Deep learning identifies pathobiological features within H&E images associated with genomic alterations and progression on anti-PD(L)1 in HUDSON, an AstraZeneca-sponsored Phase II clinical trial

Laura Dillon[1]‡, Marylens Hernandez[1]‡, Ben Glass[2], Guillaume Chhor[2], Sara Hoffman[2], Varsha Chinnaobireddy[2], Sai Chowdary Gullapally[2], Kris Sachsenmeier[1], Andy Beck[2]*, Jason Hipp[1]*

[1]Translational Medicine, Oncology R&D, AstraZeneca, Gaithersburg, MD, USA; [2]PathAI, Boston, MA, USA

‡ These authors contributed equally; *Senior and corresponding authors

#LB016

Background

- Novel biomarkers are needed to better predict which patients will respond to immunotherapy.
- Machine learning (ML) models have the potential to quantitatively characterize the tumor and tumor microenvironment (TME). However, most ML models require training to be done on a subset of samples from a novel dataset before deployment on the remainder of the dataset. This training-test approach requires a larger number of samples than is generally available for small clinical trials. There is a need for pre-trained ML models that can be applied to small datasets.
- PathAI previously trained ML models on NSCLC samples from commercial and clinical datasets to identify and quantify cellular composition, tissue architecture, and blood vessel features in the TME.
- Here we deployed PathAI's ML models to H&E images from HUDSON (NCT03334617), an AstraZeneca Phase II platform clinical trial, to identify morphologic features associated with genomic alterations and time to progression on anti-PD(L)1 therapies.

Methods

HUDSON trial samples

- HUDSON is an international, multi-site AstraZeneca-sponsored Phase II platform clinical trial of novel anti-cancer agents in subjects with mNSCLC who have progressed on anti-PD(L)1-containing therapy prior to entering HUDSON. Biopsies were collected across multiple body sites and taken both pre- and post-checkpoint progression, as well as following treatment with novel compounds in HUDSON.

HIF generation

- With no additional training, PathAI's ML models for tissue, cell, and blood vessel identification were deployed on digitized whole slide images (WSIs) of H&E-stained biopsies from HUDSON (**Fig. 1**).
- Human interpretable features (HIFs) were generated which characterize the cell composition and tissue architecture of each sample.

Model performance evaluation

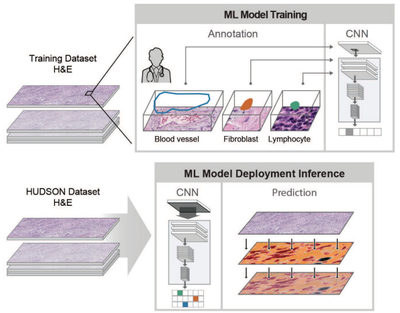
- Model performance was validated using 300, 150x150-micron-sized "frames" or regions of WSI, and exhaustive annotations were generated by the model and from 5 pathologists. Pathologist consensus scores were compared with the ML-model scores for evaluation of agreement.

HIF association analysis

- Only baseline samples, obtained prior to enrollment in HUDSON, which met minimum quality thresholds were used for HIF association analyses.
- To identify HIFs associated with HLA loss of heterozygosity (LOH), a linear model was applied followed by FDR correction at the feature level.
- Following clustering analysis to reduce feature redundancy, HIFs significantly associated with weeks to progression on anti-PD(L)1 therapy prior to enrollment in HUDSON were identified using Cox regression analysis. To determine cluster-level associations, the Browns method was used to merge p-values and FDR correction was applied at the cluster level.

Footnotes:
Abbreviations: ESI – epithelial-stromal interface; HIF – human interpretable feature; LOH – loss of heterozygosity; ML – machine learning; NSCLC – non-small cell lung carcinoma; TME – tumor microenvironment; WSI – whole slide image

Figure 1. PathAI model deployment on the HUDSON dataset

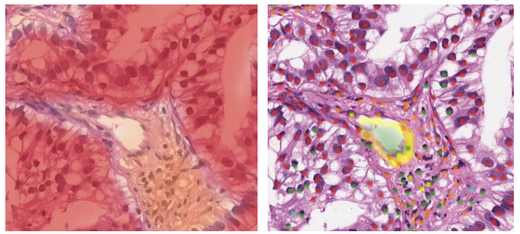


Results

Performance of ML models on the HUDSON dataset

- H&E images from 169 samples were analyzed to identify tissue regions, cells types, and blood vessels (**Fig. 2**).
- More than 4500 raw and derived HIFs were generated to describe the features of the TME from each sample.
- Quality assessment based on pathologist annotations confirmed that the ML models showed high performance when deployed on the HUDSON dataset without additional training (**Table 1**).

Figure 2. Visualization of cell and tissue features on H&E image



Left: Tissue level predictions of cancer epithelium (red) and cancer stroma (orange). **Right**: Blood vessel (yellow with green lumen) and cell predictions of cancer cells (red), lymphocytes (green), plasma cells (yellow), macrophages (aqua), and fibroblasts (orange).

