Performance validation of an artificial intelligence-powered Programmed death-ligand 1 (PD-L1) combined positive score analyzer in urothelial cancer

Jeong Hwan Park¹, Soo Ick Cho², Kyu Sang Lee³, Euno Choi⁴, Wonkyung Jung², Sukjun Kim², Gahee Park², Sanghoon Song², Cholmin Kang², Minuk Ma², Donggeun Yoo², Kyunghyun Paeng², Chan-Young Ock²

EUMC 이대목동병원 SNUH 🗗 분당서울대학교병원

> **Contact (Jeong Hwan Park):** hopemd02@daum.net

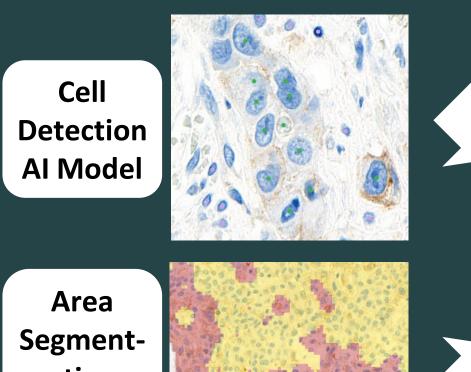
Background

- Programmed death-ligand 1 (PD-L1) expression is a predictive marker for immune checkpoint inhibitors (ICI) treatment in urothelial carcinoma (UC).
- The combined positive score (CPS) is a method to evaluate the expression level of PD-L1 in UC.
- Recently, artificial intelligence (AI) algorithms have been applied to pathology reading or interpretation and have reported pathologist-level performance.
- This study aimed to evaluate the performance of an artificial intelligence (AI)powered PD-L1 CPS analyzer on UC compared to the pathologists.

Methods

- Lunit SCOPE PD-L1 CPS was developed with 3.02 x 10⁵ tumor cells and 3.49 x 10⁵ immune cells from PD-L1 immunohistochemistry-stained whole-slide images (WSI) of UC from multiple institutions, annotated by 94 pathologists.
- The tissue area segmentation and cell detection AI models were developed based on a semantic segmentation algorithm, which includes an atrous spatial pyramid pooling block.
- To validate the model, a total of 543 PD-L1 stained UC WSIs were obtained from three university hospitals (Seoul National University Bundang Hospital [n = 245], Ewha Womans University Mokdong Hospital [n = 205], and Boramae Medical center [n = 93] in each).
- Three uropathologists evaluated slide-level CPS and assigned CPS high or low (10% cutoff value). The agreement (high or low) or correlation (continuous value) of CPS between the pathologists and the AI prediction was evaluated.

Schematic workflow of AI model development and uropathologist reader study



AI model

CPS calculation = 100 X PD-L1 staining cells (tumor cells, lymphocytes, macrophages) Total viable tumor Al model result on external validation set, compared to uropathologists SNUBH: n = 245

Urothelial carcinoma slide Ewha: n = 205Boramae: n = 93 Total: n = 543



- All pathologists agreed on the CPS level in 446 out of 543 cases (82.1%). The agreement or correlation between either of the two pathologists was 87.1%-89.9% (Table 1).
- Al model accuracy compared to pathologists' consensus was 88.8%, and the Intraclass correlation coefficient (ICC) value between the AI model CPS value and the average CPS value of pathologists was 0.94 (95% confidence interval [CI] 0.93–0.95) (Table 1, Figure 1).
- The performance of the AI model was similar with each individual pathologist 0.92 [0.90–0.94], 87.8% / 0.95 [0.93–0.97], and 89.2% / 0.92 [0.87–0.95], respectively) (Figure 3).

(accuracy / ICC; 85.1% / 0.94 [0.93–0.95] , 86.6% / 0.90 [0.87–0.92], and 87.1% / 0.93 [0.92-0.94], respectively) (Figure 2) and individual hospital dataset (89.4% /

Conclusion

Conflict of Interest (Presenting author): none to declare

¹Department of Pathology, Boramae Medical Center, Seoul, Republic of Korea

⁴Pathology, Ewha Womans University Mokdong Hospital, Seoul, Republic of Korea

³Pathology, Seoul National University Bundang Hospital, Seongnam-si, Republic of Korea

- This study demonstrates that an Al-powered PD-L1 CPS analyzer can classify CPS levels in UC comparable to pathologists.

Reference

Average of CPS (Hostpial A)

Choi S et al., Artificial intelligence-powered programmed death ligand 1 analyser reduces interobserver variation in tumour proportion score for non-small cell lung cancer with better prediction of immunotherapy response. Eur J Cancer. 2022;170:17-26.

Bellmunt J et al., A review on the evolution of PD-1/PD-L1 immunotherapy for bladder cancer: The future is now. Cancer Treat Rev. 2017;54:58-67.

Ghate K et al., PD-L1 expression and clinical outcomes in patients with advanced urothelial carcinoma treated with checkpoint inhibitors: A metaanalysis. 2019;76:51-56.



