



AI-Driven Survival Statistics for Progression-Free and Overall Survival

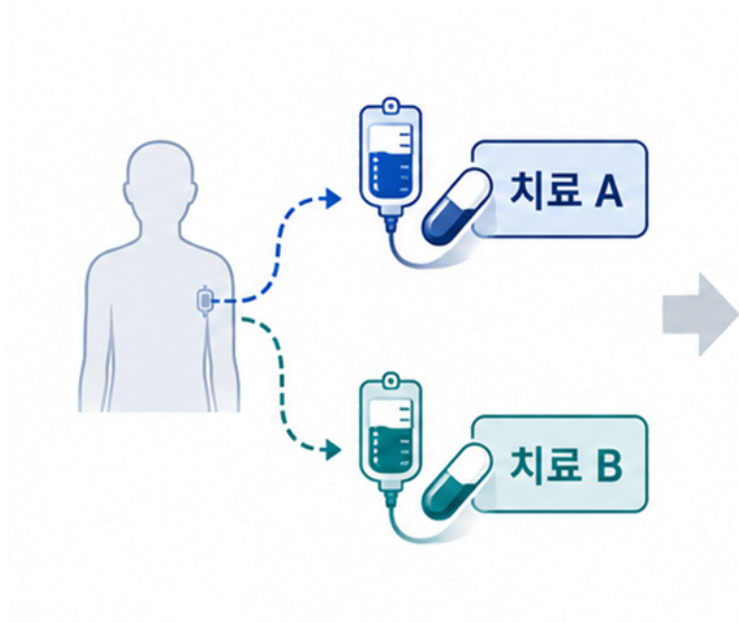
가톨릭대학교
심진아

DECLARATION OF INTERESTS

Nothing to declare

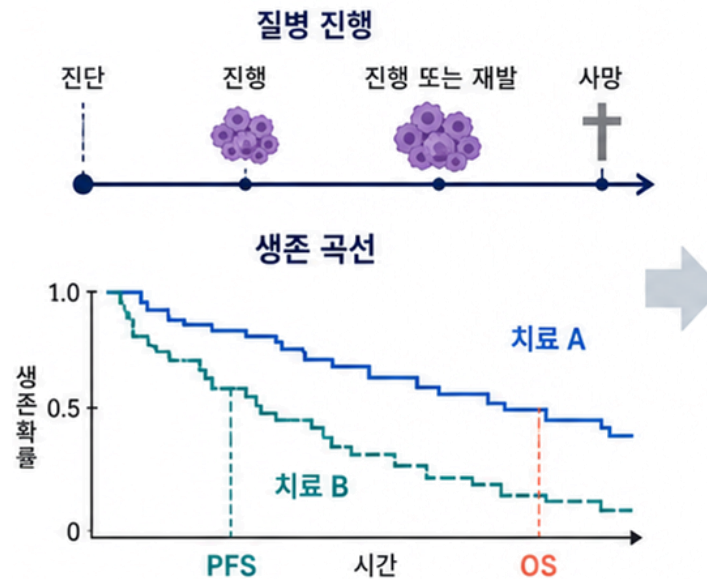
임상에서의 3가지 질문

1 두 항암치료 중 어느 치료가 환자의 생존을 더 연장하는가?



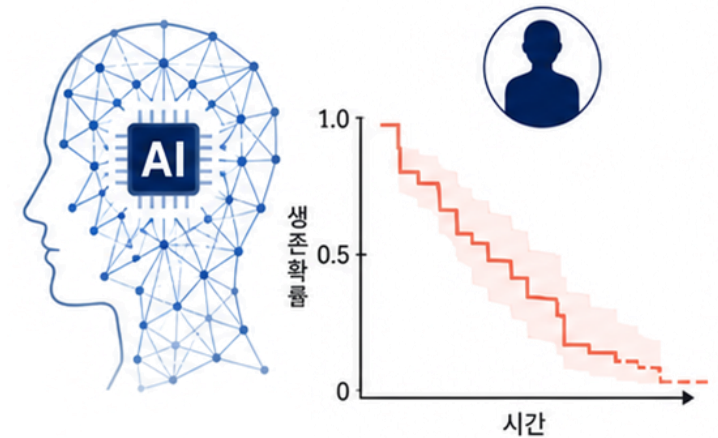
Overall Survival

2 질병 진행은 늦었지만, 전체 생존은 연장하지 못했다면 효과가 있을까?



Progression Free Survival

3 AI는 개별 환자의 생존확률을 예측 할 수 있는가?



AI기반 개인별 예측

왜 생존분석이 필요한가? 왜 일반적인 통계분석으로는 안 되는가?

사망여부만 분석할 경우



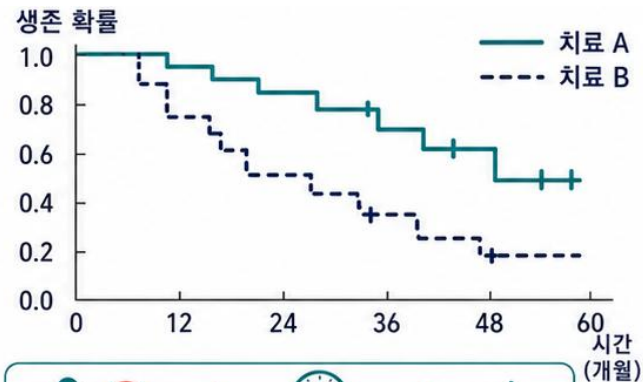
⚠️ 사건이 언제 발생했는지 반영하지 못함

평균 생존 기간 비교



⚠️ 중도절단을 처리하지 못함

그렇다면, 생존분석은?



ⓧ + 🕒 + |
사건 여부 발생 시간 중도절단

✅ 사건 발생 여부와 시간 정보를 동시에 이용함

사건 여부 + 발생 시간 + 중도절단

실제 연구 사례: Disease-free 암 생존자에서 HRQOL이 장기 생존을 예측



자궁경부암
생존자

- 856명 · 사망 30명 · 추적 중앙값 6.3년
- 낮은 신체기능: **aHR 4.25**
- 낮은 역할기능: **aHR 4.82**

Kim, Sim et al. Int J Gynecol Cancer. 2016



수술 후 폐암
생존자

- 809명 · 사망 96명 · 4,509 person-years
- 낮은 신체기능: **aHR 2.39**
- 호흡곤란: **aHR 1.56**

