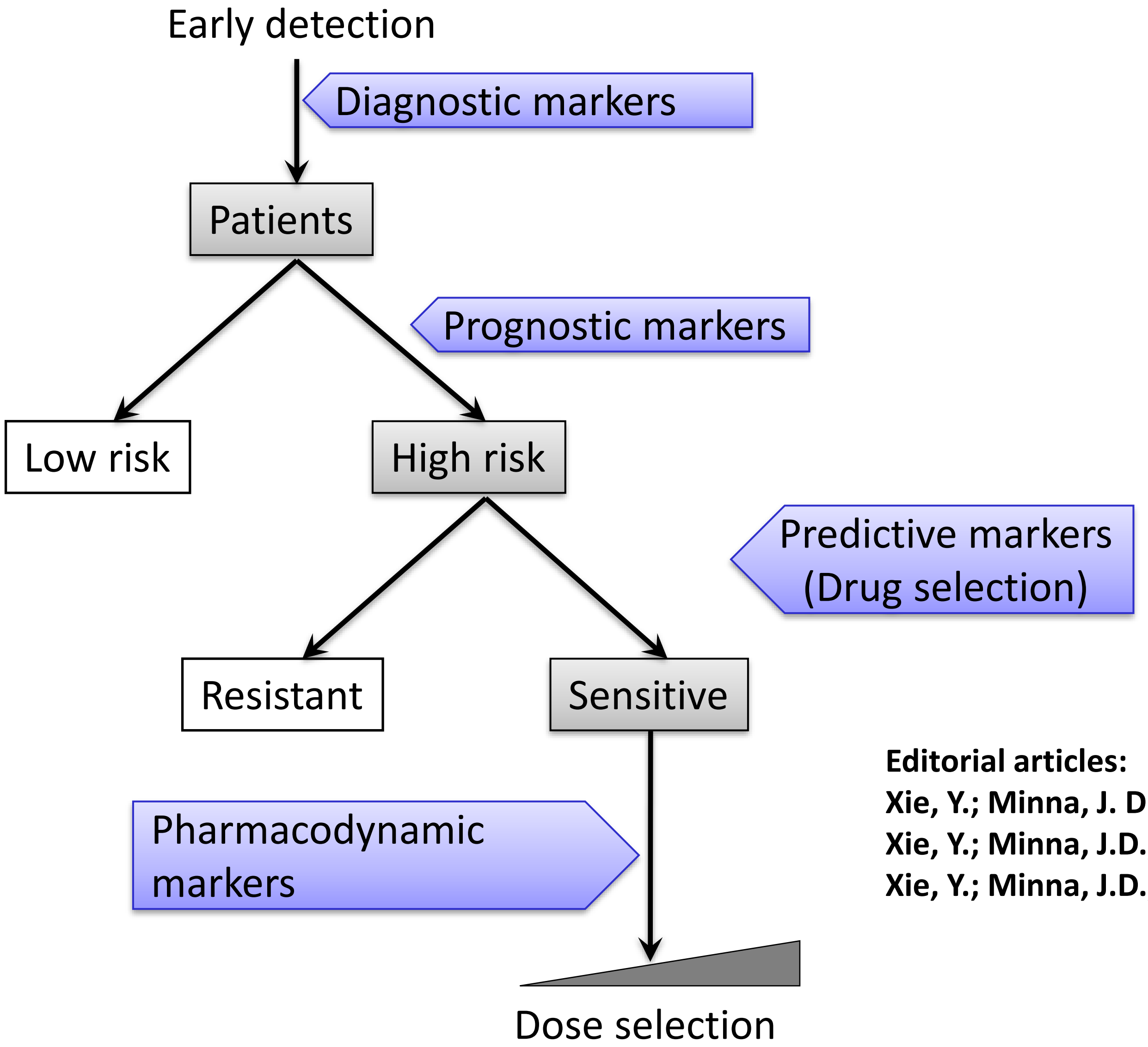
AI in Precision Health

Yang Xie, Ph.D, M.D.
Professor & Associate Dean of Data Science
Raymond D. and Patsy R. Nasher Distinguished Chair in Cancer Research
Chair, Department of Health Data Science and Biostatistics
Peter O’Donnell Jr. School of Public Health
UT Southwestern Medical Center at Dallas

2025 UT System AI Symposium in Health Care
Houston, Texas
May 15, 2025

UTSouthwestern
Medical Center

AI Can Identify Different Types of Biomarkers for Precision Health



Editorial articles:
Xie, Y.; Minna, J. D., Nat Med, 2008
Xie, Y.; Minna, J.D., JCO, 2010
Xie, Y.; Minna, J.D., Lancet, 2012

Genomic based Precision Health

Genomic signature for treatment response prediction (Adjuvant Genomics)

