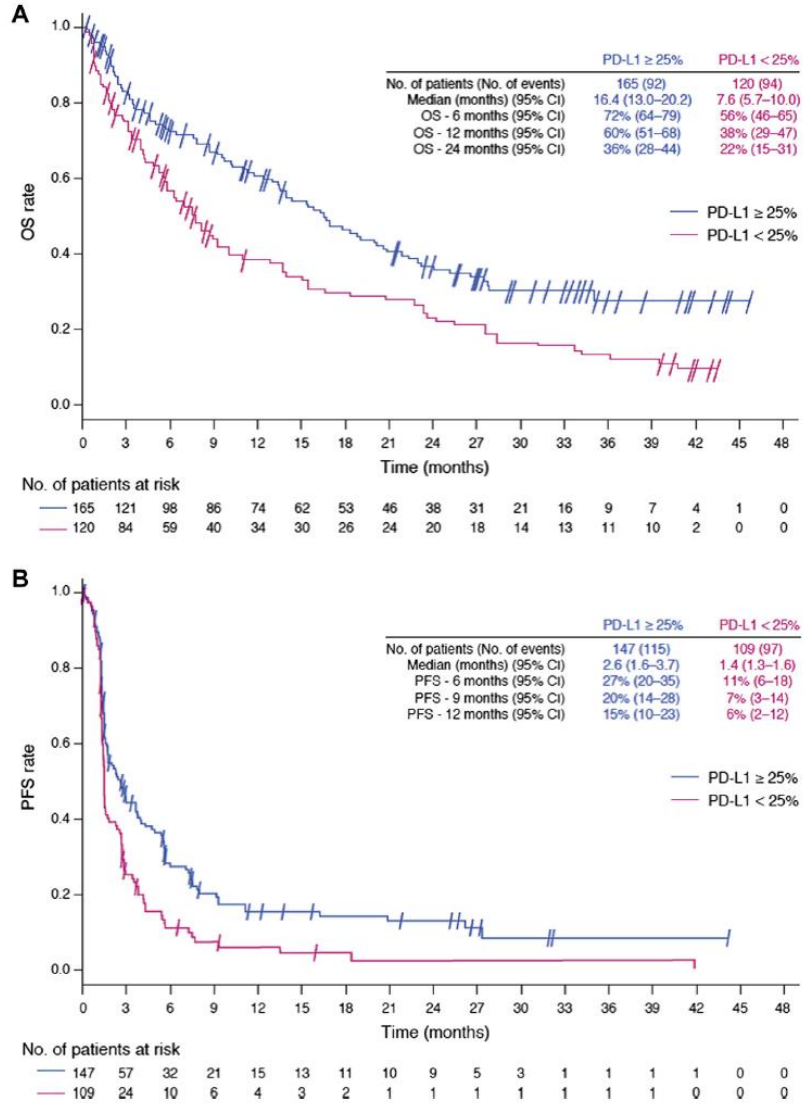
Clinical Activity, Tolerability, and Long-Term Follow-Up of Durvalumab in Patients With Advanced NSCLC

Check for updates

Scott J. Antonia, MD,[a,*] Ani Balmanoukian, MD,[b] Julie Brahmer, MD,[c] Sai-Hong I. Ou, MD, PhD,[d] Matthew D. Hellmann, MD,[e] Sang-We Kim, MD, PhD,[f] Myung-Ju Ahn, MD,[g] Dong-Wan Kim, MD, PhD,[h] Martin Gutierrez, MD,[i] Stephen V. Liu, MD,[j] Patrick Schöffski, MD, MPH,[k] Dirk Jäger, MD,[l] Rahima Jamal, MD, BSc, FRCPC,[m] Guy Jerusalem, MD, PhD,[n] Jose Lutzky, MD, FACP,[o] John Nemunaitis, MD,[p] Luana Calabrò, MD,[q] Jared Weiss, MD,[r] Shirish Gadgeel, MD,[s] Jaishree Bhosle, MD,[t] Paolo A. Ascierto, MD,[u] Marlon C. Rebelatto, PhD,[v] Rajesh Narwal, PhD, MS,[v] Meina Liang, BS,[v] Feng Xiao, BS,[v] Joyce Antal, MS,[v] Shaad Abdullah, MD, FACP,[v] Natasha Angra, PharmD,[v] Ashok K. Gupta, MBBS, MD, PhD,[v] Samir N. Khleif, MD,[w] Neil H. Segal, MD, PhD[e]