Yun, Kim, Sim et al. BMC Cancer. 2016



조사 시점에는 모두 disease-free였지만, 이후 사건 발생 시점과 중도절단을 반영하여 예후요인을 확인



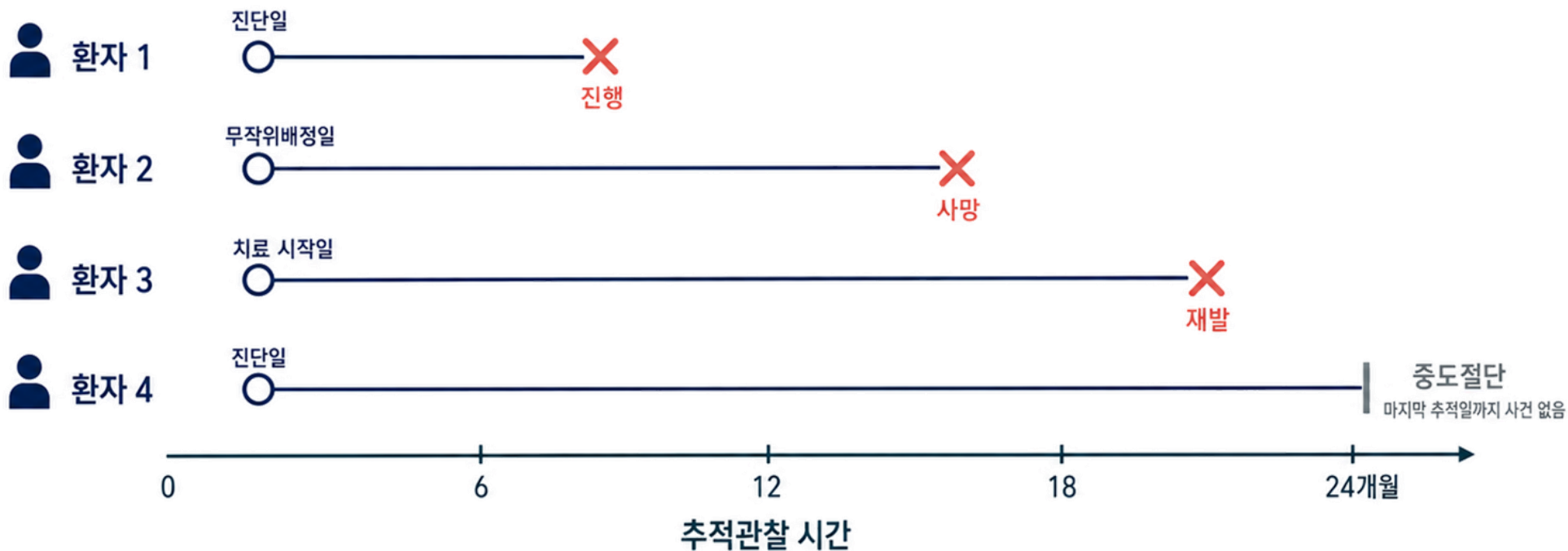
왜 생존분석이 필요한가?

왜 일반적인 통계분석으로는 안 되는가?

○ 시작 시점: 진단일, 무작위배정일, 치료 시작일

✕ 사건: 사망, 진행, 재발

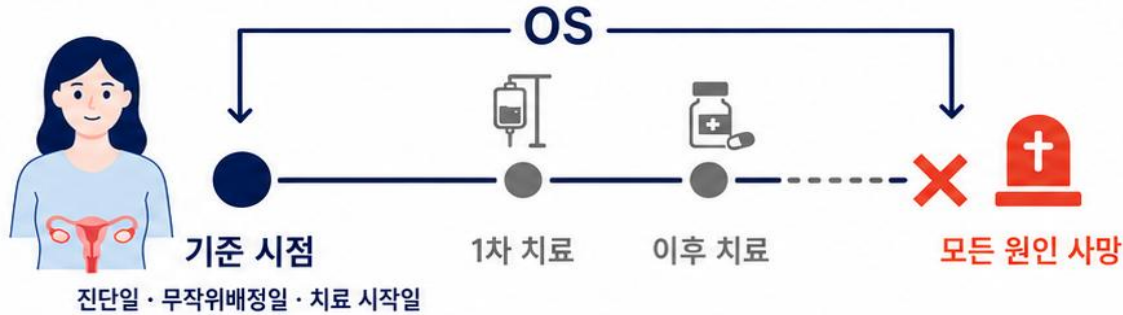
| 중도절단: 마지막 추적일까지 사건이 관찰되지 않음



관찰된 시간과 사건
발생 여부를 함께 분석

Overall Survival, OS

기준 시점부터 모든 원인에 의한 사망까지의 시간



장점

- 가장 객관적이고 임상적으로 중요한 평가변수
- 사망 원인 판정이 필요하지 않음

고려할 점

- 장기간 추적과 많은 표본이 필요
- 이후 치료나 교차치료의 영향을 받을 수 있음



명확한 평가변수이지만
결과 확인까지 시간이 오래 걸릴 수 있음

실제 연구 사례

같은 809명, 다른 분석



수술 후 disease-free 폐암 생존자 809명
사망 96명

2016 Cox 생존분석

시간 + 사건 + 중도절단

낮은 신체기능: aHR 2.39
호흡곤란: aHR 1.56

질문: 어떤 요인이
사망위험과 연관되는가?

2020 머신러닝 예측

임상정보 + HRQOL

DT · LR · Bagging · RF · AdaBoost
5년 생존 · 사망 이진분류
내부검증 AUC 최대 0.981

질문: 개인의 5년 생존 여부를
예측할 수 있는가?



Cox: 시간과 중도절단 반영

VS



AI 분류: 5년 생존 여부 예측



주의

- 중도절단 27명 제외
- 내부 5-fold 검증만 시행 · 외부검증 없음
- 업샘플링 후 자료 분할: 데이터 누출 가능성 점검 필요

Yun et al. BMC Cancer. 2016;16:354.

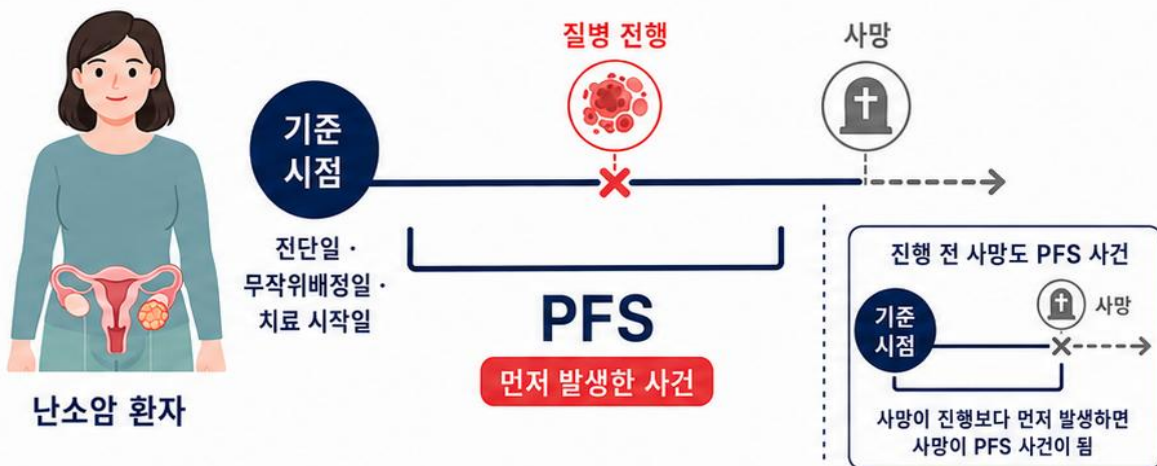
Sim et al. Scientific Reports. 2020;10:10693.