Table 1. Performance of PathAI models

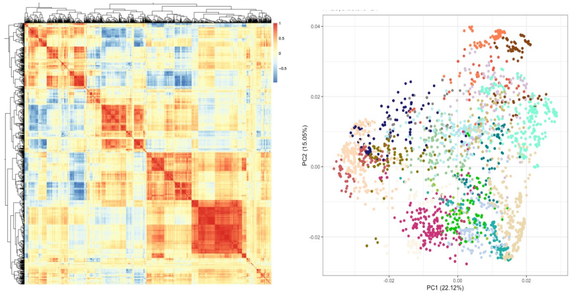
Cell Type	Model	Pathologist
Cancer Cells	0.88 [0.85 – 0.90]	0.88 [0.83 – 0.90]
Immune Cells (Lymphocytes, Plasma Cells, Macrophages)	0.92 [0.90 – 0.94]	0.87 [0.83 – 0.90]
Fibroblasts	0.66 [0.59 – 0.72]	0.63 [0.57 – 0.68]

Pearson correlation values for cell counts in evaluated frames. Model column indicates predicted ML model count vs the consensus of 5 pathologists. Pathologist column indicates each individual pathologist vs the consensus of the remaining pathologists.

Cluster analysis

- Cytology samples, samples without significant tissue, and those taken after the start of treatment with novel compounds as part of the HUDSON trial were excluded. Data from 89 samples were analyzed.
- Clustering analysis (n = 20 clusters) revealed significant correlation between individual HIFs, providing rationale for dimensionality reduction in subsequent analyses (**Fig. 3**).

Figure 3. HIF clustering analysis

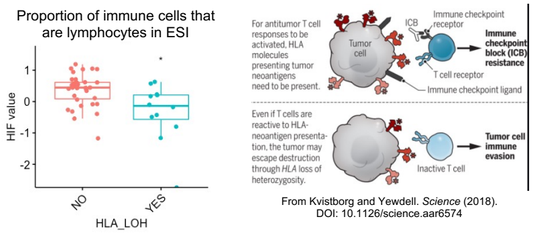


Heatmap visualization (**Left**) and PCA visualization (**Right**) of clustering analysis with preselected n=20 clusters. Each row on the heatmap represents a HIF. Each point on the PCA represents a HIF, colored by cluster number (from 1-20).

Association with class I HLA LOH

- Following correction for biopsy timing and location, a total of 30 HIFs were significantly associated ($p < 0.05$) with class I HLA LOH including increased proportion of plasma cells and decreased proportion of lymphocytes (**Fig. 4**).

Figure 4. Example feature associated with class I HLA LOH



From Kvistborg and Yewdell. *Science* (2018). DOI: 10.1126/science.aar6574

Association with time to progression on anti-PD(L)1 therapy

- Following correction for biopsy timing and location, a total of 59 HIFs were significantly associated ($p < 0.05$) with weeks to progression on anti-PD(L)1 therapy, prior to enrollment in HUDSON.
- Significant HIF clusters include those related to the proportion of cancer cells in the epithelial-stromal interface (ESI), the presence of macrophages/fibroblasts near cancer cells (**Fig. 5**), plasma cell infiltration, and blood vessel compression (**Table 2**).

Figure 5. Example feature associated with weeks to progression on anti-PD(L)1 therapy

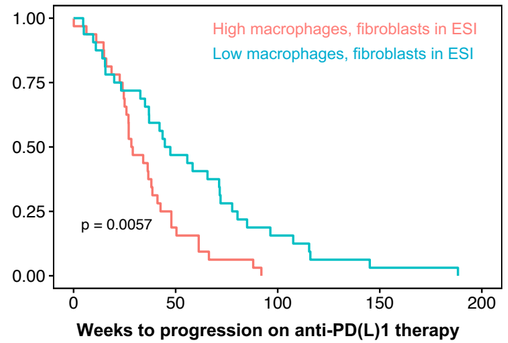


Table 2. Clusters associated with progression on anti-PD(L)1 therapy

HIF Cluster Description	HR (conf.lower, conf.higher)
Proportion of cells that are cancer cells in ESI	1.8 (1.2,2.6)
Proportion of macrophages proximal to cancer cells in ESI	1.7 (1.2, 2.4)
Proportion of plasma cells proximal to cancer cells in ESI	1.6 (1.2,2.1)
Proportion of macrophages, fibroblasts proximal to cancer cells in ESI	1.5 (1.1, 2.1)
Compression of blood vessel lumen or vessel with lumen in tissue	0.65 (0.45, 0.84)

Conclusions

- PathAI models were able to identify TME-associated features from H&E-stained WSIs from a Phase II clinical trial which were associated with genomic alterations and progression on checkpoint inhibitor therapy.
- These results suggest the power of deploying pre-trained ML-based systems in a clinical trial setting to identify pathobiological features associated with tumor characteristics and treatment response from only H&E images.

References: Available upon request

Acknowledgements

This study was funded by AstraZeneca. We wish to acknowledge HUDSON Study Team members Helen Ambrose, Jan Cosaert, Zoltan Szucs, Rosalind Hobson, Marlene Dressman, and Carolyn Hayward, as well as Maria Udriste for preparation of the HUDSON WSIs.

Disclosures

LD, MH, KS, and JH are employees of, and hold stock in AstraZeneca. BG, GC, SH, VC, SCG, and AB are employees of, and hold stock in PathAI.

Contact Email: andy.beck@pathai.com, jason.hipp@astrazeneca.com

Presented at the American Association for Cancer Research (AACR) Virtual Meeting; April 10–15, 2021