PREDICTIVE BIOMARKERS AND PERSONALIZED MEDICINE | MARCH 14 2013

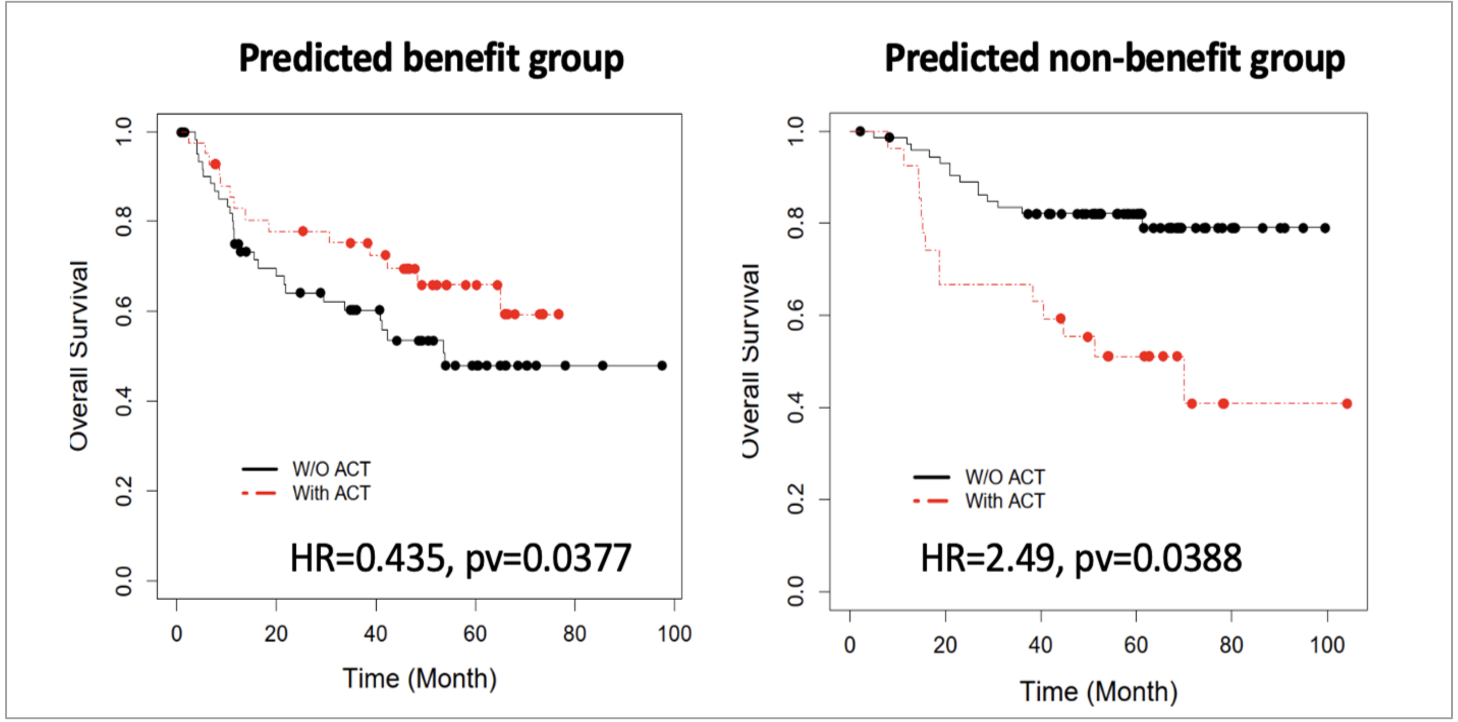
A 12-Gene Set Predicts Survival Benefits from Adjuvant Chemotherapy in Non–Small Cell Lung Cancer Patients

Hao Tang; Guanghua Xiao; Carmen Behrens; Joan Schiller; Jeffrey Allen; Chi-Wan Chow; Milind Suraokar; Alejandro Corvalan; Jianhua Mao; Michael A. White; Ignacio I. Wistuba; John D. Minna; Yang Xie

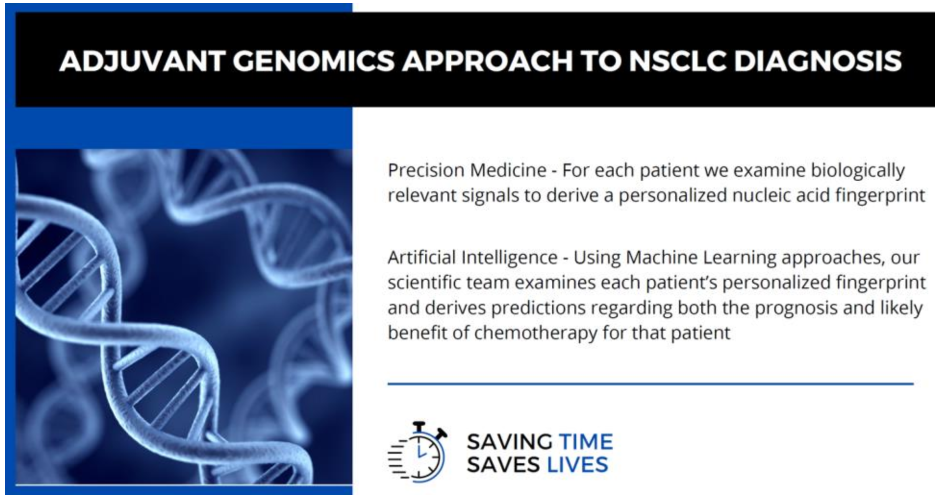
PRECISION MEDICINE AND IMAGING | JANUARY 03 2019