AI가 생존 여부를 잘 분류해도
시간과 중도절단을 처리하지 않으면
엄밀한 생존분석은 아니다

Progression-Free Survival, PFS

기준 시점부터 질병 진행 또는 사망 중 먼저 발생한 시점까지의 시간



장점

- ✓ OS보다 사건이 일찍 발생하여 결과를 빠르게 평가 가능
- ✓ 후속 치료의 영향을 상대적으로 덜 받음



고려할 점

- ! 영상평가 주기와 진행 판정 기준에 영향을 받음
- ! PFS 개선이 반드시 OS 개선을 의미하지는 않음



질병 억제 효과를 빠르게 평가하지만
임상적 생존 이득과는 구분해야 함



실제 연구 사례

유전정보로 PFS 고위험군을 구분할 수 있는가?



대장암 수술 환자
960명



수술일



면역 관련
SNP 37,398개



Cox
생존분석



PFS
위험층화

PFS: 수술일부터 진행 · 재발 · 사망 중 먼저 발생한 시점까지

NOS3 변이:
PFS HR 4.67

병기 1-3
탐색 코호트

PRS 최고군:
PFS HR 9.81

높은 유전위험 점수와
불량한 PFS 연관



- 주의: 외부 코호트에 재발 자료가 없어 PFS 검증은 제한적
- AI 모델이라기보다 고차원 유전정보 기반 Cox · PRS 분석

Yun et al. BMC Cancer. 2025;25:456.



PFS는 바이오마커 기반
예후예측에도 활용된다

PFS와 OS 비교

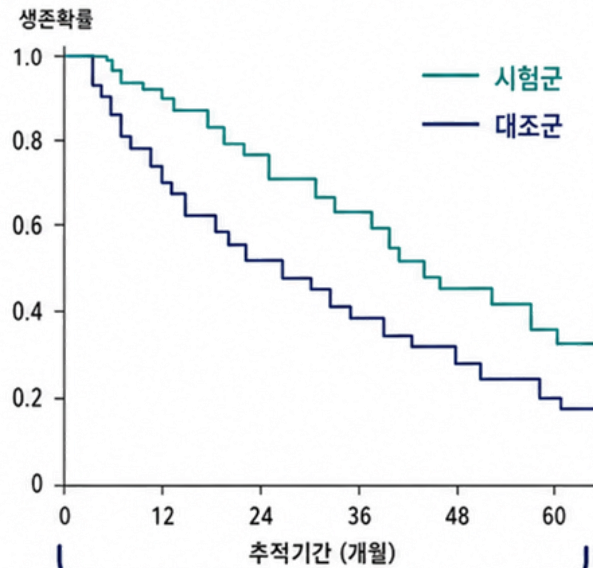


진행 또는 사망	사건	모든 원인 사망
✓ 조기 평가 가능, 사건 수가 많음	✓ 장점	✓ 객관적이고 임상적 의미가 명확
⚠ 평가 시점·판독 기준의 영향	⚠ 한계	⚠ 장기 추적 및 큰 표본 필요
➡💊 상대적으로 적음	➡💊 후속 치료 영향	➡💊 큼
? 질병을 얼마나 오래 억제하는가?	? 임상적 질문	? 생명을 얼마나 연장하는가?

PFS는 질병 억제를, OS는 생존 연장을 평가

Log-rank test 와 Hazard Ratio, HR

Log-rank test

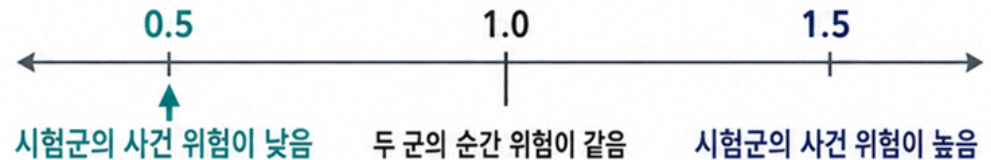
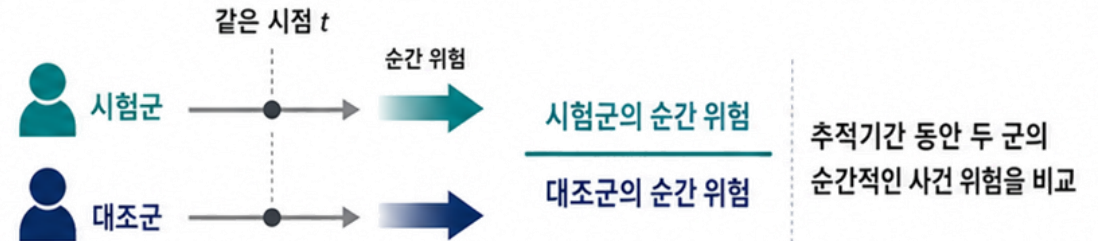


