#2633: AI based PD-L1 CPS quantifier software identifies more patients for checkpoint therapy in gastric cancer at pathologist-level interobserver concordance

Sunil S. Badve[1,*], George Louis Kumar[2,3,*], Tobias Lang[4,*], David Mulder[4], Diego Calvopiña[4], Patrick Frey[4], Maria Karasarides[2,5]

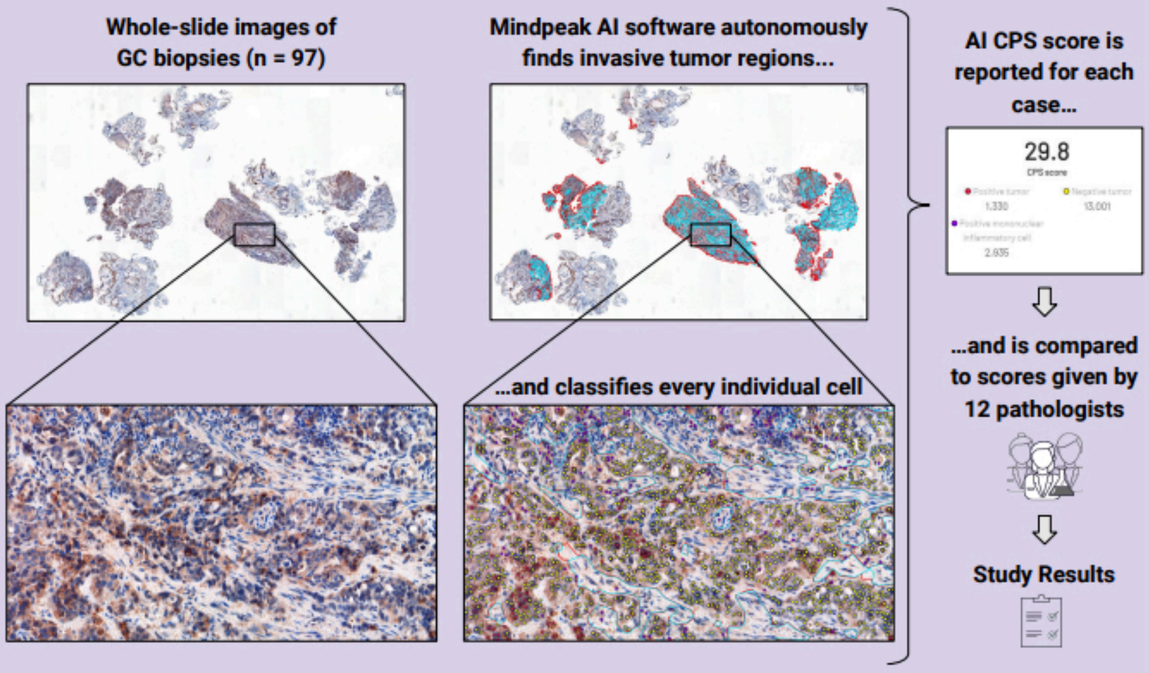
(*) Authors contributed equally to this work. **1** Winship Cancer Institute, Emory University, Atlanta, GA, USA; **2** Affiliation pertains to the time the analysis was conducted: Bristol Myers Squibb, Princeton, NJ, USA; **3** AstraZeneca, US Medical Affairs, Wilmington, DE, USA; **4** Mindpeak GmbH, Hamburg, Germany; **5** The Delphina Group

EMORY UNIVERSITY

2024 ASCO ANNUAL MEETING

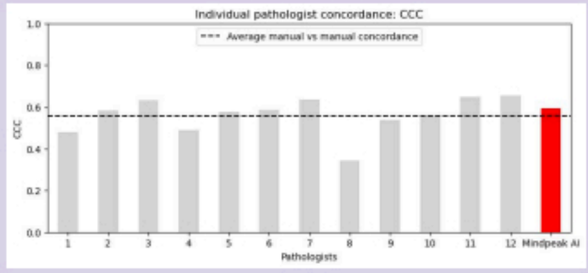
Background: To determine whether gastric cancer (GC) patients are eligible for immunotherapy, PD-L1 expression is analyzed by immunohistochemistry (IHC) using the "Combined Positive Score" (CPS). This requires quantification of PD-L1 stained tumor and tumor-associated immune cells as well as all viable tumor cells. However, manual CPS scoring on whole-slide images (WSIs) is time-consuming and prone to error as evidenced by low interobserver concordance. While the use of artificial intelligence (AI) holds the promise to ameliorate this key challenge in clinical practice, AI models have not yet met the required accuracy thresholds for PD-L1 CPS scoring on GC biopsies.

