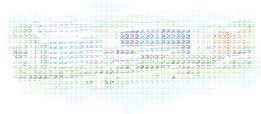


The background features a dark blue circular graphic with concentric lines, resembling a target or a microscope's field of view. In the center, there is a pair of human lungs. The image is overlaid with several microscopic sections of lung tissue, stained in shades of purple and pink, showing cellular structures and possibly tumor cells.

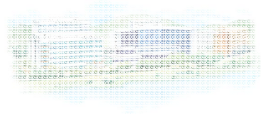
# Next-Generation Pathology: The Present and Future of AI Powered Lung Cancer Pathology

Jung Wook Yang, MD, PhD  
Department of Pathology  
Gyeongsang National University College of Medicine  
Jinju, Republic of Korea

Korean Association of Tuberculosis and Respiratory Diseases – Lung Cancer Symposium, April 2026



**I have no conflicts of interest to disclose.**

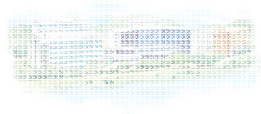


## 1. Why AI in lung cancer pathology now?

## 2. AI in lung cancer pathology: present and future

- Quantitative pathology (IHC)
- Driver mutation prediction from H&E
- Treatment-response prediction
- Prognosis prediction
- Multimodal foundation models
- Pathology large language models

## 3. Take-home messages



## Clinical pressure points

- Increasing histologic and molecular complexity in lung cancer pathology.
- Expanded therapeutic biomarkers.
- Need for more precise and complex pathology results.
- Limited biopsy tissue and competing downstream tests.

## What AI can add

- Standardized, reproducible quantitative pathology.
- High-dimensional analysis of morphology and biomarker data



## A. Quantitative pathology

- Automated IHC scoring for Ki-67, PD-L1, TROP2.

## B. Driver mutation prediction

- ALK, EGFR prediction from H&E to prioritize molecular tests.

## C. Immune checkpoint inhibitor (ICI)-response prediction

- Immune phenotyping for immunotherapy.

## D. Prognosis prediction

- WSI-based risk models refining recurrence and survival.

**Key point: AI provides decision support, not autonomous diagnosis.**

# A. Quantitative Pathology – automated IHC scoring

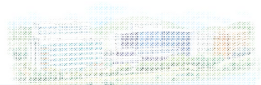
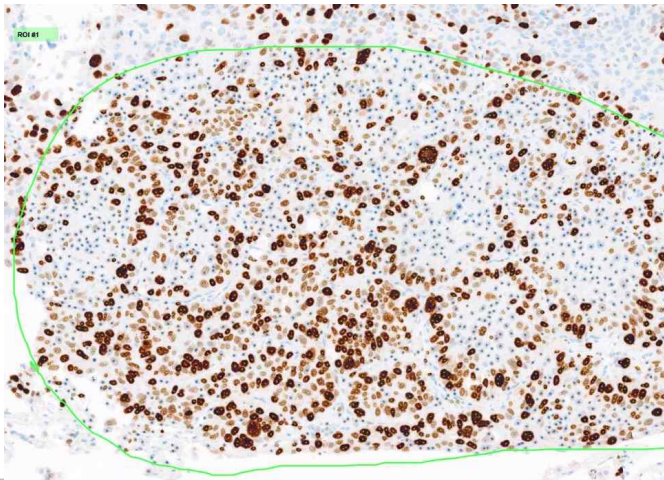


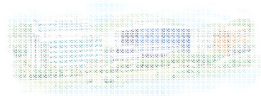
Marker	Clinical role	Example AI/DIA tools	Relevance in lung cancer
Ki-67	Proliferation index; aggressive biology	Roche uPath Ki-67(Breast), AIVIS Quanti Breast Ki-67, and other DIA platforms (ex. QuPath)	Not lung-specific, but technically deployable for reproducible Ki-67 quantification
PD-L1	Immunotherapy (ICI) selection	Roche uPath PD-L1 (SP263) Lunit SCOPE PD-L1 22C3 TPS	<b>Most mature AI-assisted application in lung cancer pathology.</b>
TROP2	ADC (Dato-DXd*) companion diagnostics	Roche VENTANA TROP2 RxDx computational pathology device.	<b>1<sup>st</sup> AI-driven CDx</b>