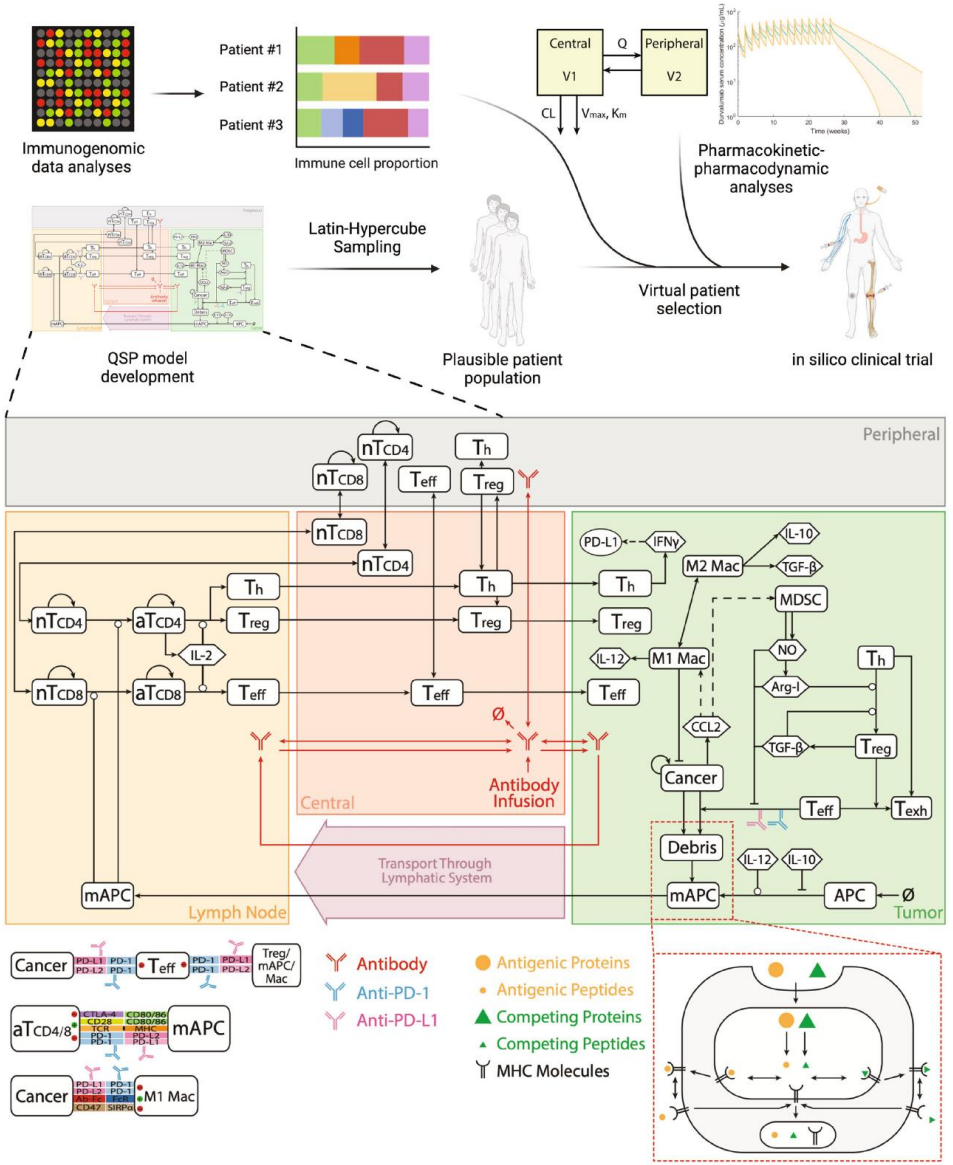
[a]*Moffitt Cancer Center, Tampa, Florida*
[b]*The Angeles Clinic and Research Institute, Los Angeles, California*
[c]*Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins, Baltimore, Maryland*
[d]*Department of Medicine, Division of Hematology-Medical Oncology, University of California Irvine School of Medicine, Irvine, California*
[e]*Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, New York*
[f]*Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea*
[s]*Sunkyunkwan University School of Medicine, Samsung Medical Center, Seoul, South Korea*

