The Utilization of Artificial Intelligence to Produce Clinically Relevant Scores on PD-L1 Immunostained Non-Small Cell Lung Cancer Biopsies

Bhavika Patel, PhD[1], Stephanie Allen[1], Sameer Talwalkar, MD[1], Navi Mehra, PhD[1], Jeppe Thagaard[2], Thomas W. Ramsing[2], Agnete Overgaard, PhD[2], Jenifer Caldara[2]

ABSTRACT

Immunotherapies targeting complex programmed death-ligand 1 (PD-L1) interactions have been groundbreaking in the treatment of non-small cell lung cancer (NSCLC), offering new strategies for patients who are likely to respond. PD-L1 binds to the checkpoint PD-1 to regulate T cells, which has an important role in boosting the immune response. Identifying patients who may benefit from PD-L1 focused care is dependent on the interpretation of the drug's associated companion diagnostic (CDx) immunohistochemistry (IHC) assay. PD-L1 IHC assays stratify NSCLC patients using a Tumor Proportion Score (TPS), a manual scoring paradigm which must be applied by a clinical pathologist. Application of the score is complex as it requires the reader to visually exclude stromal and tumor-infiltrating immune cells across a full sample, retaining PD-L1 expression information for true tumor cells only.

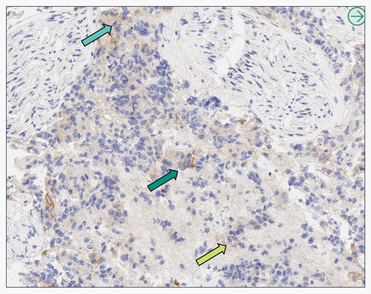
The intricacy of TPS application is time consuming for the pathologist and introduces opportunity for inter- and intra-reader variability. Our novel image analysis approach utilizes artificial intelligence (AI) to automatically denote tumor nests and exclude tumor-infiltrating immune cells. The remaining true tumor cells may then be quantified for PD-L1 expression across the full biopsy, producing a TPS for each sample. This method may allow for increased accuracy and time-efficiency in providing a TPS over traditional methods alone, which may result in improved patient care.

STANDARDIZED SPECIMEN PROCESSING

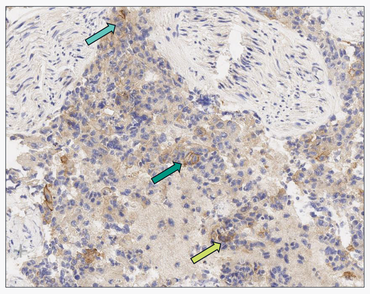
Immunohistochemical Optimization

Formalin-fixed, paraffin-embedded (FFPE) breast cancer specimens were obtained and stained using the PD-L1 [28-8] pharmDx monoclonal antibody. Staining was performed on the Leica Bond III™ automated stainer (Leica Biosystems) and optimized using a matrix of dilutions and antigen retrieval conditions on control and disease state tissues. After staining optimization, a board-certified MD pathologist selected the 1:50 dilution as the optimal condition for staining, which aligns with vendor recommendations. Additionally, antigen retrieval (AR) performed in a pH9 buffer was selected over a pH6 buffer, notable as the vendor recommends use of the lower pH AR buffer. Brightfield 40X whole-slide imaging was performed using the AT2 scanner (Leica Biosystems). All instruments were calibrated and validated for use according to applicable Lanterne Dx SOPs and quality standards. All work was performed at Lanterne Dx's single-site CLIA-certified and CAP-accredited laboratory.

The 28-8 clone was originally optimized using human FFPE commercially purchased blocks. However, antigens may have been compromised, which made the antibody difficult to optimize. Furthermore, compared to the 22C3 clone, the 28-8 clone consistently showed weaker staining, making it difficult to score.



1:150 dilution. This presented less notable background blush, but true membrane staining was more challenging to identify due to the lighter intensity.



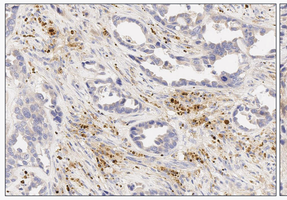
1:50 dilution. This presented a stronger background blush, but true membrane staining was darker and more easily identified.

[1] Lanterne Dx, Boulder CO. [2] Visiopharm Corporation, Westminster, CO.

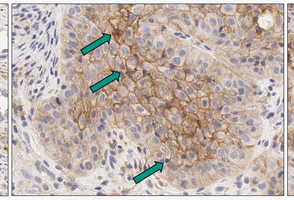
A COMPLEX ASSESSMENT

Manual Interpretation

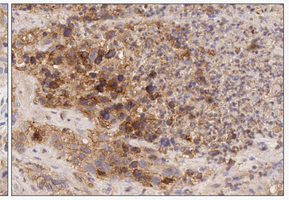
Interpretation of the PD-L1 NSCLC assay requires the pathologist to visually disregard numerous cell types that are similar in appearance to or are intermixed with tumor cells, many of which can also express PD-L1. The identification of only tumor cells is performed in-tandem with tallying the positive to negative tumor cell ratio to produce a percent positive, yielding a Tumor Proportion Score (TPS).



DAB staining in the tumor microenvironment, even though the tumor in the sample is mostly negative (TPS=3%).

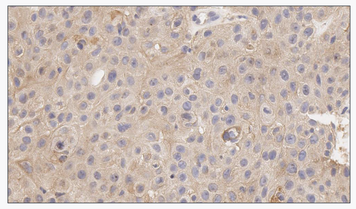


Tumor nest region with likely tumor infiltrating lymphocytes. Non-tumor cells must be excluded during interpretation.

