Using Artificial Intelligence to Predict Response to Immunotherapy

Navi Mehra, Joseph Krueger*, Karen Ryall, Jeni Caldera, Will Paces, Kelsey Weigel†
Flagship Biosciences, Inc, Westminster, CO.
* Corresponding author † Presenter

P90

Overview

BACKGROUND

The programmed death ligand 1 (PD-L1)/programmed cell death 1 (PD-1) checkpoint blockade is the backbone for the myriad of combination therapies being developed; thus, effective PD-L1 immunohistochemistry (IHC) testing remains critical for predicting patient response to these therapies. Pathological interpretations applied to PD-L1 IHC as a response biomarker are becoming more complex, going beyond simple tumor proportion score (TPS) and requiring more complex diagnostic algorithms that evaluate the role of PD-L1 expression in tumor cells, immune cells in the tumor microenvironment (TME), and tumor-infiltrating lymphocytes (TILs), all whose spatial relationships are critical for understanding the immune contexture. This complex matrix of several different biological cell types and spatial relationships can quickly become impossible for a pathologist to record and report successfully.

METHODS

Recent advances in computational ability, machine learning algorithms, and data science now allow the application of artificial intelligence (AI) methods to create data-rich profiles from whole-slide images (WSIs) of tissue that capture key tissue context information about PD-L1. In this study, tissue image analysis was applied to increase objectivity and reproducibility of scoring, and AI interpretation of the results was used to increase the accuracy and sensitivity of the diagnostic performance of existing PD-L1 IHC testing methods. In this manner, existing FDA-approved PD-L1 IHC tests can be reevaluated by AI approaches to create value in the clinical setting without changing the IHC assay or causing significant disruption in the normal procedures performed in pathology laboratories.

RESULTS

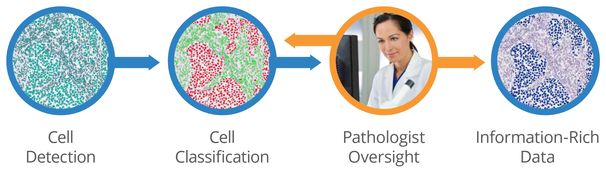
In this study, we demonstrate how our WSI AI platform captures the different attributes of PD-L1–stained slides to create a summary score or output for decision-making. Flagship's computational tissue analysis (cTA®) records the PD-L1 staining, morphological, organizational, and spatial features of a tissue section. When using a cohort of patient tissues with clinical response, the cTA can be measured against clinical response to create a more sophisticated scoring system and a better diagnostic cut point.

CONCLUSIONS

By applying Flagship's cTA AI to existing PD-L1 IHC companion diagnostics, clinical laboratories can now go beyond using tissue image analysis for improving objectivity and reproducibility and can also create entirely new scoring approaches from existing PD-L1 IHC companion diagnostics to improve clinical performance.

FIGURE 1.
Flagship's AI System

Flagship's cTA AI solution incorporates image analysis, machine learning, statistical analysis, and pathologist input throughout the workflow.



	Pathology	Image Analysis	Machine Learning	Data Science
Remove artifacts, necrosis, nontumor areas	☑			
Generate cell data		☑		
Tumor/TME separation			☑	
Biomarker scoring				☑
Patient classification			☑	☑
Data acceptance	☑			

Figure 1. Flagship's AI system integrates image analysis, machine learning, and data science with oversight by a team of board-certified pathologists to produce information-rich data for pathologist review and delivery to client. Pathologists apply the proper controls for cell detection and classification on each slide and accept the final data.

FIGURE 2.
Flagship Biofeatures Generation

The AI builds a Biofeatures profile network to characterize the tissue sample and identify patterns that create distinct patient classes.