두 생존곡선 전체가
같은지 검정

귀무가설: 두 생존곡선이 같다



Hazard Ratio, HR



$$HR = 0.70$$

시험군의 순간적 사건 위험이
대조군보다 약 30% 낮음

⚠ 주의: 이렇게 해석하면 안 됨



생존기간이 30% 증가했다



사망확률이 30% 감소했다



HR은 생존기간이나
누적 사망확률의 비가 아니라
순간 위험의 비

HR만으로 충분하지 않은 이유

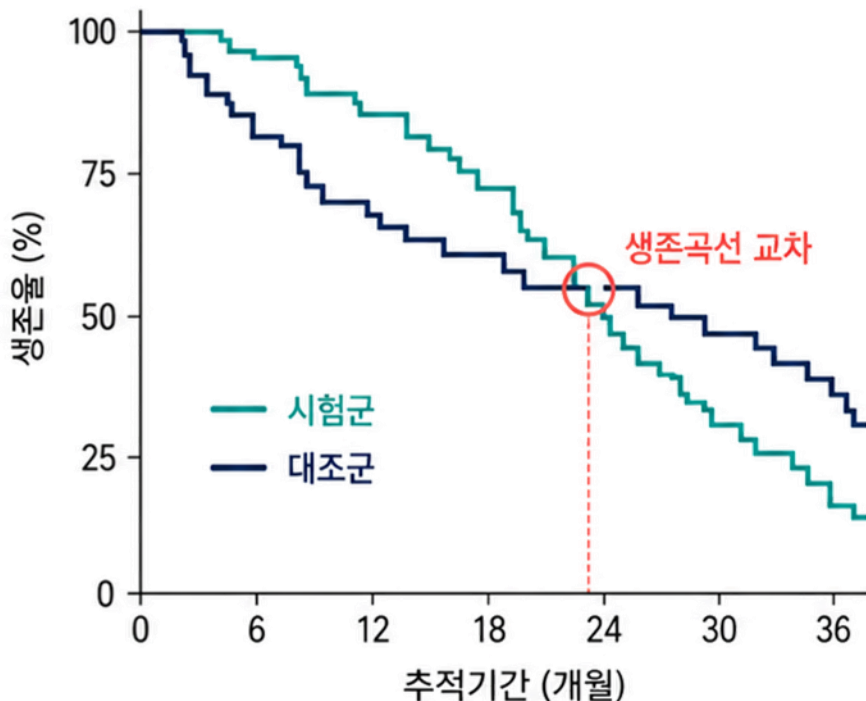
HR = 하나의 평균적 요약



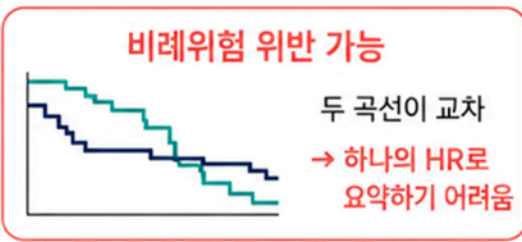
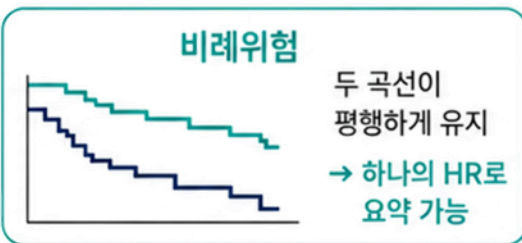
시간에 따라 치료효과가 달라지면
하나의 HR로 요약하기 어려움

함께 제시할 결과

p-value와 HR만 제시하지 말고



HR은 비례위험 가정을 전제로 함

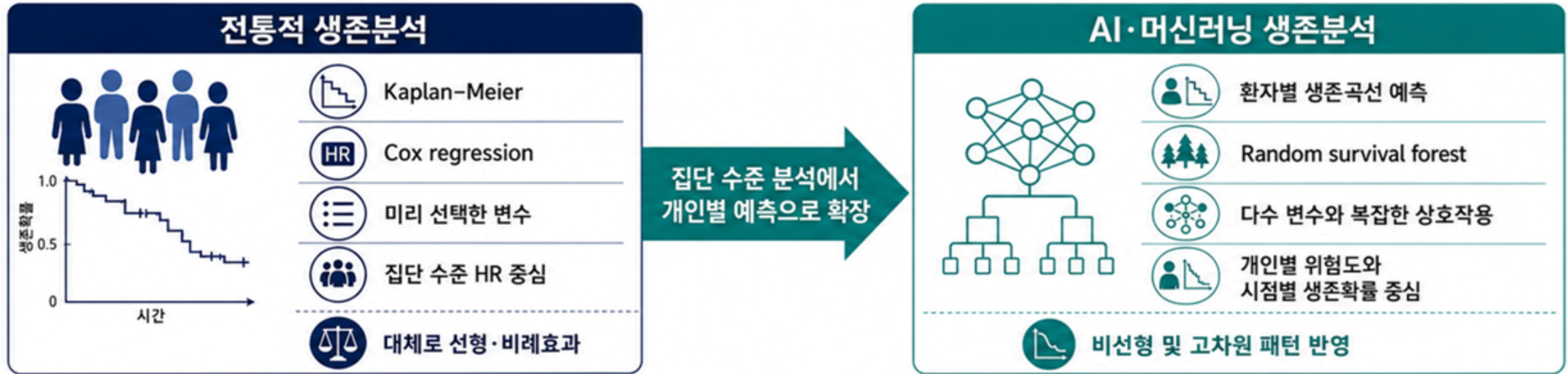


※ 교차가 있다고 해서 비례위험 가정이
항상 위반되는 것은 아님(위반 가능)

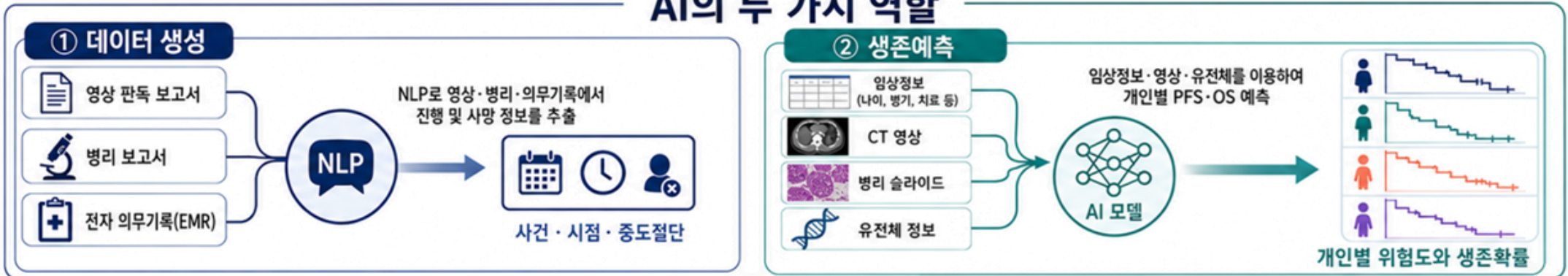
- 1 Kaplan-Meier 곡선
- 2 중앙 PFS 또는 OS
- 3 특정 시점 생존율
- 4 95% 신뢰구간
- 5 필요할 경우 RMST

치료 효과의 크기와 시간적 변화를 함께 해석

AI-driven survival analysis?



