

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

Jeong Hwan Park<sup>1</sup>, Soo Ick Cho<sup>2</sup>, Kyu Sang Lee<sup>3</sup>, Euno Choi<sup>4</sup>, Wonkyung Jung<sup>2</sup>, Sukjun Kim<sup>2</sup>, Gahee Park<sup>2</sup>, Sanghoon Song<sup>2</sup>, Cholmin Kang<sup>2</sup>, Minuk Ma<sup>2</sup>, Donggeun Yoo<sup>2</sup>, Kyunghyun Paeng<sup>2</sup>, Chan-Young Ock<sup>2</sup>

<sup>1</sup>Department of Pathology, Boramae Medical Center, Seoul, Republic of Korea  
<sup>2</sup>Oncology, Lunit, Seoul, Republic of Korea  
<sup>3</sup>Pathology, Seoul National University Bundang Hospital, Seongnam-si, Republic of Korea  
<sup>4</sup>Pathology, Ewha Womans University Mokdong Hospital, Seoul, Republic of Korea  
Conflict of interest (Presenting author): none to declare

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

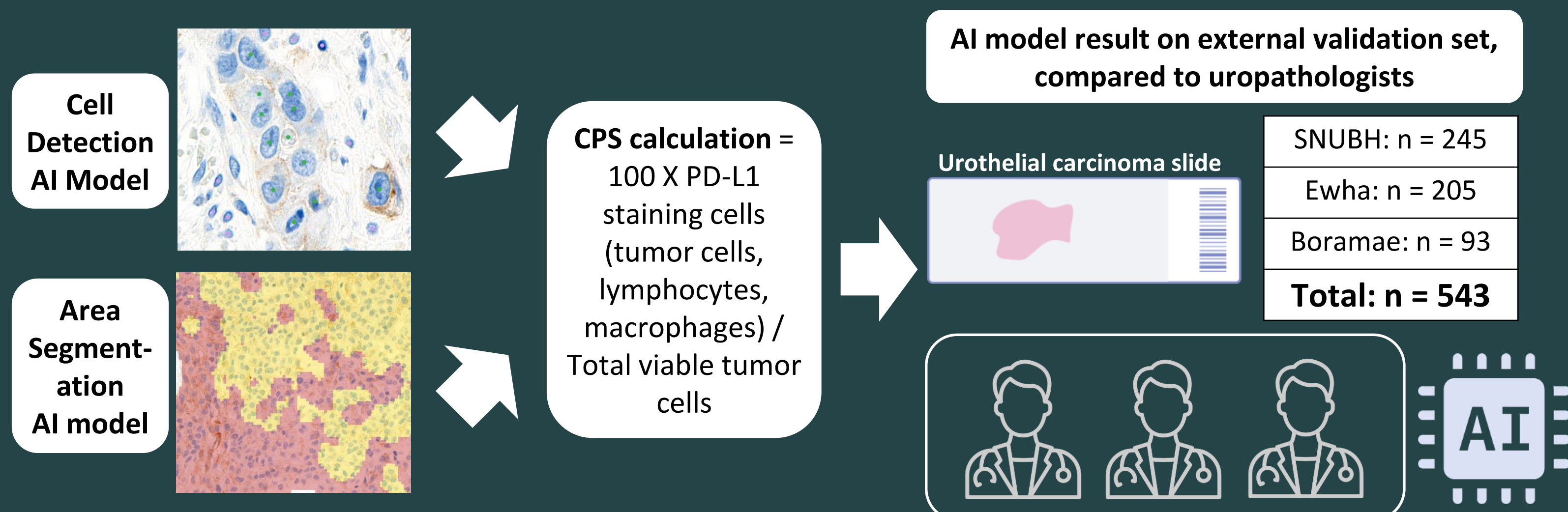
## Background

- 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  $3.02 \times 10^5$  tumor cells and  $3.49 \times 10^5$  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