Average of CPS (Hostpial C)

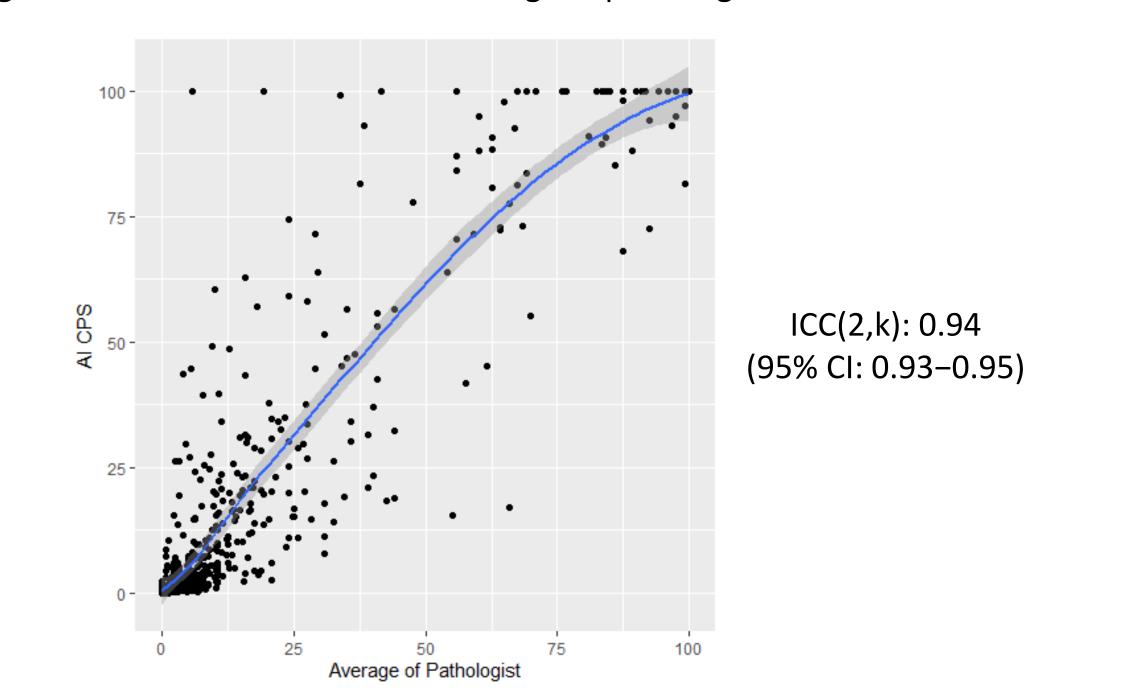




Table 1. Comparison of CPS evaluation between each pathologist or between pathologists and the AI model

Agreement (more than CPS 10 or not)	Correlation (Intraclass coefficient [2,k], 95% CI)
88.8%	0.94 (0.93-0.95)
87.1%	0.91 (0.90-0.93)
89.9%	0.96 (0.96–0.97)
87.3%	0.92 (0.90-0.93)
	(more than CPS 10 or not) 88.8% 87.1% 89.9%

Figure 1. CPS value between the average of pathologists and the AI model



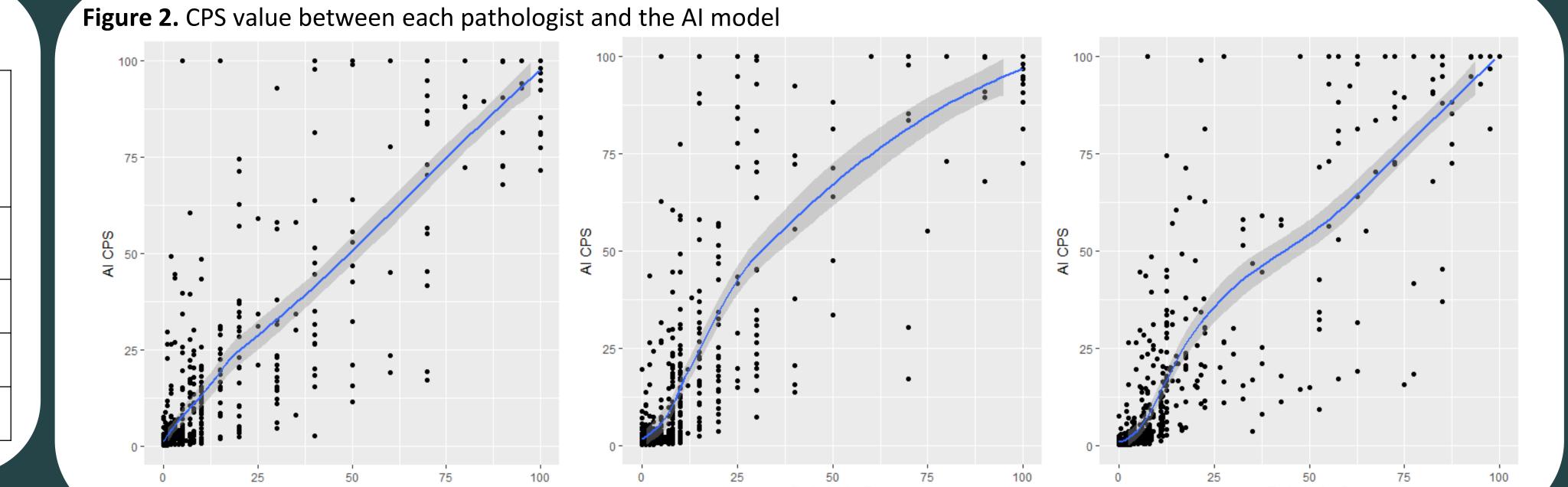


Figure 3. CPS value comparison in individual hospital dataset

Average of CPS (Hostpial B)