Validation of the 12-gene Predictive Signature for Adjuvant Chemotherapy Response in Lung Cancer

Yang Xie; Wei Lu; Shidan Wang; Ximing Tang; Hao Tang; Yunyun Zhou; Cesar Moran; Carmen Behrens; Jack A. Roth; Qinghua Zhou; David H. Johnson; Stephen G. Swisher; John V. Heymach; Vassiliki A. Papadimitrakopoulou; Guanghua Xiao; John D. Minna; Ignacio I. Wistuba



Predictive performance in MD Anderson dataset (n=207)
(Similar validation in three independent cohorts)



ADJUVANT GENOMICS APPROACH TO NSCLC DIAGNOSIS

Precision Medicine - For each patient we examine biologically relevant signals to derive a personalized nucleic acid fingerprint

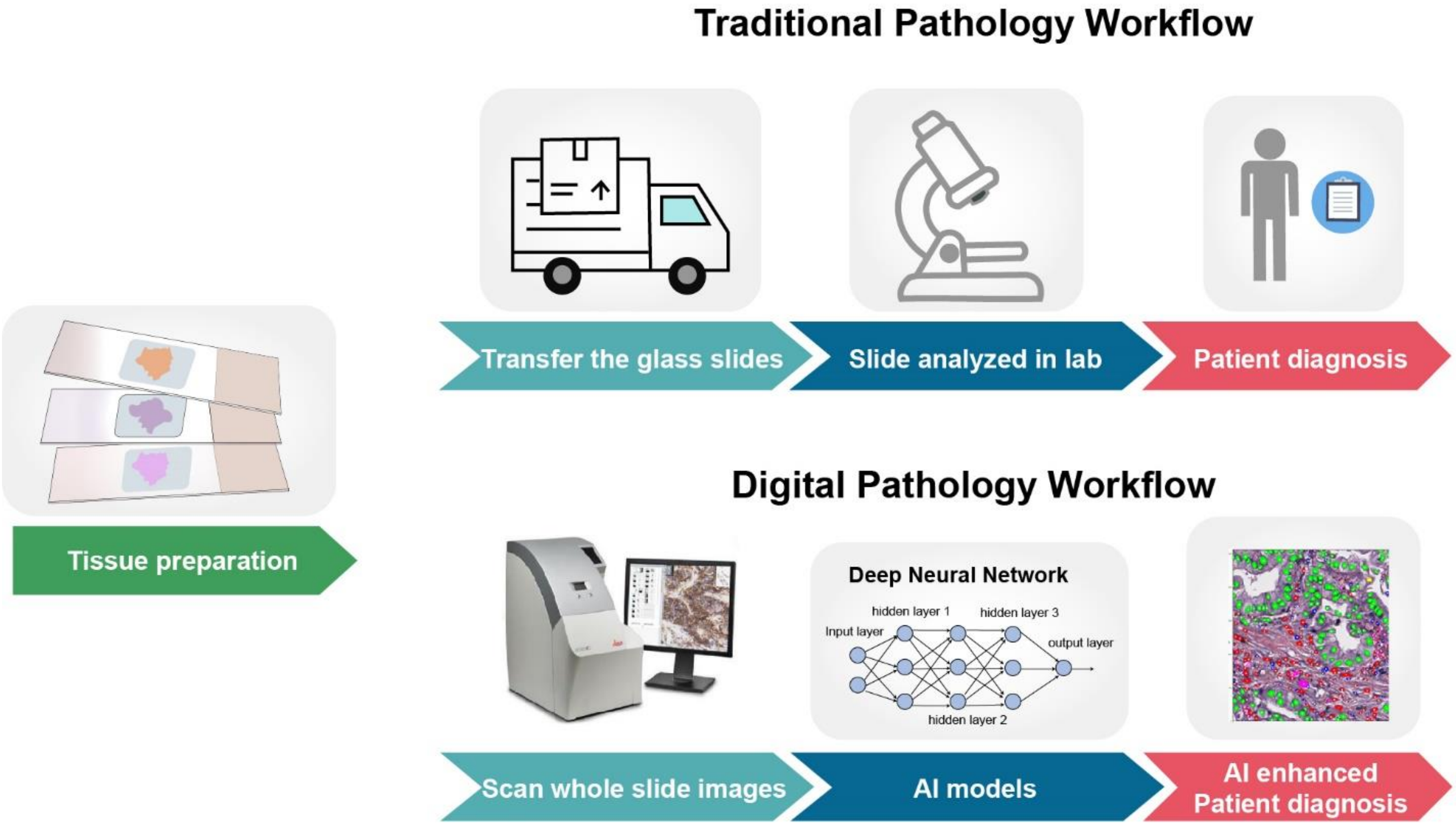
Artificial Intelligence - Using Machine Learning approaches, our scientific team examines each patient's personalized fingerprint and derives predictions regarding both the prognosis and likely benefit of chemotherapy for that patient