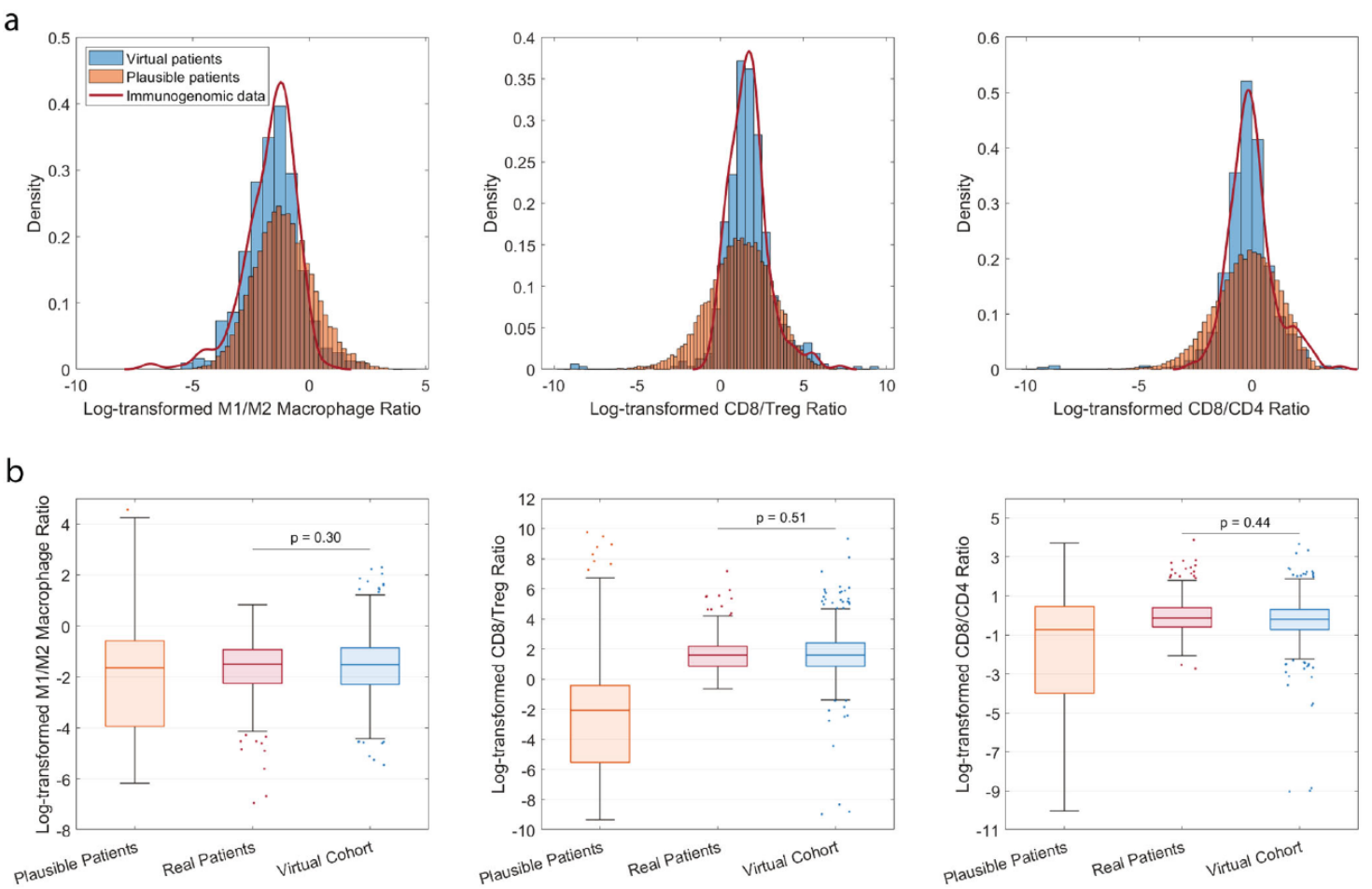




Antonia et al., *J Thorac Oncol*, 2019

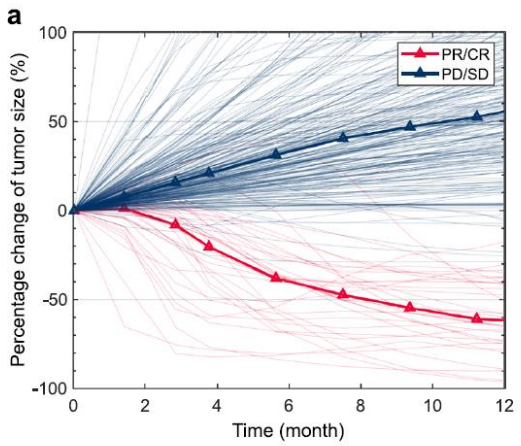
Digital twin

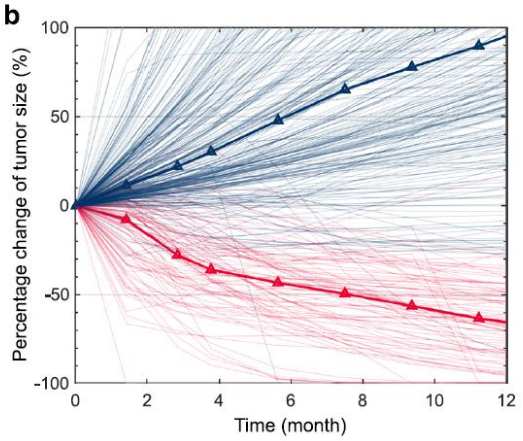