AI의 두 가지 역할

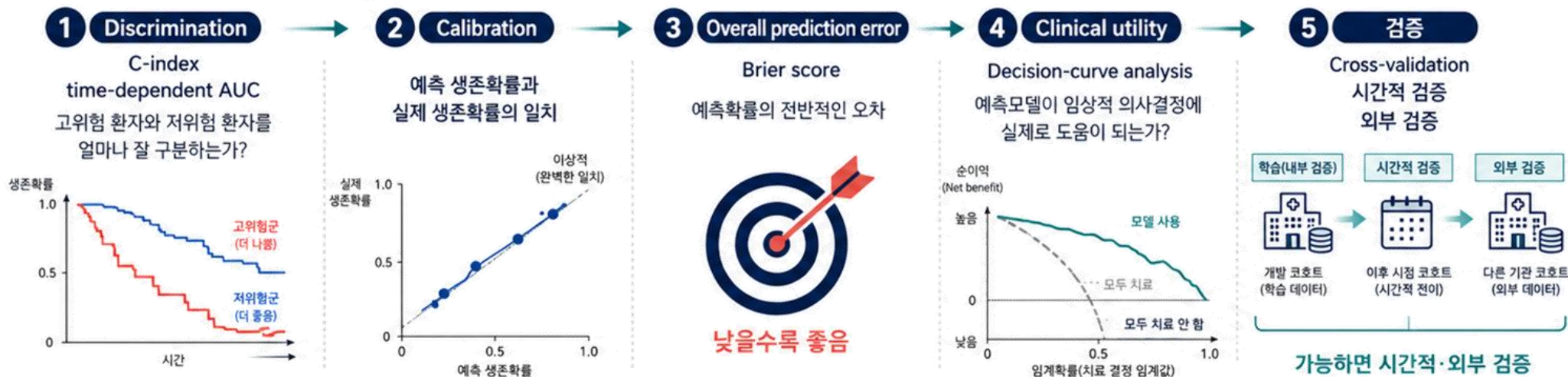


AI는 생존분석을 대체하는 것이 아니라, 더 많은 정보를 이용하여 개인별 생존확률을 예측하도록 확장한다.

AI 기반 생존모델의 평가 방법



일반적인 accuracy만으로 평가할 수 없음



함께 확인해야 할 주의점



중도절단을 고려해야 함

생존분석 지표는 중도절단을 반영하여 계산되어야 함



데이터 누출을 방지해야 함

학습·검증 데이터의 정보 누출은 과도한 성능을 야기함



치료 선택에 따른 교란이 존재할 수 있음

비무작위적 치료 배정으로 예측에 교란이 생길 수 있음



높은 예측성능이 치료효과나 인과효과를 의미하지 않음

예측은 인과를 증명하지 않으며, 모델 성능이 치료 효과를 보장하지 않음

좋은 AI 기반의 생존분석 모델은 accuracy뿐 아니라, calibration (보정), 임상적인 유용성 및 외부 타당성을 함께 갖추어야 함

AI적용 생존분석 논문 사례 (1)

하나의 HR이 전체 기간동안의 치료효과를 충분히 설명할 수 있는가?

JN JAMA Network

ICON7-A cohort 745명

Bevacizumab 384명

Control 361명

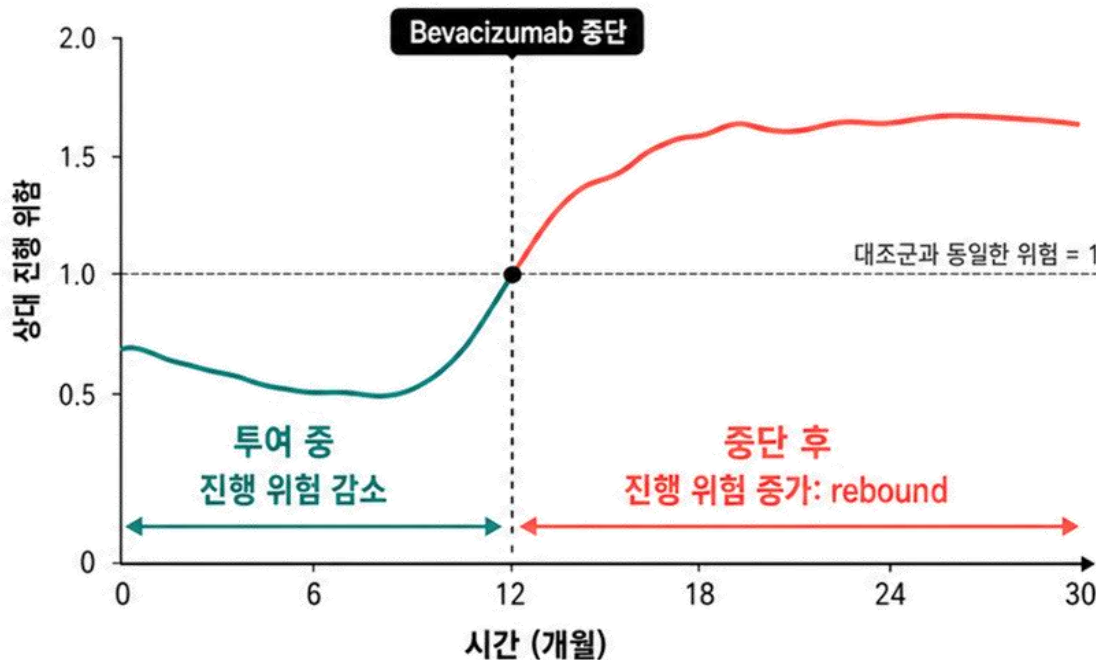
ICON7-GOG-0218 등 7개 3상 시험 분석

ICON7 전체 PFS HR = 0.81
(95% CI, 0.70-0.94)

? 이 하나의 숫자로
시간에 따른 효과 변화를
설명할 수 있는가?

Takamatsu et al. JAMA Network Open. 2023

시간에 따른 상대 진행 위험 (Bevacizumab / Control)



RMST로 나누어 보면

전체 기간

RMST ratio 1.07
95% CI 0.97-1.17
 $P = .16$

중단 전

RMST ratio 1.08
95% CI 1.05-1.11
 $P < .001$

중단 후

RMST ratio 0.79
95% CI 0.69-0.90
 $P < .001$

중단 전 PFS 이득,
중단 후 효과 역전



비례위험 가정이 성립하지 않으면 하나의 HR보다 시간별 위험과 RMST가 치료효과를 더 잘 보여줄 수 있음

AI적용 생존분석 논문 사례 (2)

AI를 통한 서로 다른 PFS를 보이는 고위험군 선별

Darbandsari et al. Nature Communications. 2024

nature
communications



Discovery 368명



독립 검증 290명

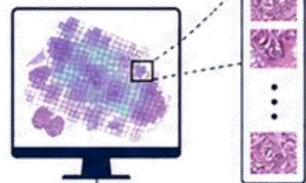
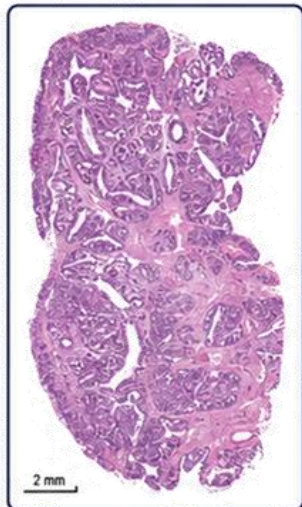