SAVING TIME
SAVES LIVES

A clinical assay has been produced and is under clinical validations

UTSouthwestern
Medical Center

Digital Pathology



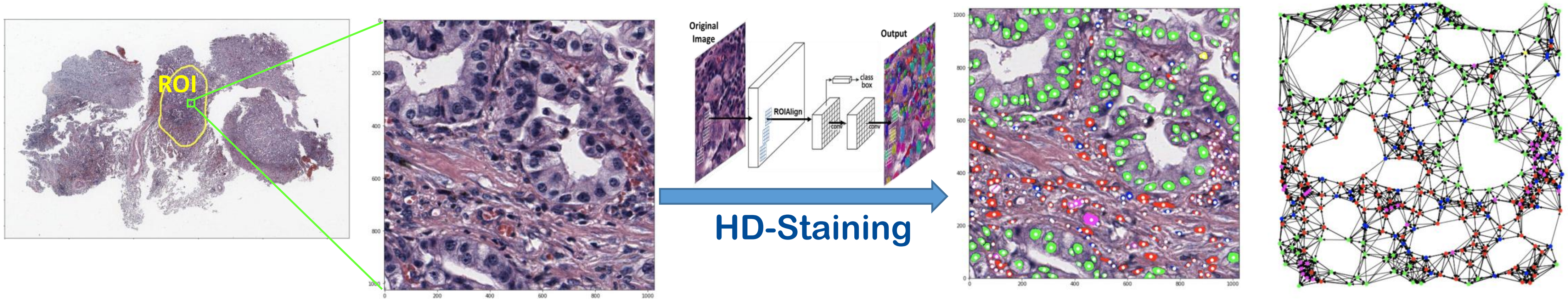
Traditional Pathology Workflow

- Tissue preparation
- Transfer the glass slides
- Slide analyzed in lab
- Patient diagnosis

Digital Pathology Workflow

- Tissue preparation
- Scan whole slide images
- AI models
- AI enhanced Patient diagnosis

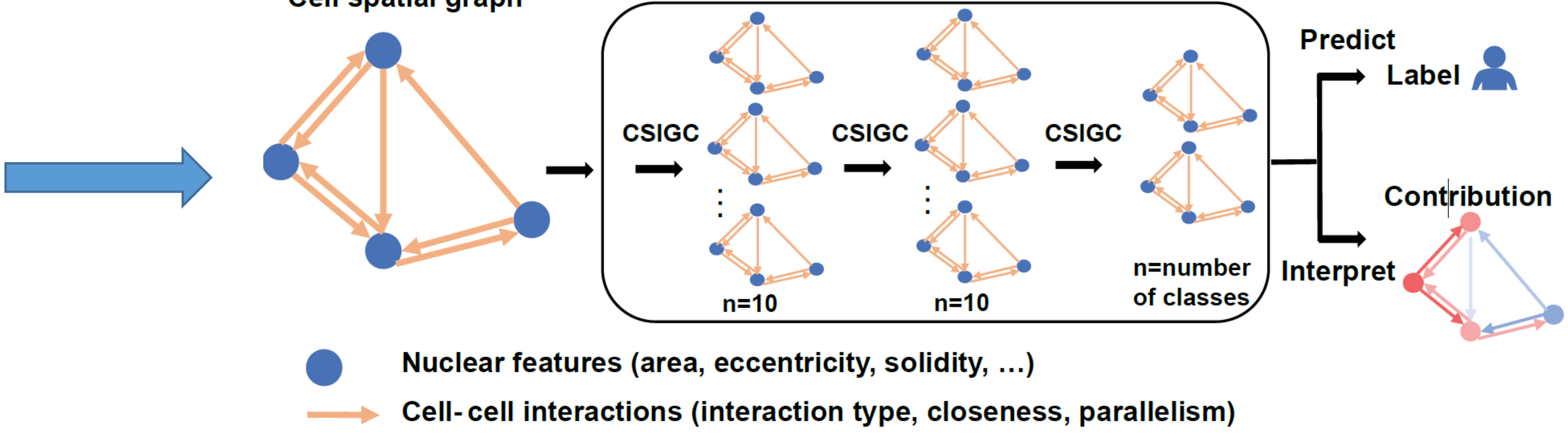
AI Models for Digital Pathology



HD-Staining

🟩 Tumor Nuclei
🟦 Lymphocyte Nuclei
🟥 Stroma Nuclei
🟨 Macrophage Nuclei
🟦 Karyorrhexis
🟪 Red Blood Cells

CeoGraph



● Nuclear features (area, eccentricity, solidity, …)

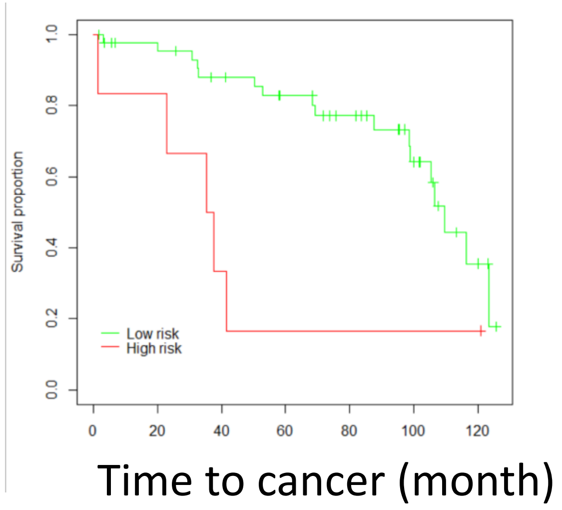
→ Cell- cell interactions (interaction type, closeness, parallelism)