\*datopotamab-deruxtecan

# Ki-67 Proliferation Index

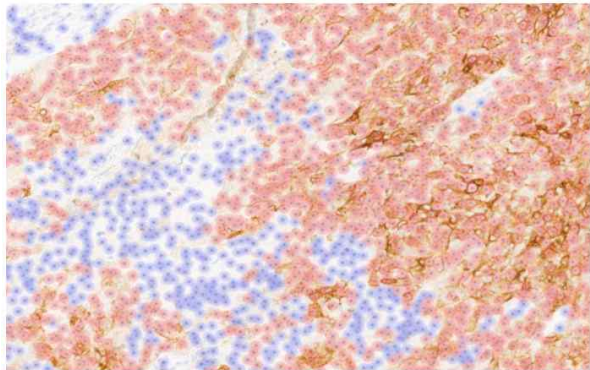
Gyeongsang National University Hospital





Tumor proportion score (TPS)

$$= \frac{\text{Any positive tumor cells}}{\text{Total tumor cells}}$$



Tumor with membrane staining (red overlay)

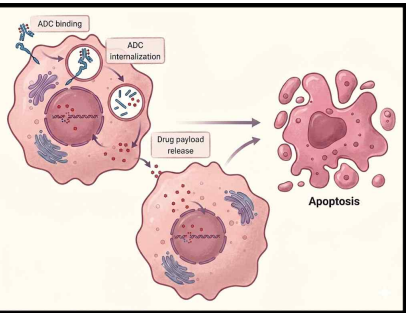
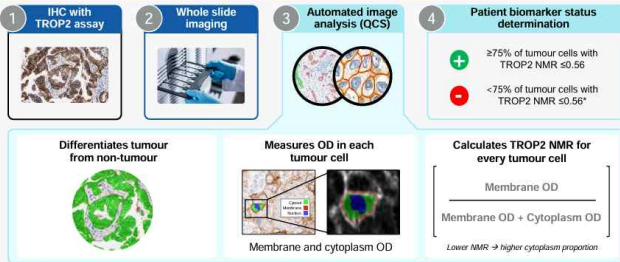
Tumor w/o membrane staining (blue overlay)

# TROP2 Normalized Membrane Ratio (NMR) Measured by Quantitative Continuous Scoring (QCS)



- QCS uses AI to calculate the NMR, accurately measuring predictive TROP2 expression on the cell membrane.
- This overcomes the limits of traditional visual scoring and helps precisely identify patients who will respond to targeted treatments.

Figure 1. TROP2 NMR measured by QCS



Antibody-Drug Conjugates(ADCs): Mechanism of Action (generated with Google Gemini)

# B. Driver Mutation Prediction: DeepPATHO



## Study design

- Evidential DL model (DeepPATHO) to predict **ALK expression** from H&E WSIs of lung adenocarcinoma, trained on SMC/GNUH datasets and validated on independent biopsy and TCGA cohorts.

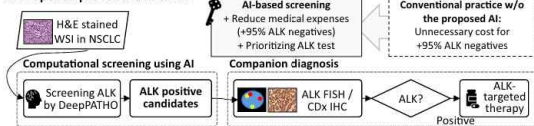
## Key results

- Patch/slide-level **AUC 0.922/0.962** on internal validation
- Very high recall** across internal and biopsy cohorts, missing only one rare mucinous-like ALK+ case in TCGA

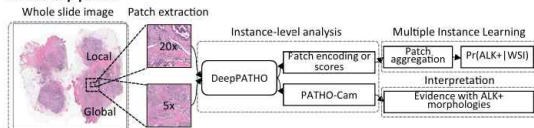
## Clinical message

- Clinically applicable H&E-based ALK screening tool that can prescreen ALK candidates and help **prioritize CDx testing** when tissue or resources are limited.

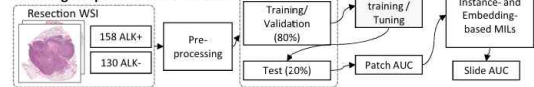
### a. Proposed procedure in clinic



### b. Overall pipeline



### c. Training DeepPATHO and evaluation



### d. Independent and external validation

Dataset	Sample Size	W. Acc	F1-score	Precision	Recall
SMC Biopsy	26 ALK+ / 66 ALK-	98.2%	98.3%	96.5%	100%
	W. Acc	98.2%			
	W. F1-score		98.3%		
	Precision			96.5%	
	Recall				100%
CGNUH Resection/biopsy	5 ALK+ / 5 ALK-	90%	90.9%	83.3%	100%
	Acc	90%			
	F1-score		90.9%		
	Precision			83.3%	
	Recall				100%
TCGA Resection	4 ALK+ / 324 ALK-	98%	97.8%	100%	99%
	W. Acc	98%			
	W. F1-score		97.8%		
	Precision			100%	
	Recall				99%