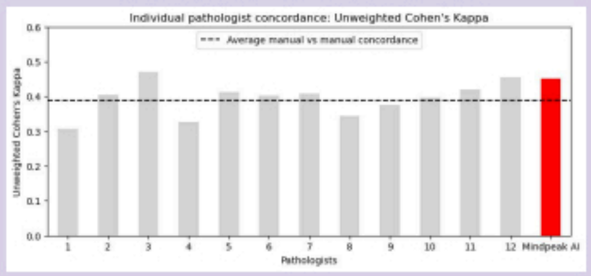
Methods: We investigated the use of an AI-based PD-L1 CPS quantifier software (Mindpeak PD-L1 CPS AI) to support pathologists in standardized PD-L1 IHC assessment on GC biopsies. An AI software for automated PD-L1 CPS scoring was deployed on WSIs from GC biopsies (n = 97) stained for PD-L1 with the 28-8 pharmDx assay and scanned on a 3DHistech P1000 scanner. Manual CPS scores from 12 pathologists on all 97 slides were available for comparison. Pairwise correlation was calculated for continuous values using Lin's concordance correlation coefficient (CCC). Pairwise concordance was measured for scores binarized at the clinically-relevant positivity cutoff of CPS ≥ 5 using unweighted Cohen's kappa.



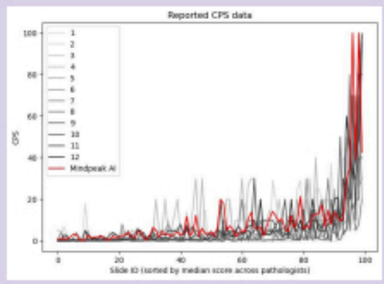
Results:

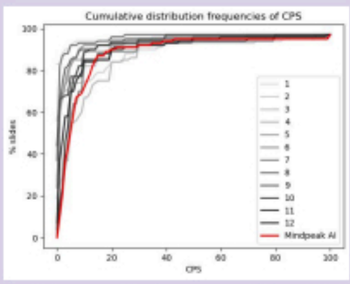
For continuous CPS scores, the **CCC between AI scores and pathologists' scores was higher (0.59) than the mean correlation among pathologists (0.56).** At a cutoff of CPS≥5, the **concordance between AI scores and pathologists' manual scores (κ=0.45) was higher than the mean concordance among pathologists' manual scores (κ=0.39) (p<0.05)**



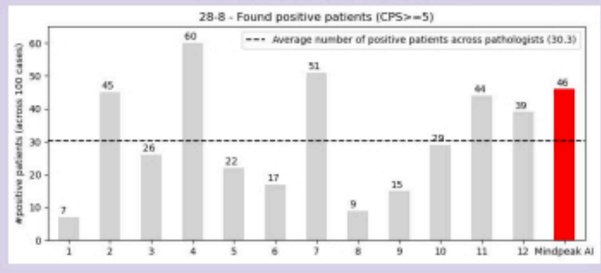


In the majority of cases, the **AI scores were found to be within the range of all pathologists, but slightly above the pathologist median**.





Substantial variability is seen among pathologists when categorizing patients as positive (CPS≥5), with approximately 30.3 ±5.0 patients classified as positive on average by manual scoring and **46 patients (>50% more) categorized as positive by the AI model.**



Conclusions:

An AI model for the assessment of PD-L1 expression in GC using CPS was applied successfully without human intervention. The correlation in continuous CPS scores as well as the concordance in clinical categories with all pathologists was higher for the AI model than for individual pathologists on average, while at the same time, the AI model found more positive patients. This shows that **by using AI more positive patients eligible for PD-L1 targeted treatment might be identified while simultaneously ensuring a level of concordance that is non-inferior to pathologists.**

Correspondence: Sunil S Badve, sbadve@emory.edu; Maria Karasarides, maria@delphina.io