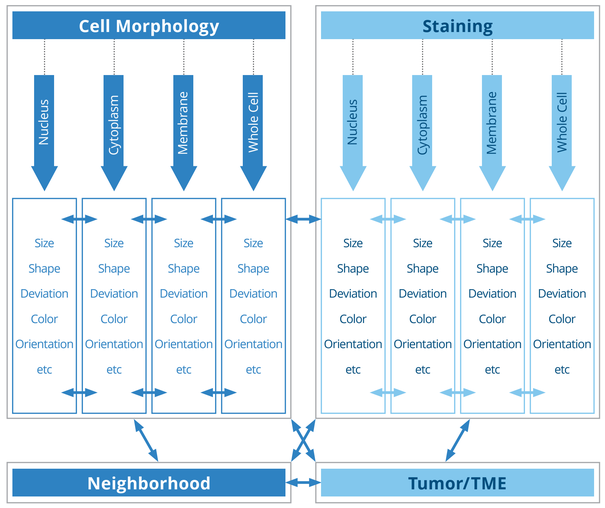


Figure 2. Flagship's Biofeatures are the biologically relevant cellular, staining, and tissue-organization measurements made by the computer algorithm. A Biofeature may be a simple measurement, such as nuclear staining, or a complex, multiparameter measurement, such as the distance between biomarker-positive immune and tumor cells.

FIGURE 3.
Patient-Specific Cell Classification and Tumor-Stroma Separation

The Flagship AI system solves a key problem for any pathology image analysis system: variation between different patient samples.

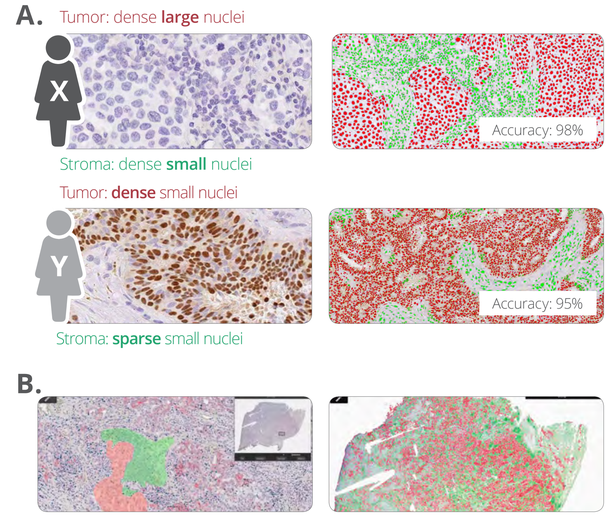


Figure 3. (A) Examples of variance between patient samples. **(B)** Using the guidance provided by the pathologist (left), the cell types in both the tumor and stroma are identified, and machine learning then detects all tumor and TME cells across the WSI (right).

FIGURE 4.
Biofeatures Prediction Classifiers for Flagship Scorecards

Flagship has developed a "Scorecard" approach that assimilates all of the data gathered by the AI system into a summary for visualization and understanding.

AI-Determined Classifier Model

Importance	Predictor	Feature Description
1	Negative	Deviation of a PD-L1 cytoplasmic staining feature in tumor cells. Less feature variability is beneficial.
2	Negative	Nuclear morphology feature of cells in the TME. Represents a feature of the leukocyte population.
3	Positive	Deviation of a PD-L1 cytoplasmic staining feature in tumor cells. Less feature variability is beneficial.
4	Negative	Deviation of a PD-L1 membrane staining feature in tumor cells. Less feature variability is beneficial.
5	Positive	Tumor cell morphology feature. Large differential from nonresponders is beneficial.
6	Positive	Morphology-dependent PD-L1 membrane staining feature in tumor cells. Large differential from nonresponders is beneficial.
7	Positive	TME leukocyte morphology feature. Larger differential from nonresponders is beneficial.
8	Negative	Tumor cell morphology feature. Smaller value is beneficial.

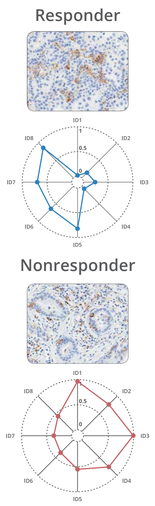


Figure 4. The top unique classifiers that define the patient cohort are established through a guided machine learning process. Different machine learning models and hyperparameters are explored to optimize model performance. Based on the top classifiers, each patient sample is assigned its own visual representation of the data, or Scorecard (see plots on right, bottom). Every Scorecard is unique per patient based on how highly or lowly they express the classifiers. The data shown here are an example of a model of response based on PD-L1 staining.

FIGURE 5.
Predictive Scorecard Model

A per-patient summary score is developed from the Scorecard model to create a prediction score that delivers the actionable data needed for decision-making.