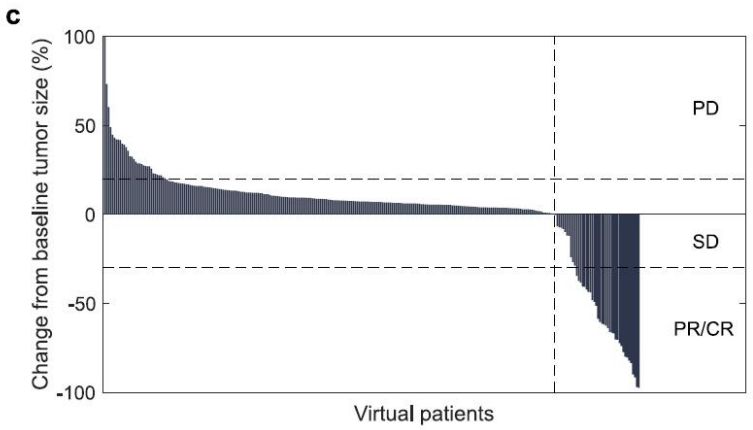


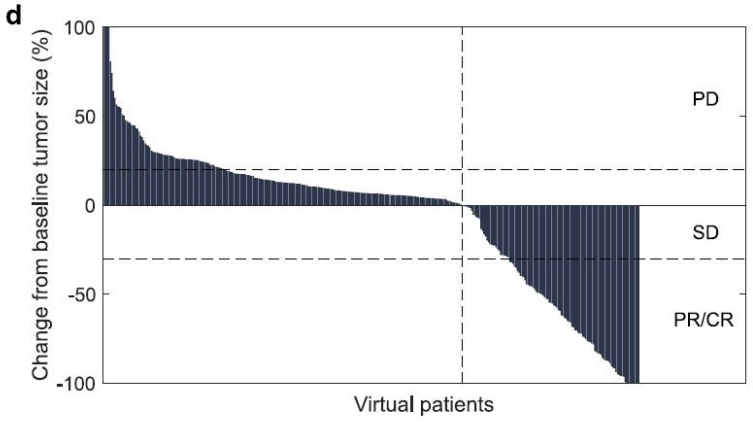
Digital twin

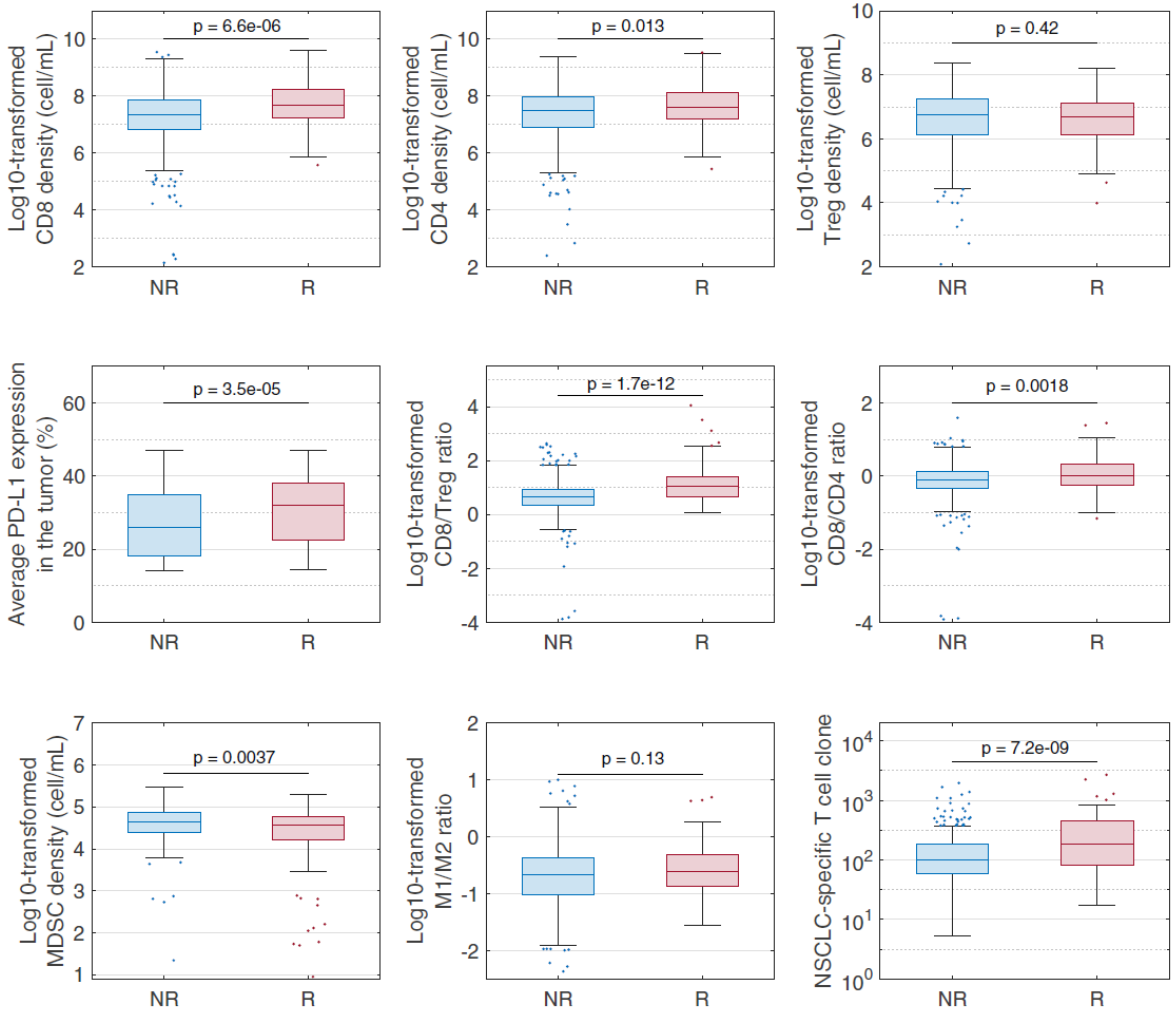