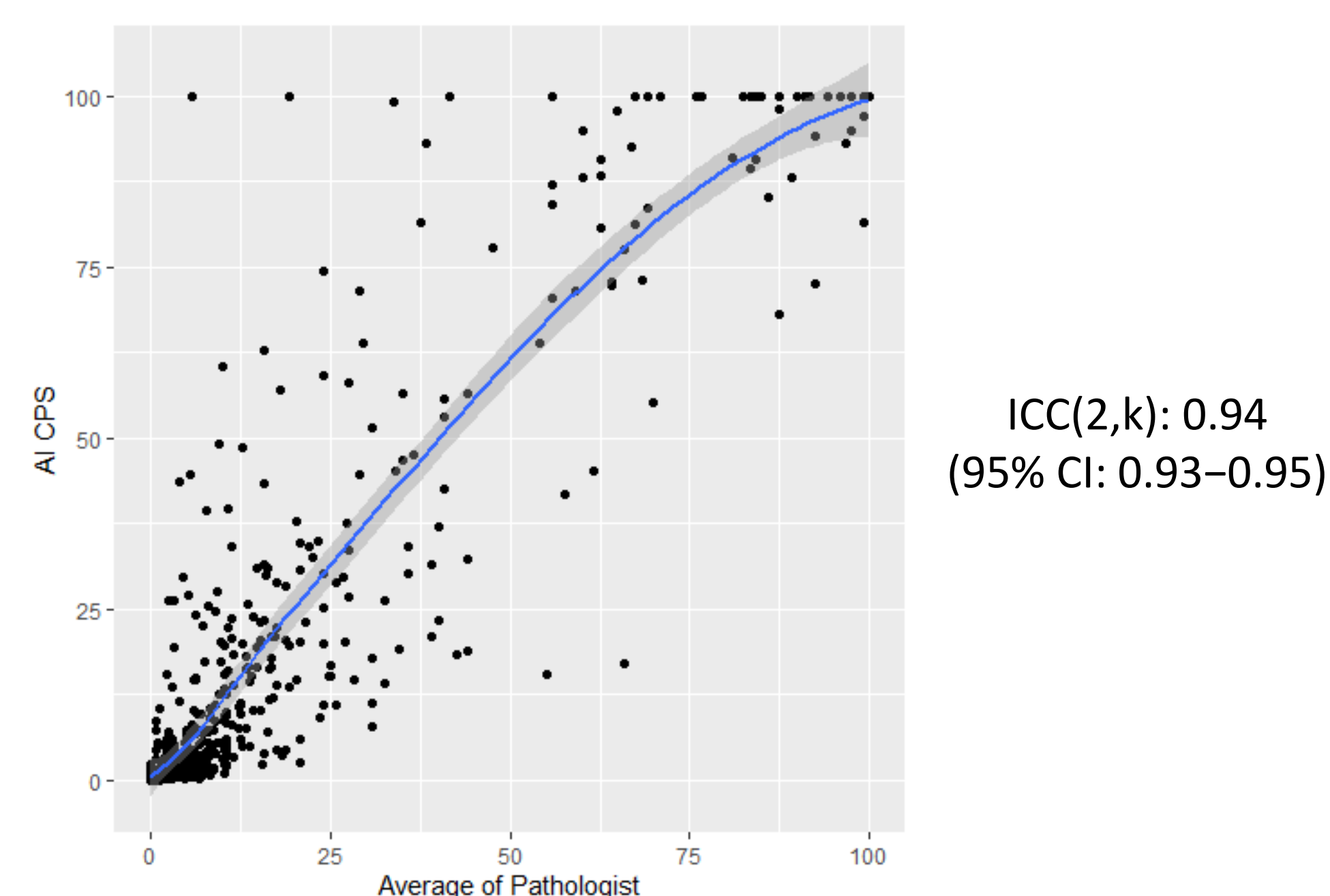
## Results

- 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).
- AI 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 (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% / 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).

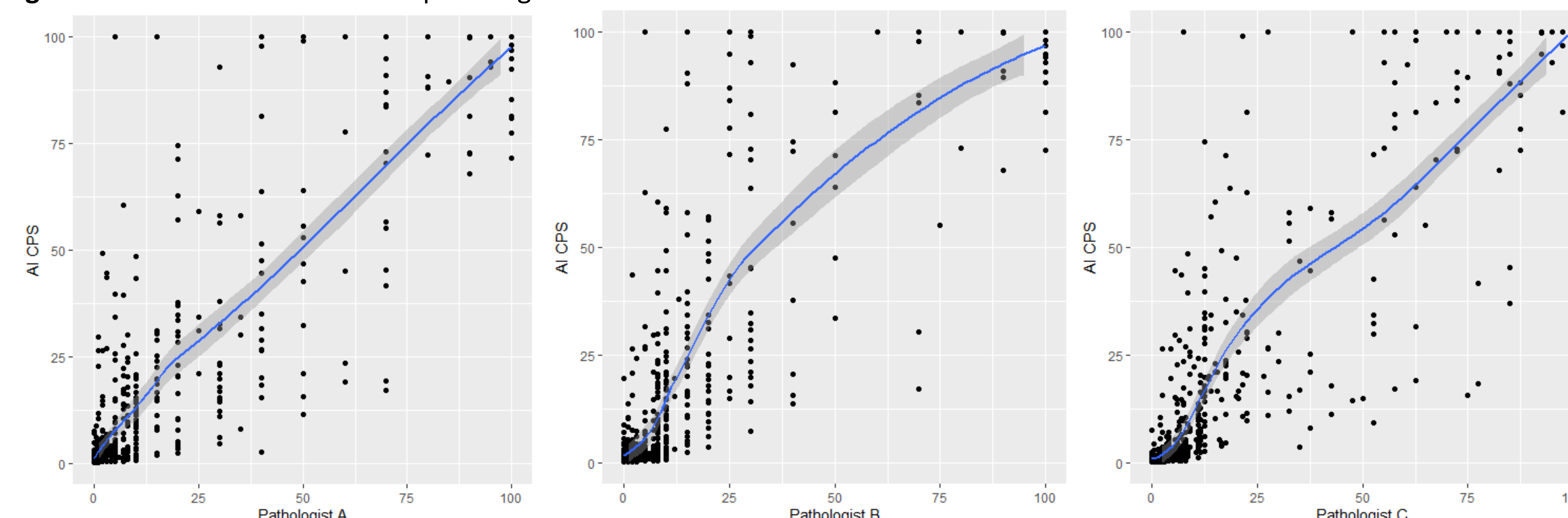
**Table 1.** Comparison of CPS evaluation between each pathologist or between pathologists and the AI model

Comparison	Agreement (more than CPS 10 or not)	Correlation (Intraclass coefficient [2,k], 95% CI)
Pathologists' consensus vs. AI model	88.8%	0.94 (0.93–0.95)
Pathologist A vs. B	87.1%	0.91 (0.90–0.93)
Pathologist A vs. C	89.9%	0.96 (0.96–0.97)
Pathologist B vs. C	87.3%	0.92 (0.90–0.93)

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



**Figure 2.** CPS value between each pathologist and the AI model



**Figure 3.** CPS value comparison in individual hospital dataset