다기관 검증 614명
26개 병원

일상적인 H&E
병리 슬라이드

AI 기반
병리영상 분석

기존 분자분류: NSMP



Multiple Instance Learning
(MIL)

NSMP 환자군
(분자분류 기반 저위험-중간위험 그룹)



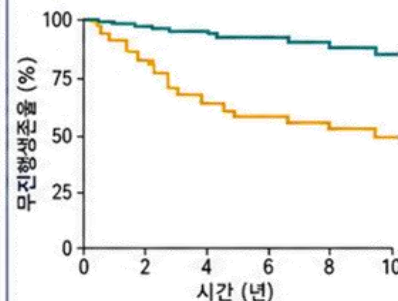
일반 NSMP
약 80%



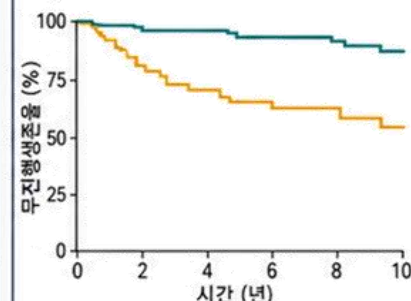
p53abn-like NSMP
숨은 고위험군
약 20%

➔ 병리 형태가 p53-abnormal과 유사

Discovery cohort: 10년 PFS



Validation cohort: 10년 PFS



병기·등급·조직형 보정 후에도

진행 위험 HR = 2.5

$p = 0.01$

HR=2.5는 병기, 등급, 조직형을 보정한 후의 무진행생존(PFS) 관련 위험비(조정 결과)



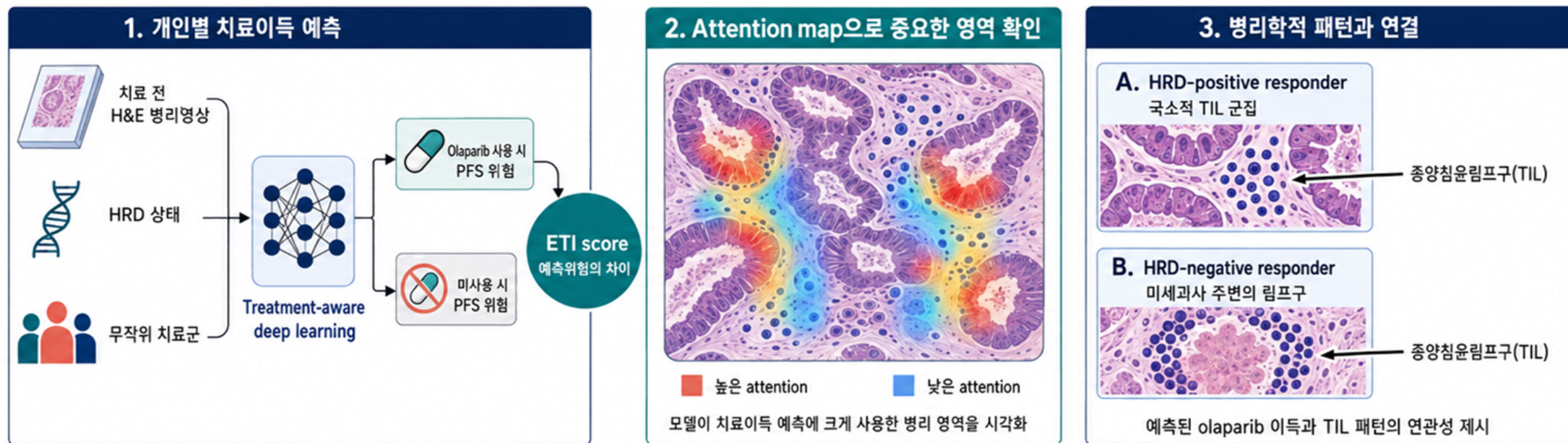
DSS에서도 일관된 고위험 양상



같은 NSMP 환자라도 AI가 병리영상에서 서로 다른 PFS를 보이는 **고위험군을 추가로 식별**

설명 가능한 AI(XAI, eXplainable AI) (1)

Attention map 기반 설명가능성



설명과 예측의 연결

ETI treatment-biomarker interaction HR 0.36
95% CI 0.22-0.59

Attention 영역에서
치료민감성과 관련될 수 있는
형태학적 가설 생성

설명가능성의 한계

- Attention map은 모델이 본 위치를 보여줄 뿐이다
- TIL이 치료효과의 원인이라는 의미는 아니다
- 독립 외부검증 없음
- 병리전문가 검토와 생물학적 검증 필요

‘설명가능한 AI는 ‘왜 그렇게 예측했는가’에 대한 단서를 제공하지만, 인과적 기전을 증명하지는 않는다

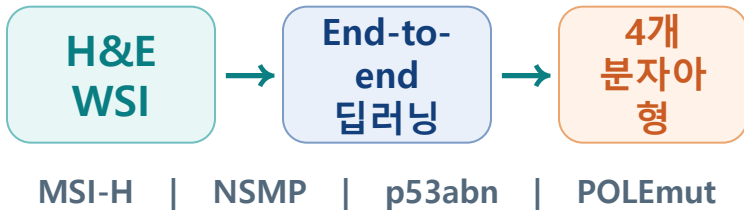
설명 가능한 AI(XAI, eXplainable AI) (2)

Grad-CAM 기반 예측 근거와 생존결과의 연결

Guo et al. *npj* | precision oncology · 2026

연구 설계

Fudan 324명 · 364 WSI
외부검증: TCGA 274명 · Suzhou 33명



분류 성능 · macro-AUROC

0.867

Fudan
5-fold CV

0.844

TCGA
외부검증

0.847

Suzhou
외부검증

p53abn 외부검증 AUROC 0.950 (TCGA)