- Faster diagnosis
- More Accurate
- Cost-effective

AI Models for Digital Pathology

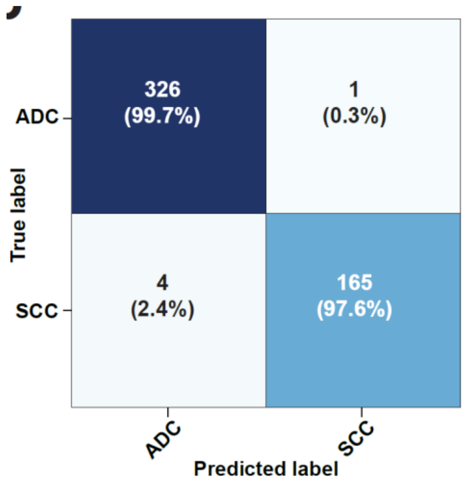
UTSouthwestern
Medical Center

1. Early detection and risk stratification

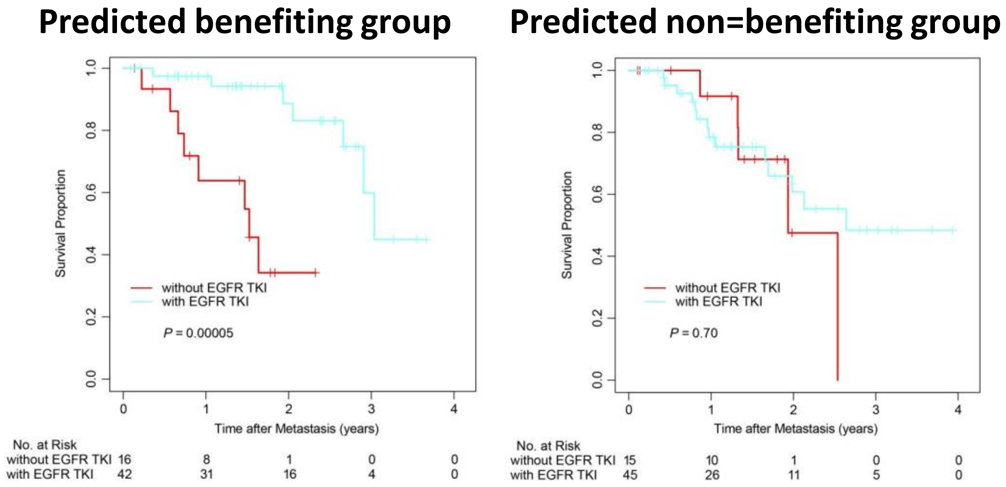


🟩 **Predicted high-risk group**
🟥 **Predicted low-risk group**

2. Disease diagnosis and subtyping



3. Treatment response prediction


Predicted benefiting group

Predicted non=benefiting group

Disease areas:

- Lung cancer
- Breast cancer
- Head & neck cancer
- Oral potentially malignant disorders
- Rhabdomyosarcoma
- Liver cancer

Key publications:

- Wang at al, Cancer Research, 2020
- Zhang et al, Am J Pathol. 2022
- Rong et al, Am J Pathology., 2023
- Wang et al, J. Clin. Invest. , 2023
- Rong et al, Modern Pathology, 2023
- Wang et al, Nature Communications, 2023

AI-assisted Design of T Cell Receptors (TCRs)

pMTnet : predicts the binding of T cell receptors (TCR) with TCR targets using AI approaches

In silico

- pMTnet and associated technologies for TCR design (IP protected)*
- Proprietary database of TCR-antigen binding relationships

Empirical validation

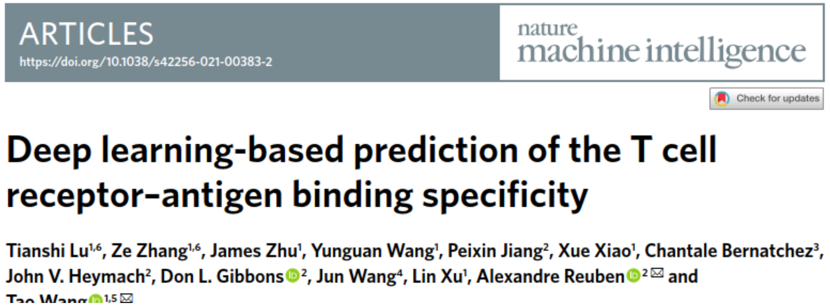
- Extensive library of TCR targets and corresponding TCRs available to enhance pMTnet
- Functional validation of enhanced TCRs
- Established workflow to isolate ultra-rare TCRs