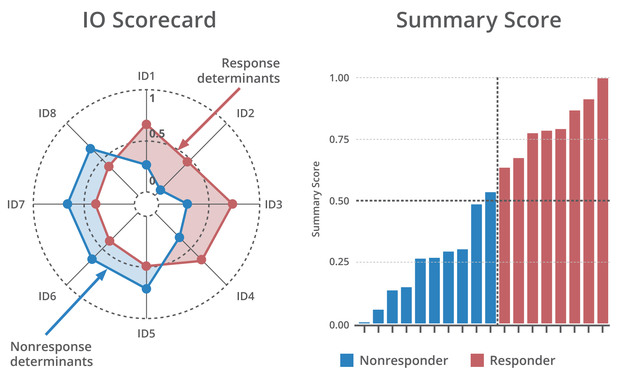
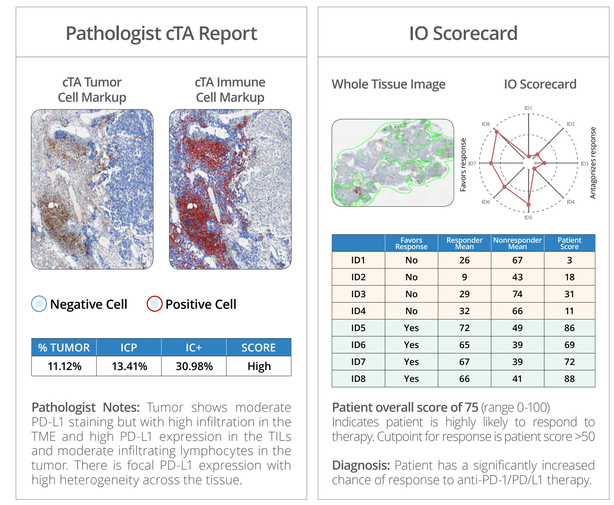


Figure 5. The individual scorecards can be assembled into one average scorecard that represents the overall cohort of patient samples examined. In this case, the top response and nonresponse determinants are shown for the cohort of patients used to model a response based on PD-L1 expression; these patients are considered either responders or nonresponders. In addition, a summary score can be established to find a unique cut point to determine whether a patient sample examined by the same processes will fall into the response or nonresponse category.

FIGURE 6.
Reporting

The AI system delivers both customary pathology and novel Scorecard data to provide patient-specific information.



% TUMOR	ICP	IC+	SCORE
11.12%	13.41%	30.98%	High

Pathologist Notes: Tumor shows moderate PD-L1 staining but with high infiltration in the TME and high PD-L1 expression in the TILs and moderate infiltrating lymphocytes in the tumor. There is focal PD-L1 expression with high heterogeneity across the tissue.

	Favors Response	Responder Mean	Nonresponder Mean	Patient Score
ID1	No	26	67	3
ID2	No	9	43	18
ID3	No	29	74	31
ID4	No	32	66	11
ID5	Yes	72	49	86
ID6	Yes	65	39	69
ID7	Yes	67	39	72
ID8	Yes	66	41	88

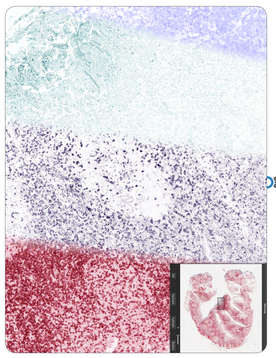
Patient overall score of 75 (range 0-100)
Indicates patient is highly likely to respond to therapy. Cutpoint for response is patient score >50

Diagnosis: Patient has a significantly increased chance of response to anti-PD-1/PD/L1 therapy.

Figure 6. The report will deliver examples of the cTA markup on the tissue slide with a computer-generated score familiar to pathologists along with any relevant pathologist notes (left). The report will also include the Scorecard and a table summarizing the top classifiers that define the patient to accompany the overall prediction score for the patient cohort. The example shown here is from the SP263 assay scoring paradigm for urothelial carcinoma.

Summary Value of Flagship's Pathology AI System to Drug Development

Flagship Partnership



Tissue *Big Data*

Flagship's AI system creates a rich data profile from existing IHC stained slides to understand drug efficacy and patient response.

1. Extract more value from precious tissue sections that cannot be achieved through traditional pathology.
2. Access meaningful data, patterns, and trends not obvious without a complex tissue-context profile.
3. Customize a data solution to your biomarker, your drug, and the indication specific for a hypothesis.
4. Understand tissue biomarker biology on a granular level and distinguish treatment and response criteria.
5. Curate the data captured for future retrospective comparison to clinical outcome data and to aid in development of diagnostic approaches.

Download the Poster