Grad-CAM이 '어디를 보고 분류했는지' 시각화

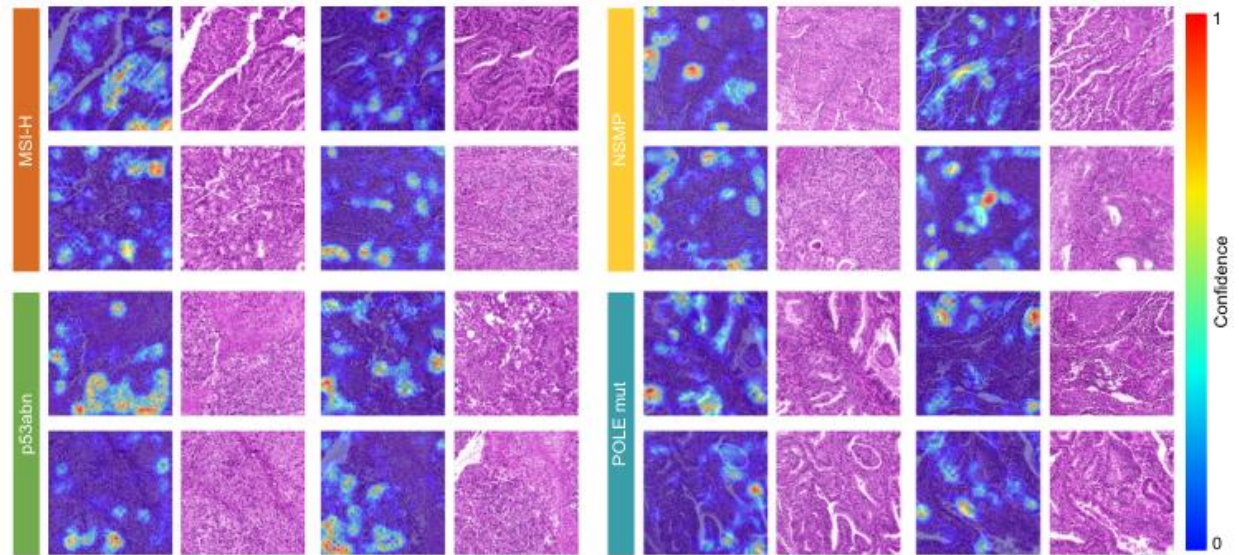


Fig. 2 | Patch-level visualization of the morpho-molecular correlates. The corresponding heatmaps were generated using the Grad-CAM method based on the subtypes. The right patch shows the original morphological features. MSI-H: microsatellite instability-high. NSMP: no specific molecular profile. p53abn:

빨강·노랑 영역 = 모델 예측에 강하게 기여한 조직 부위

형태학적 단서 | MSI-H: 기질 림프구 ↑ · p53abn: 유두상·장액성 형태
POLEmut: 이질성·고형 성장 · NSMP: 기질 세포성 ↑

핵심 | Grad-CAM 기반 XAI이며, 분자아형 분류 연구이다 · OS·PFS의 시간과 중도절단을 직접 모델링한 생존분석은 아니다.

설명 가능한 AI(XAI, eXplainable AI) (3)

SHAP value 기반 설명가능성

XGBoost-Cox modeling with SHAP analysis · Feng et al., BMC Cancer, 2026

설명가능한 AI 생존분석

단일기관 난소암 환자 313명
2012-2025 후향적 코호트



✓ 사건 시간 + 중도절단을 함께 반영
예측 성능

0.738

OS C-index
5년 AUC 0.814

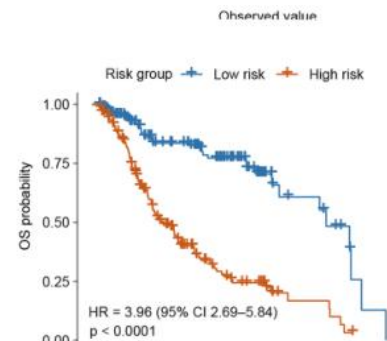
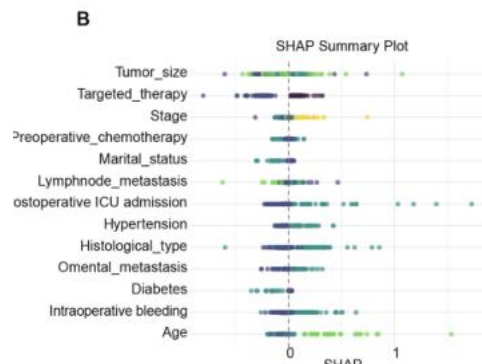
0.743

PFS C-index
5년 AUC 0.853

SHAP의 역할

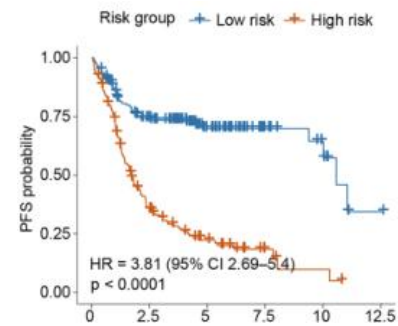
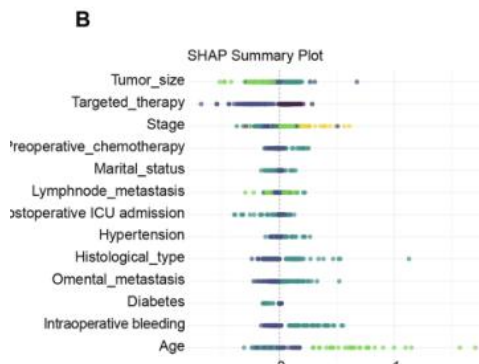
- 전체적으로 중요한 변수
- 변수값이 위험을 높이거나 낮추는 방향
- 환자별 예측에 대한 기여도

OS 주요 변수: 종양 크기 · 표적치료 · 병기 · 연령



KM: HR 3.96 (95% CI 2.69-5.84)

PFS 주요 변수: 연령 · 종양 크기 · 병기 · 표적치료 · 출혈



KM: HR 3.81 (95% CI 2.69-5.40)