In silico clinical trial simulation











Digital twin

서울대학교 암연구소
Seoul National University Cancer Research Institute

암연구소 소개 | 대학원 | 연구 | 예약신청 | 암정보교육관 | 커뮤니티 | 발전기금

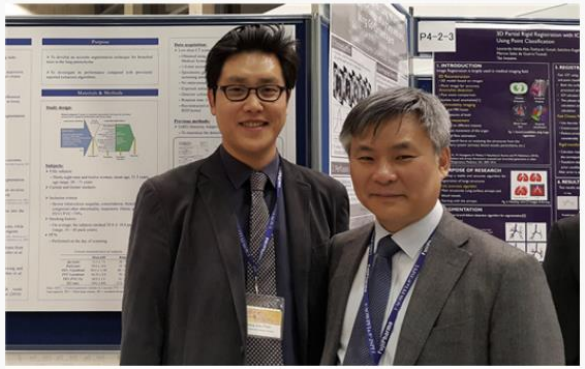
암연구소 소개 | 연구실 소개

인공지능 디지털 트윈 연구실

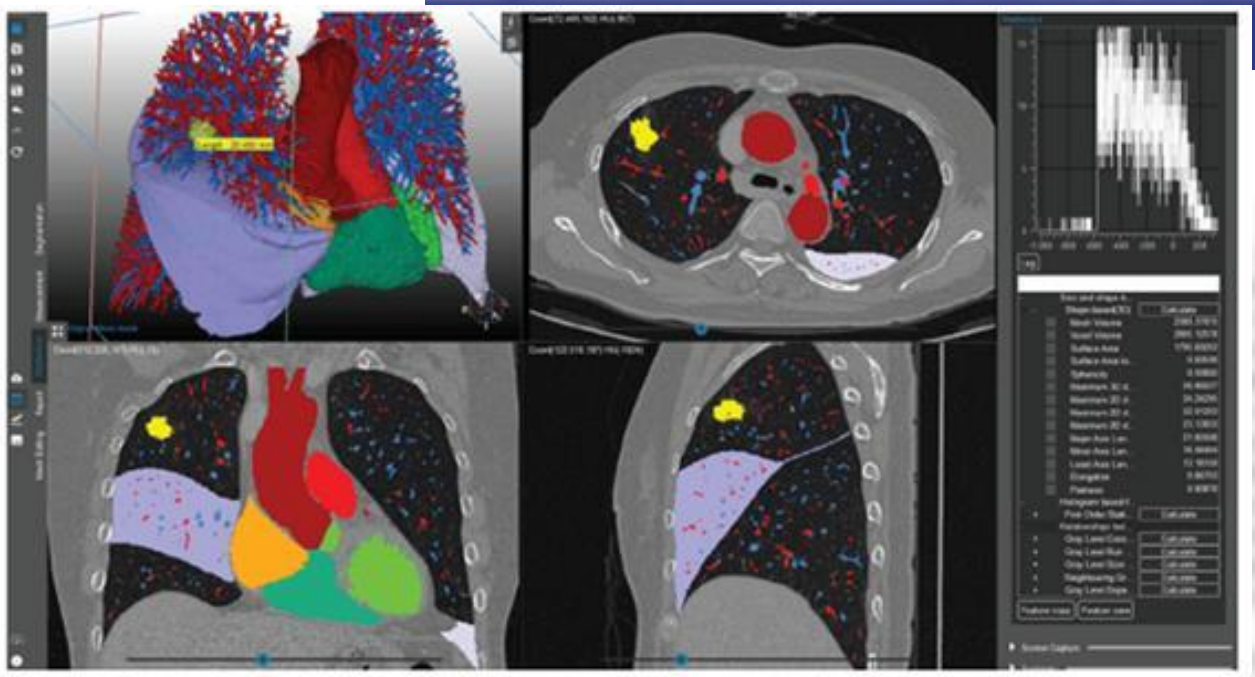
Artificial Intelligence Digital Twin Lab (AIDTL)