# B. Driver Mutation Prediction: Lunit SCOPE Genotype Predictor

## Study design

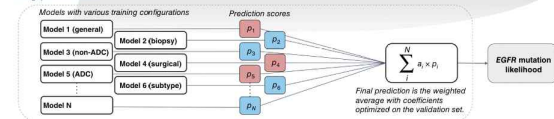
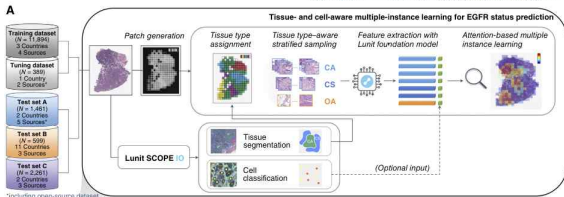
- Multicenter retrospective study of NSCLC H&E WSIs to develop and validate an **EGFR mutation prediction model**.
- Training/validation: TCGA + CPTAC + AstraZeneca biobank etc, total 12,000 WSIs.

## Key results (AUC)

- Primary test cohort (n≈1,461): **0.905**
- Independent 2nd test cohort (n≈599): **0.860**
- Multi-scanner concordance across 2,261 cases: agreement rate 90.4% between AI predictions across scanners.

## Clinical message

- H&E-based deep learning can reliably discriminate EGFR-mutant vs wild-type NSCLC from routine slides.
- Best used as a triage tool to **prioritize molecular testing** in settings with limited tissue or incomplete genotyping.



	Test set A			Test set B		
	(n = 1,461)			(n = 599)		
	AUROC	Specificity	Sensitivity	AUROC	Specificity	Sensitivity
	(95% CI)	At sensitivity 90% (95% CI)	At specificity 90% (95% CI)	(95% CIs)	At sensitivity 90% (95% CI)	At specificity 90% (95% CI)
Total	0.905	0.731	0.743	0.860	0.611	0.546
EGFR mutated	(0.884-0.922)	(0.650-0.828)	(0.671-0.815)	(0.828-0.888)	(0.510-0.699)	(0.442-0.654)
EGFR wild type						



## Study design

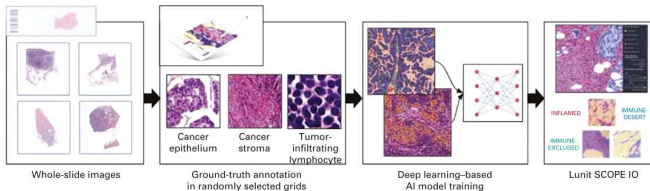
- Classification of **immune phenotypes (inflamed, immune-excluded, immune-desert)** and relate them to ICI outcomes in advanced NSCLC using an AI-based whole-slide TIL analyzer (Lunit SCOPE IO)

## Key Results (treatment prediction)

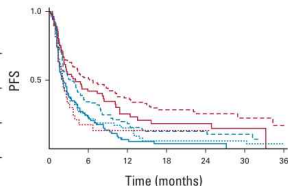
- Inflamed immune phenotype** was associated with **higher response rates and longer PFS/OS to ICI** than immune-excluded or immune-desert phenotypes in advanced NSCLC.

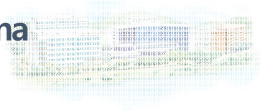
## Clinical message

- AI-based spatial TIL phenotyping can **complement PD-L1 TPS—particularly in the TPS 1–49% range**—by refining selection of NSCLC patients for immune checkpoint inhibitors, rather than replacing PD-L1 as a stand-alone biomarker.



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)		



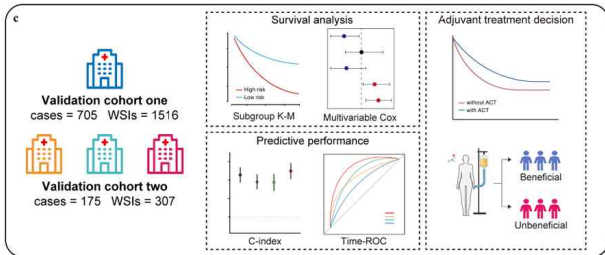


## Study design

- WSI-based deep learning **model predicting recurrence risk after resection in lung adenocarcinoma**, trained without manual annotations and validated in two independent multicenter cohorts.