Real-time feedback to improve in silico assets

TCRs predicted and validated iteratively in-house

Licensable TCRs

*PCT nationalized/filed in the US in March 2023



ARTICLES
https://doi.org/10.1038/s42256-021-00383-2

nature machine intelligence

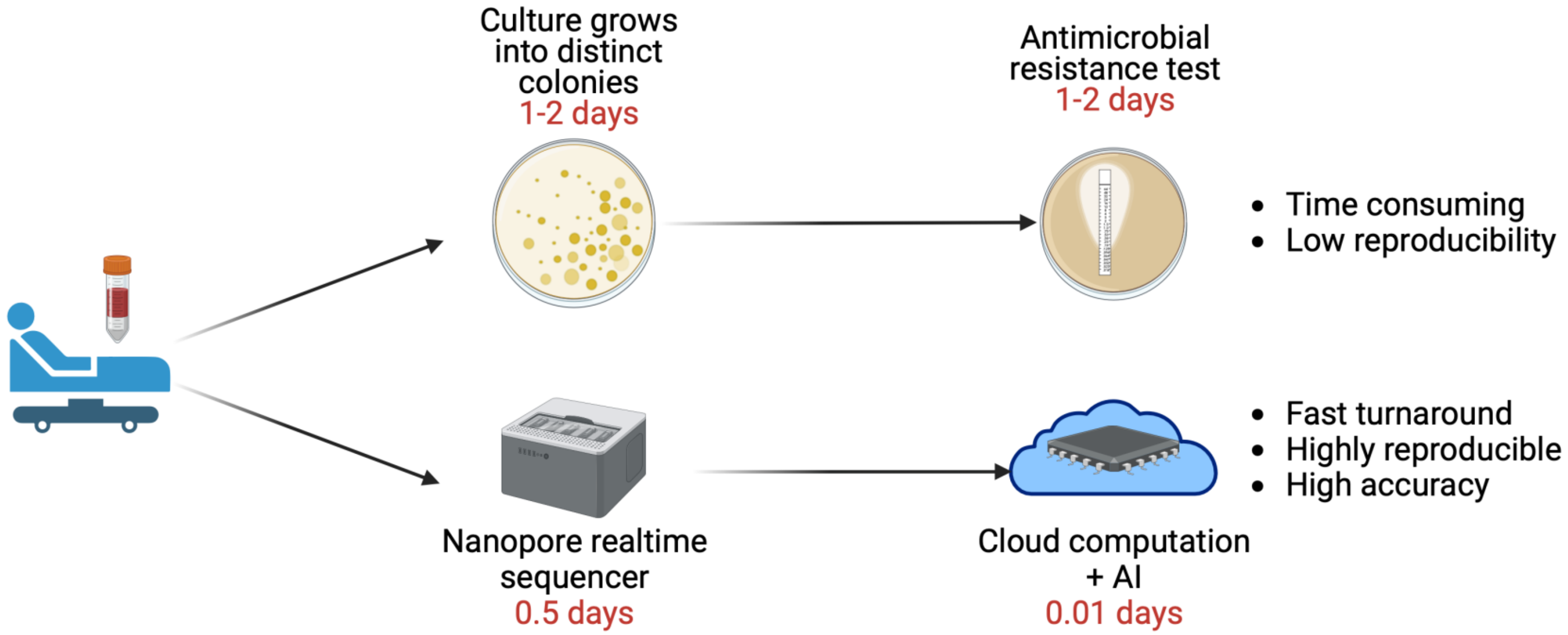
Check for updates

Deep learning-based prediction of the T cell receptor–antigen binding specificity

Tianshi Lu[1,6], Ze Zhang[1,6], James Zhu[1], Yunguan Wang[1], Peixin Jiang[2], Xue Xiao[1], Chantale Bernatchez[3], John V. Heymach[2], Don L. Gibbons[2], Jun Wang[4], Lin Xu[1], Alexandre Reuben[2] ✉ and Tao Wang[1,5] ✉

AI-assisted Antimicrobial Resistance Tests

PARP (Pan-Antibiotic Resistance Prediction) improves antimicrobial resistance tests



Culture grows into distinct colonies
1-2 days

Antimicrobial resistance test
1-2 days

- Time consuming
- Low reproducibility

Nanopore realtime sequencer
0.5 days

Cloud computation + AI
0.01 days

- Fast turnaround
- Highly reproducible
- High accuracy

PLOS COMPUTATIONAL BIOLOGY

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

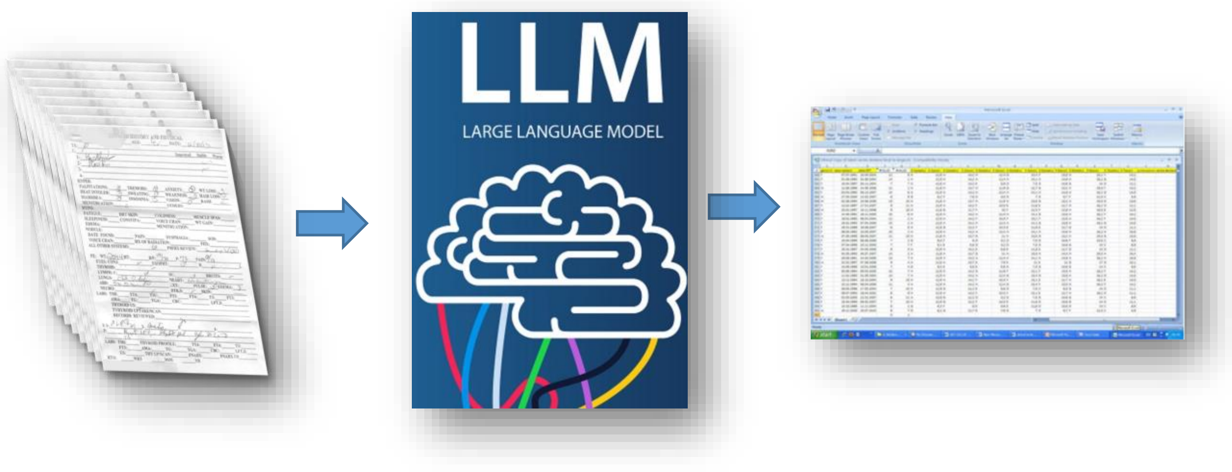
VAMPr: VAriant Mapping and Prediction of antibiotic resistance via explainable features and machine learning