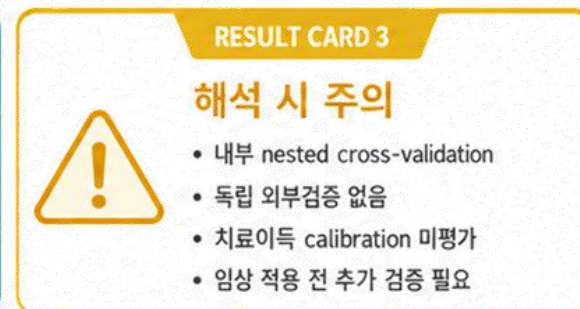
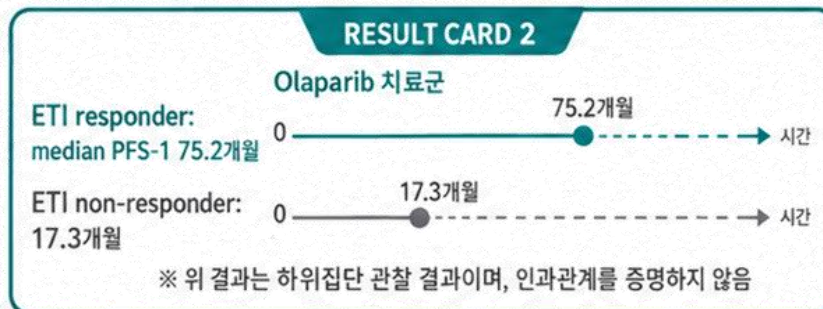
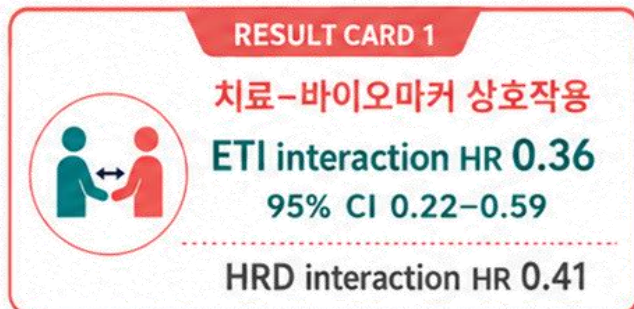
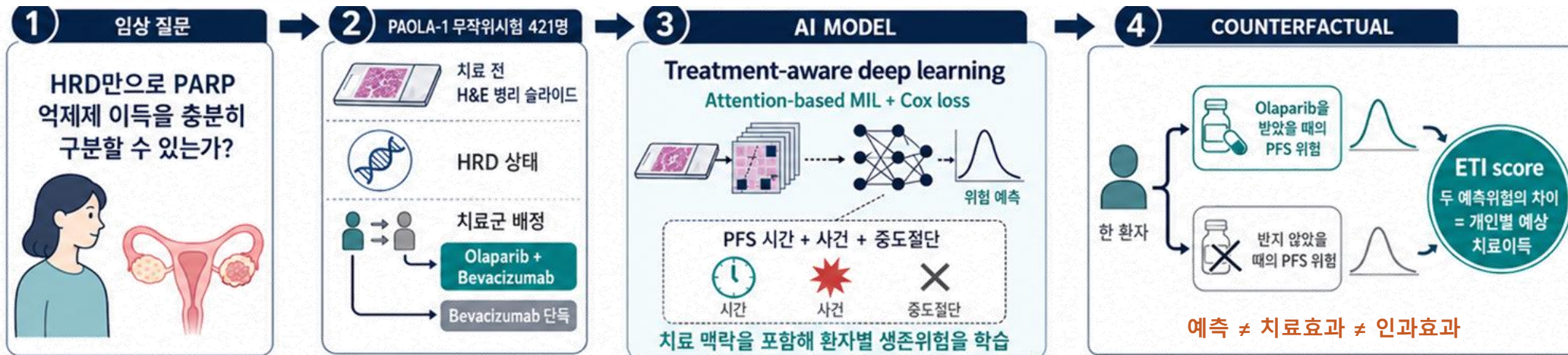
해석의 한계 | SHAP value는 모델의 예측 근거를 설명하지만, 치료효과의 인과성을 증명하지 않는다.
· 단일기관·내부검증 연구일 경우, 외부검증이 필요하다.

AI적용 생존분석 논문 사례 (3)

AI생존분석모델의 활용: 누가 PARP 억제제로 이득을 얻는가?

npj | Digital Medicine

난소암에서 병리영상 기반 개인별 치료의 효과 및 이익 예측



★ 예후가 좋은 환자를 찾는 것을 넘어, 각 환자에게 어떤 치료가 더 유리한지를 예측

환자별 PFS-OS 예측과 설명가능한 위험요인 요약

 연령 63세

FIGO 병기 III

조직형 장액성

 종양 크기 8.2 cm 

표적치료 ☒ 시행

림프절 전이 ☒ 있음

추적 정보

기준 시점: 치료 시작일

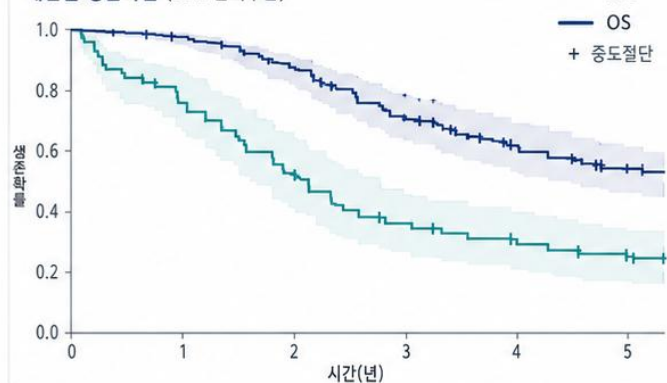
중도절단 고려

PFS

	1년	3년	5년
PFS	78%	52%	34%

OS	1년	3년	5년
	92%	71%	56%

개인별 생존곡선 (95% 신뢰구간)



C-index	3년 AUC	5년 AUC	Brier score
0.743	0.830	0.853	0.142

1년 3년 5년

3년 PFS 위험 기여도



① SHAP은 이 환자의 예측에 각 변수가 기여한 방향과 크기를 보여줍니다.

검증 정보

내부검증 완료 · 외부검증 필요



해석 주의

높은 예측위험은 치료효과 또는 인과효과를 의미하지 않습니다.

중도절단을 반영한 XGBoost-Cox 기반 예측 예시

AI 기반 생존분석, 꼭 기억해야 할 3가지!

TAKE-HOME MESSAGES

1

시간을 올바르게
측정할 것!

사건 여부만이 아니라
발생 시간 + 중도절단을
함께 분석

PFS와 OS는 서로 다른 임상
질문에 답한다

2

개인별 예측으로 확장!

AI는 비선형 패턴을 학습해서
환자별 생존곡선과
시점별 확률을 제시

SHAP은 예측에 기여한 방향과
크기를 설명한다

3

성능보다 신뢰성을
확인할 것!

C-index·time-AUC·
Brier score와
calibration을 함께 평가

외부검증 없는 높은 성능은
임상 적용의 근거가 아니다



좋은 AI 생존모델은 단순히 AUC가 높은 모델이 아니라,
시간과 중도절단을 올바르게 다루고 검증된 개인별 예측을 제공하는 모델입니다.

예측 ≠ 치료효과 ≠ 인과효과

Korean Society of
Gynecologic Oncology



2026년 대한부인종양학회 하계학술대회 제20차 젊은의사 워크숍 제9차 분과전문의 연수강좌 제1차 부인종양간호사 연수강좌 & 카데바 워크숍

일자 2026년 8월 8일(토) - 9일(일)

장소 위례 밀리토피아 마린 호텔 / 가톨릭대학교 응용해부연구소



대한부인종양학회
Korean Society of Gynecologic Oncology