## Key Results (treatment prediction)

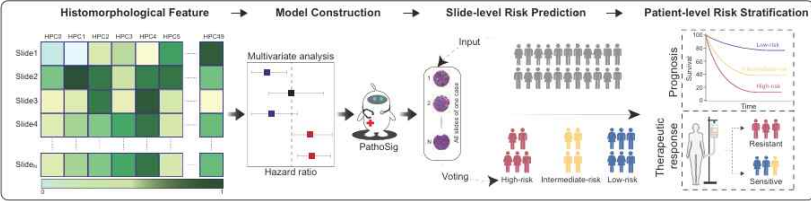
- DL-defined high vs low risk groups **significantly ( $p \leq 0.001$ )** separated DFS in both validation cohorts (**HR=1.95, 3.04**) and remained an **independent predictor** alongside IASLC grade and TNM stage (**HR=1.82, 2.96**).
- Combining the DL score with TNM stage identified **high-risk stage II-III patients** who gained clear or potential **DFS benefit from adjuvant chemotherapy**, while low-risk patients showed little benefit.



## Clinical message

- A WSI-based DL risk score can **complement TNM stage and IASLC grade** to refine prognosis and guide adjuvant chemotherapy use in resected lung adenocarcinoma.

# D. Prognosis Prediction: Small cell carcinoma



c.Pathomics Signature

## Study design

- Multicenter retrospective study of pure and combined SCLC (n≈380) treated mainly with postoperative chemoradiotherapy (predominantly **limited-stage disease**)
- A deep learning model extracted **histomorphologic clusters** from these cores to compute the **PathoSig** risk score.

## Key Results

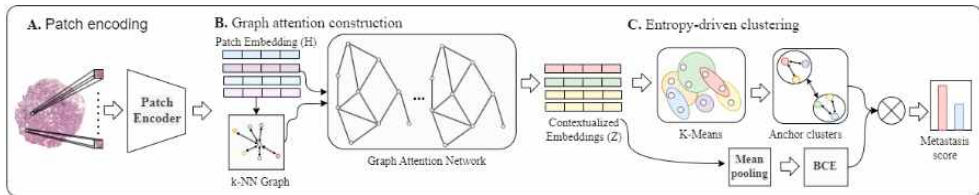
- **The PathoSig-defined risk groups** had significantly **shorter OS/DFS** than low-risk patients across all four cohorts.
- **Multivariable HRs** in the three validation cohorts  
 OS: **5.03, 9.96, and 2.48** (all P<0.02) / DFS: **3.76, 3.46, and 2.63** (all P≤0.041)

## Clinical message

- PathoSig is a **robust and independent prognostic factor** for SCLC-LD

## Study design

- Pilot study using H&E WSIs from the TCGA-LUAD cohort to **predict metastatic relapse in stage I lung adenocarcinoma**.
- Patients were labeled using a 2-year metastasis-free standard:
  - **High-risk:** 208 patients with lymph node or distant metastasis at surgery or within 2 years postoperatively.
  - **Low-risk:** 58 patients with no nodal or distant metastasis at surgery and at least 2 years of metastasis-free follow-up.
  - **Late-recurrence set:** 16 stage I patients with metastasis beyond 2 years (13 at 2–5 years, 3 after 5 years) were excluded from training and reserved as an **independent test set**.



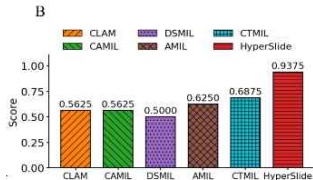
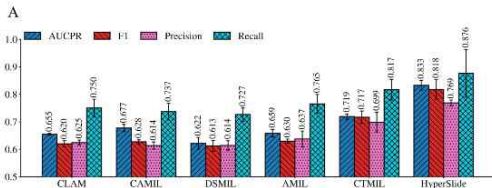
Overview of HyperSldie

## Key results (AUC)

- HyperSlide achieved **AUCPR 0.876**
- On the **late-recurrence test set (n = 16)**, **HyperSlide correctly identified 15 cases**, versus 11 for CTMIL.

## Clinical message

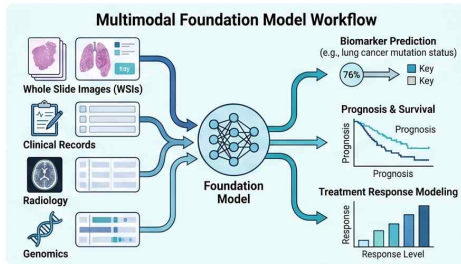
- These WSI-derived risk scores may serve as a **clinical decision-support tool for adjuvant chemotherapy in resected stage I LUAD.**