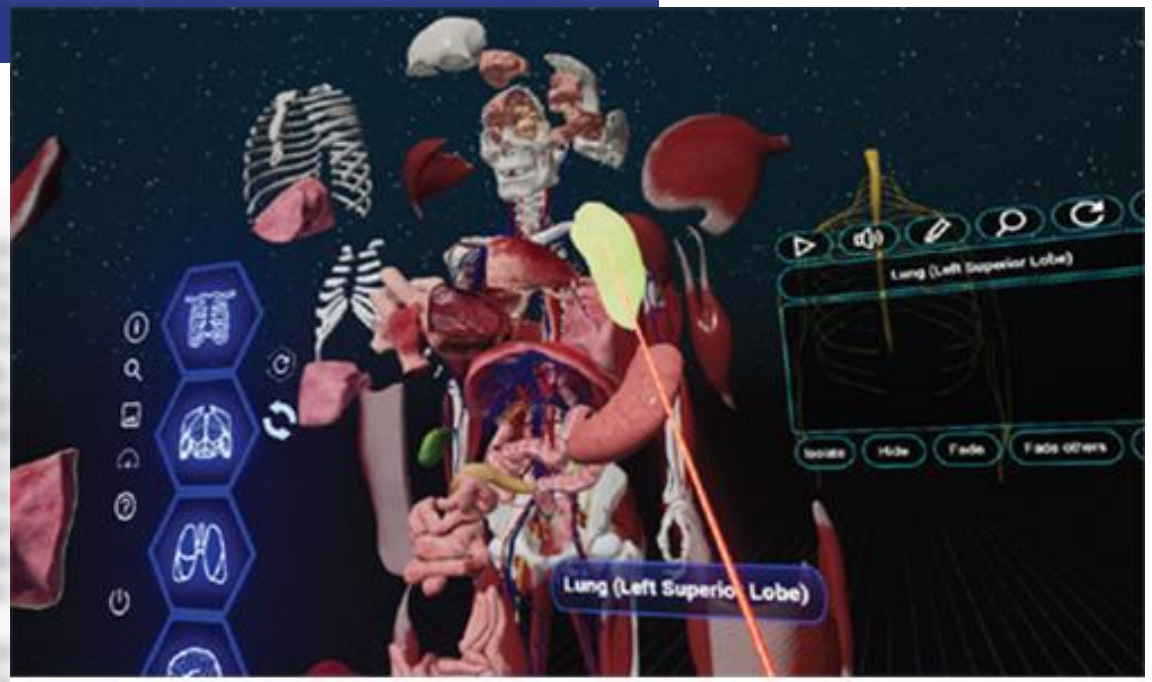
디지털 트윈 시장은 2025년까지 295억 7,000만 달러(약 30조원)로 성장할 것으로 예상되었으며, 글로벌 조사기관 가트너(Gartner) 또한 디지털 트윈을 `17~'19년 3년 연속 디지털 변혁시대를 주목해야 10대 전략기술로 선정하였다. 성공사례의 등장과 기술성숙 등으로 하이프사이클의 '기대 정점(Peak of Inflated Expectations)'에 위치하여 향후 10년 내 본격 시장 확산이 전망되고 있다. 특히 의학 분야에서의 디지털 트윈은 3차원 가시화, AR, VR, 메타버스 등의 신기술들과 만나며 교육과 시뮬레이션 등에서 혁신적 기회를 창출할 수 있을 것이라 기대되고 있다. 본 연구실은 이러한 학내의 미충족 필요성에 힘입어 2011년 9월 창설되었으며, 신종감염병 및 암정복 연구의 새로운 시대를 열기위해 활발한 연구를 진행하고 있다. 특히 인공지능, AR, VR, 메타버스 등의 첨단 기술을 활용해 임상에 적용할 수 있는 가치 실현을 목표로 연구활동이 활발히 이루어지고 있으며, 서울대학교병원의 다양한 연구자들과 여러 형태의 공동연구를 진행하고 있다. 인공지능 디지털 트윈 연구실은 명실상부 미래 의료IT기술 연구실로서 서울대학교병원 내 데이터 분석 기술 개발 및 임상 적용에 있어 중심적 역할을 해내는 중이다.

또한 본 연구실은 SW연구개발실, AI연구실, Deep Learning Training실 등이 유기적으로 한 공간에 배치되어 여러 연구진들이 실시간으로 소통하고 있다. 이를 통해 인공지능 연구의 효율성을 극대화할 수 있으며, 참여자들 또한 쉽게 이해하고 참여하여 관련 분야에 대한 빠른 이해가 가능하도록 구성되어 있다.



Digital twin





Summary

- Hallmarks of AI in Lung cancer research
  - Pattern recognition
  - Sensitivity to subtle changes

- Application of AI into Lung cancer research
  - Clinical research: Radiomics, Pathologic data
  - Multiomics data: Spatial transcriptomics
  - Discovery of treatment target: Synthetic lethality
  - Prediction of protein structure: Alphafold
  - Digital twin

Thank you for your attention