Jiwoong Kim, David E. Greenberg, Reed Pifer, Shuang Jiang, Guanghua Xiao, Samuel A. Shelburne, Andrew Koh, Yang Xie, Xiaowei Zhan

Large Language Model: Extract Information from Clinical Notes

UTSouthwestern Medical Center

Prompt engineering design



A critical assessment of using ChatGPT for extracting structured data from clinical notes

Overall performance of ChatGPT on data extraction from pathology reports

Attribute	Accuracy	F1	Kappa	Recall	Precision	Coverage
primary tumor features (pT)	0.87	0.87	0.76	0.87	0.89	0.97
regional lymph node involvement (pN)	0.91	0.91	0.84	0.91	0.92	0.94
overall tumor stage	0.76	0.76	0.61	0.76	0.77	0.94
histological diagnosis	0.99	0.99	0.98	0.99	0.99	0.96
Average	0.89	0.88	0.80	0.89	0.89	0.95

Huang et al, npj Digital Medicine, 2024

Data Integration Platform

Isolated silos

Healthcare
UT Southwestern

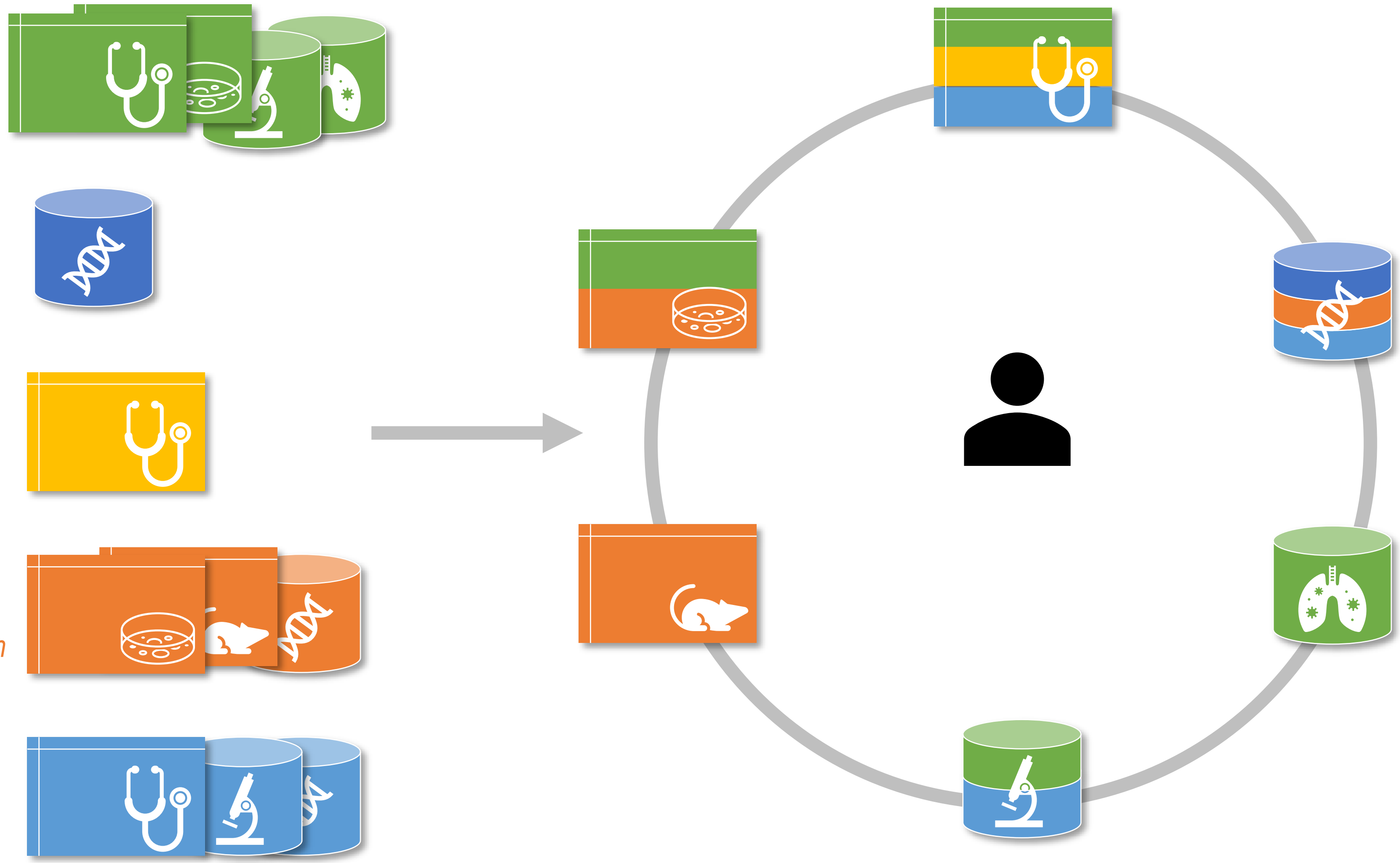
Testing facility
Foundation Medicine

Research lab
Clinical research

Research lab
Translational research

Public domain
TCGA

Data Commons



Friendly interface

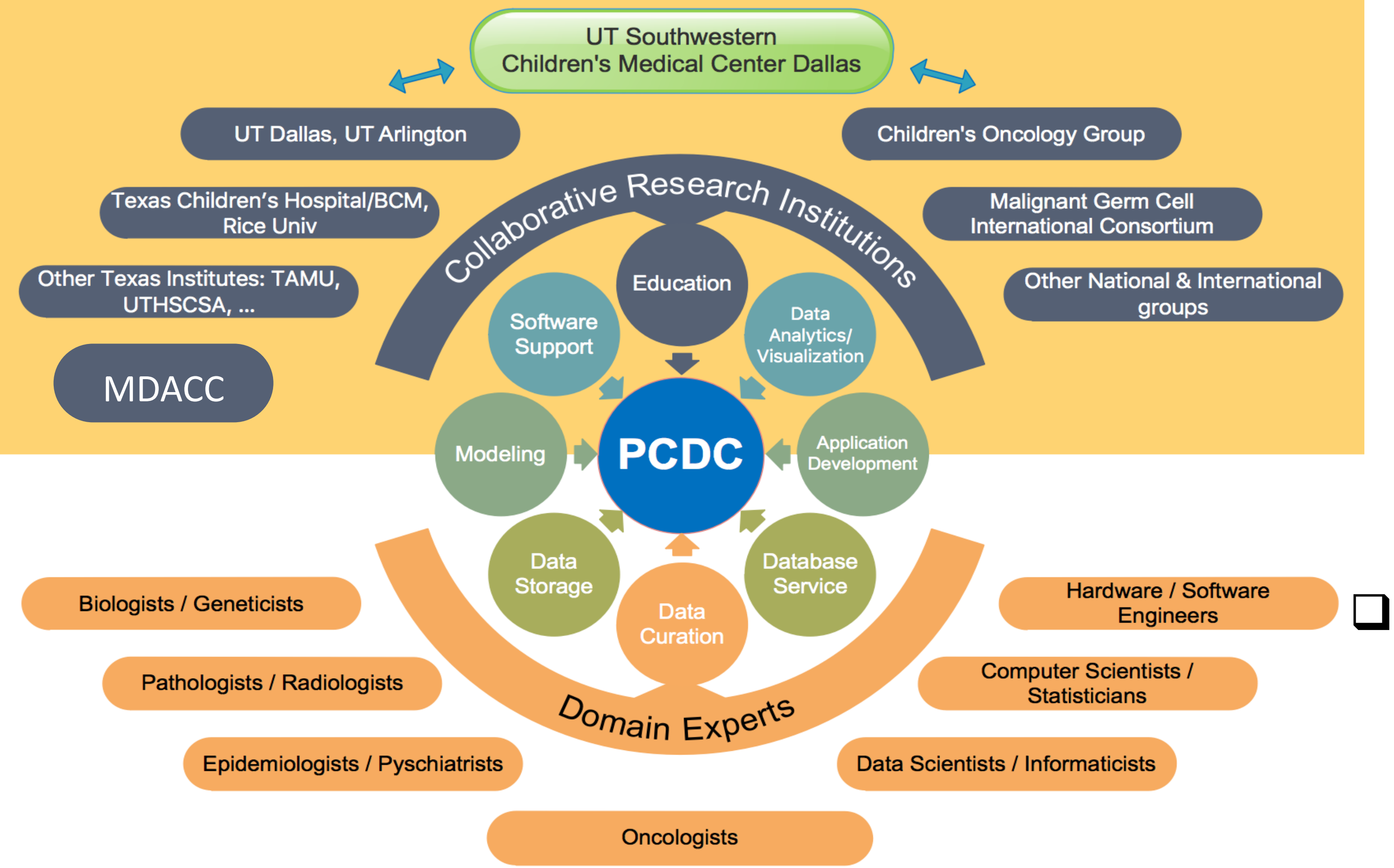
Security

Sharing

Analysis-ready

PCDC: Pediatric Cancer Data Core (CPRIT RP180805)

UTSouthwestern Medical Center



- ☐ **Major goal: To develop Pediatric Cancer Data Commons**

Acknowledgments

UTSouthwestern
Medical Center

Funding

NCI, NIGMS, NIAID, CPRIT, UTSW Endowment

Quantitative Biomedical Research Center (QBRC)

https://qbrc.swmed.edu/