# Long-term future: Foundation Models and Pathology LLMs for Decision Support

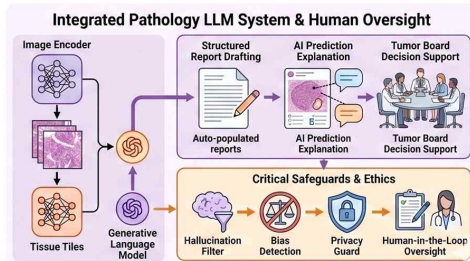
## Multimodal pathology foundation models

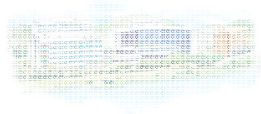
- Pretrained on very large collections of **WSIs together with basic clinical, radiological or genomic information.**
- Learn general image features that can later be **fine-tuned for specific cancer tasks** such as biomarker prediction, treatment response, or prognosis.



## Pathology large language models (LLMs)

- Combine image encoders for pathology slides with language models that can read and write text.
- Potential uses: drafting structured pathology reports, **explaining AI predictions in natural language**, and supporting tumor boards or resident education.





- For IHC markers such as **Ki-67, PD-L1 and TROP2**, **quantitative pathology** and AI are already in routine use or ready to use as decision-support tools that improve reproducibility.
- **H&E-based deep learning models** for **driver mutations, immunotherapy response and prognosis** are reaching clinically useful performance as **clinical decision-support systems**, but still require careful external and local validation before broad deployment.
- **Multimodal foundation models and pathology-specific LLMs** may eventually act as **co-pilots for report drafting and clinical decision-support**, but in patient care they must remain tightly constrained under strong human oversight.

# 경청해 주셔서 감사합니다



- 형태계측검사의 도입과 현재 사용
  - 2000년경 골·신경·골격근 조직에서 image analyzer로 세포주기·핵산·핵면적·세포 수 등을 계측하는 목적으로 허가.
  - 2018년 '나551 계측병리검사'에서 '나569 세포주기 및 핵산분석검사' 하위 항목으로 이동, 현재는 거의 Ki-67 형태계측 위주로 청구.
  
- 현행 인정 원칙(나569나)
  - '나569나 세포주기 및 핵산분석(형태계측검사)'는 정확한 수치(index) 계측이 임상적으로 직접 의미가 있을 때만 산정.
  - WHO 분류·가이드라인에서 증식률에 따라 약제 선택을 권고하는 신경내분비종양·유방암의 Ki-67이 대표 사례.
  
- 대한병리학회 의견: Ki-67 전용 구조 + 재검토 필요
  - 재분류 이후 다양한 형태계측검사가 있음에도 세포주기 관련 Ki-67만 청구 가능해 **사실상 Ki-67 전용 코드화**.
  - 2000년 허가 이후 디지털 병리 도입 등 환경 변화가 커, 현 시점에서 구조적 재검토가 필요하다는 의견.
  - 제안:
    - Ki-67 형태계측은 현행 '나569나' 유지.
    - 이외 임상적 의미가 있는 marker의 정량 계측은, '나567 IHC Level I(C5673)' 시행 후 index 계측한 경우 '나567 IHC Level II(C5674)'로 산정하는 체계 필요.



## ● AI 병리 수가 평가 원칙

- HIRA·복지부가 병리 분야 AI 기반 의료기술 급여 평가 가이드라인을 마련해, 진단 정확도 향상·비용 효과성이 입증되면 추가 가치(가산료, 별도 행위)를 인정하겠다고 밝힘.

## ● 제한적 실제 급여 사례

- 현재 건강보험에서 인정된 AI 분석료는 주로 영상 분야(뇌 MRI 등)와 일부 비급여 코드(안저·심전도)에 머물러 있고, **병리 AI는 아직 정식 수가**가 거의 없는 상태.

## ● 디지털 병리 인프라와 AI 소프트웨어 비용이 수가에 반영되지 않음

- 슬라이드 스캔·서버·소프트웨어 유지비 등 초기·운영 비용에 대한 명확한 보상 구조가 없어, 병원 입장에서 디지털 병리와 AI 도입이 재정적으로 손해가 되기 쉽다는 점이 학술·언론에서 반복 지적됨.

## ● '새로운 의료행위'로 인정 범위가 좁아 적용이 지연

- 세포 수 계수 같은 일부 계측 기능은 기존 계측 병리 수가 안에서 처리되는 반면, **예후·치료 반응 예측 등 새로운 정보를 주는 AI는 신의료기술·혁신의료기술 평가를 거쳐야** 하고, 장기적인 치료 효과·비용 효과 자료가 부족해 급여 결정이 늦어지는 상황