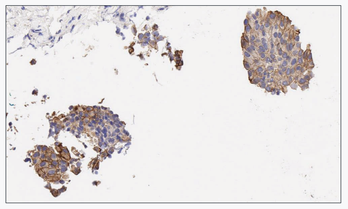


Possible PD-L1 positive tumor cells beginning to intermix with necrotic areas, granular staining appearance.

Additionally, vendor interpretation criteria specifies that only complete circumferential or partial linear plasma membrane staining is considered positive for scoring purposes. Cytoplasmic staining, if present, should not be incorporated. In this study, any sample displaying membrane expression also displayed some level of cytoplasmic expression, often at similar intensities which makes deciphering only membrane staining challenging. Furthermore, there is no standardized approach for interpretation of PD-L1 staining among the most commonly used antibody clones (22C3, 28-8, SP142 and SP263). The 22C3 clone is the most used in pathology labs and therefore, preferred by most pathologists.



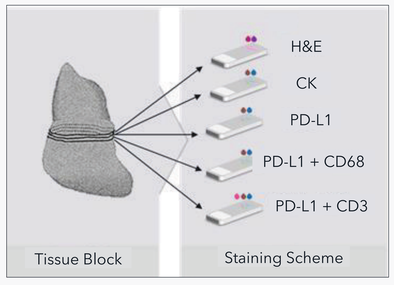
Specimen with notable cytoplasmic staining that sometimes gives the visual impression of where the membrane begins, even though the membranes are, generally, negative.



Specimen where even though cytoplasmic staining is present, linear membrane expression is distinguishable.

Designing an AI Model

To mimic this complex interpretation, it is key that the algorithm's (APP's) design can identify cell types of interest, as well as distinguish true membrane staining from cytoplasmic blush. A deep learning network was trained to identify tumor nest regions for further analysis and exclude the tumor microenvironment (TME), areas of necrosis, and normal tissue. A secondary network was trained to classify cell types within tumor nest regions so that only tumor cells are then quantified for PD-L1.



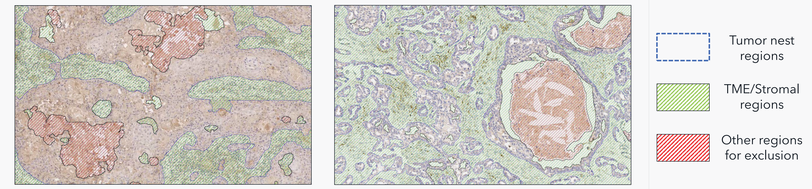
Serial section staining scheme performed on specimens that were used to place annotations to train the AI APP.

Serial sections that were strategically stained to identify the areas and cell types noted in the manual paradigm above. Images of the serial sections were virtually aligned using Visiopharm's TissueAlign module, a patented method for virtual multiplex staining. Transfer learning was then used to relay annotations of stained cell types (ground truth) to the PD-L1 stained section, allowing the AI model to be trained on and for use in the PD-L1 IHC stained monoplex.

RESULTS

Regional Segmentation

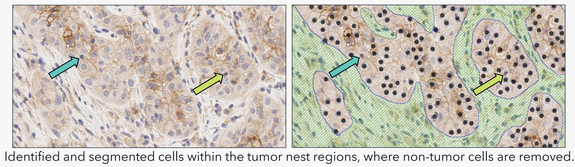
Visiopharm's deep learning network for PD-L1 regional segmentation in this research was not previously exposed to or trained on the sample set obtained and processed by Lanterne Dx. This APP was utilized to automatically find tumor nest areas for subsequent cellular assessment, while also identifying TME, necrosis, and other non-tumor areas for exclusion.



Two specimens demonstrating varying tumor morphology and staining patterns with AI markups overlayed.

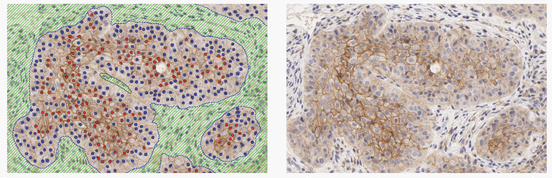
Cellular Identification and Quantification

The AI APP for finding tumor cell to produce a TPS has two key requirement. First, the APP must identify only tumor cells and overlook other cell types, such as lymphocytes, macrophages, and other non-tumor cells that may be present in the tumor nest regions.



Identified and segmented cells within the tumor nest regions, where non-tumor cells are removed.

Second, the APP must distinguish cytoplasm staining from membrane staining, even when the membrane staining is partial. A DAB staining intensity threshold alone is not sufficient for the identification. Instead, this APP is designed to have greater control over the recognition of positive membranes using requirements for the amount a cell must express staining and restricting this to the membrane area only through specific denotation of nuclear, cytoplasm, and membrane areas per cell.



Tumor region with AI markup overlayed to denote cells with positive membrane expression (red) and negative cells with no expression or only cytoplasmic expression (blue)

CONCLUSIONS

We utilized virtual multiplexing to develop an automated, AI workflow to assess PD-L1 Tumor Proportion Scores (TPS) and investigated it on NSCLC specimens that have undergone standardized processing.

- Due to the wide range of PD-L1 expression location and intensity, we recommend the utilization of a laboratory which employs assay optimization methods while following quality assurance procedures.
- The complexity of this PD-L1 scoring paradigm and initial results of the AI suggest that there is utility of this tool to properly assess TPS. To further verify this performance, comparison of the AI-supported scores to manual reads with inter- and intra-reader variables is proposed as a Phase 2 of this study.