Concordance analysis of AI-powered CD8 quantification and automated CD8 topology with manual histopathological assessment across seven solid tumor types

Abstract #1282

Maria Guramare[1]*, Nishant Agrawal[1]*, George Lee[2]*, Adam Stanford-Moore[1], Abhik Lahiri[1], Diksha Meghwal[1], Darren Fahy[1], Raymond Biju[1], Archit Khosla[1], Jacqueline Brosnan-Cashman[1], Dimple Pandya[2], Scott Ely[2], Jimena Trillo-Tinoco[2], John Wojcik[2], Falon Gray[2], Benjamin Chen[2], Sergine Brutus[1], Ben Glass[1], Cyrus Hedvat[1], Ilan Wapinski[1], Mike Montalto[1], Andrew H. Beck[1], Charles Biddle-Snead[1]^, Vipul Baxi[2]^

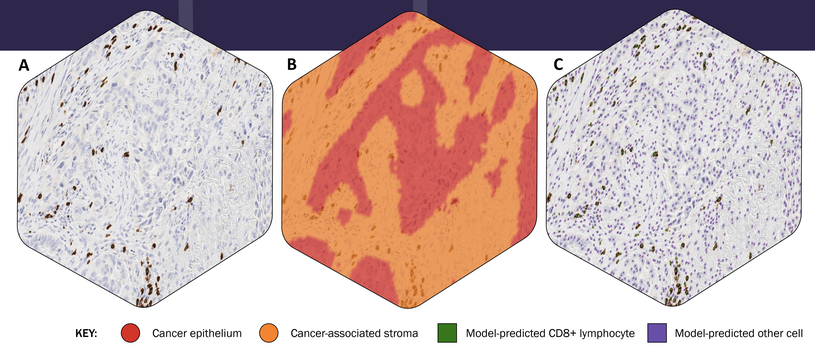
STUDY BACKGROUND

The degree of CD8+ lymphocyte infiltration into the tumor microenvironment, as well as the distribution of lymphocytes within the tumor and surrounding stroma (inflamed, excluded, or desert immunophenotypes), are key determinants for the potential efficacy of immunotherapy[1-3]. Thus, accurate characterization of the tumor immune microenvironment is essential. However, manual histopathological assessment of CD8 topology is subject to many challenges, including subjectivity and reproducibility[4,5].

We previously developed ML-based models for the identification and quantification of CD8+ lymphocytes[6] and topology[7] in melanoma. Here, we developed ML-based models for the identification and quantification of CD8+ lymphocytes and CD8 topology classifiers across seven cancer types: urothelial carcinoma (UC), head and neck squamous cell carcinoma (HNSCC), non-small cell lung cancer (NSCLC), gastric cancer (GC), colorectal cancer (CRC), pancreatic cancer (PC), and hepatocellular carcinoma (HCC).

PATHAI CD8 MODELS

Figure 2. Example representations of PathAI cell and tissue models as deployed in NSCLC. A) CD8 immunohistochemistry. B) Model-detected cancer epithelium and cancer-associated stroma. Regions of necrosis, also detected by the model, are not pictured. C) Model-detected CD8+ lymphocytes and other cells.



METHODS

Patient samples and model generation. Our general workflow is summarized in **Figure 1**. ML algorithms were developed to quantify CD8+ lymphocytes in UC, HNSCC, NSCLC, GC, CRC, PC, and HCC specimens from clinical trials and commercial sources (N=1603) using CD8+ cell detection models trained on digitized whole slide images of CD8 immunohistochemistry (Agilent-Dako clone C8/144b). Annotations were provided by the PathAI network of expert pathologists to train algorithms for classifying tissue regions (including parenchyma and stroma) and cell types (**Figure 2**).

Figure 1. Study workflow. Convolutional neural network (CNN)-based ML models were trained using manual annotations provided by pathologists identifying features including cancer, necrosis, cancer stroma, non-CD8+ lymphocytes and CD8+ lymphocytes. Inference on study samples generates tissue- and cell-level predictions. Examples of tissue-level predictions include stroma and parenchyma. Examples of cell-level predictions include CD8+ lymphocytes and non-CD8+ lymphocytes.

Frames analysis. To evaluate the performance of the CD8+ cell model, AI-predicted counts were compared to a consensus count (ground truth) from five independent pathologists for representative fields of view (frames) using Pearson correlation[8] (**Figure 3**). Inter-pathologist agreement was also calculated.

Development and assessment of CD8 topology classifier. Using a distinct cohort of samples for each cancer type (N=1690), a simple, two-parameter CD8 topology classifier was trained to predict slide-level labels of inflamed, excluded, and desert using pathology-provided labels and derived features. Classifier-predicted topology labels were compared to pathologist-generated labels using unweighted Cohen's kappa.

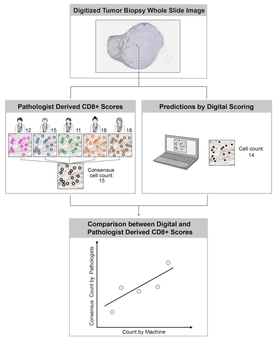


Figure 3. Schematic of frames validation workflow.

RESULTS

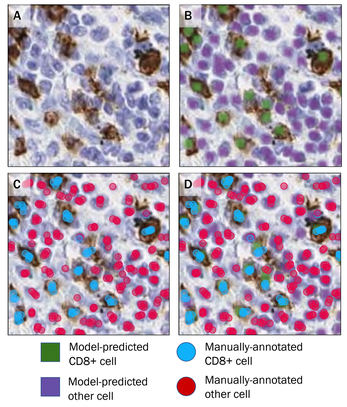


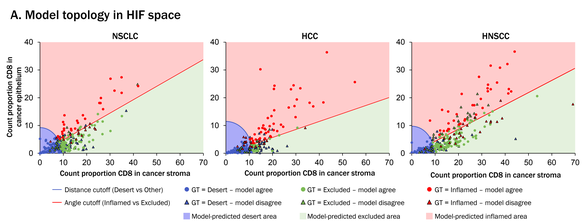
Figure 3. Example of frames analysis for CD8+ cell model performance evaluation in NSCLC. A) Sampled frame within a WSI of CD8 immunohistochemistry in NSCLC. B) Deployment of ML cell model on sampled frame for model prediction of CD8+ lymphocytes (green) and other cells (purple). C) Ground truth consensus pathologist-derived annotations of CD8+ lymphocytes (red) and other cells (blue) from 5 board-certified pathologists. D) Merged overlay of ML-based and pathologist-based CD8 predictions.

Table 1. Frames-based quantitative validation of CD8+ cell model performance.

Cancer type	Inter-pathologist (Mean pathologist vs ground truth) [Pearson, 95% CI]	PathAI Model vs. Ground Truth [Pearson, 95% CI]
UC	0.93 [0.90 – 0.95]	0.95 [0.92 – 0.97]
HNSCC	0.95 [0.92 – 0.97]	0.95 [0.92 – 0.97]
NSCLC	0.95 [0.92 – 0.97]	0.96 [0.93 – 0.97]
GC	0.89 [0.82 – 0.93]	0.91 [0.86 – 0.95]
CRC	0.95 [0.92 – 0.96]	0.95 [0.92 – 0.97]
PC	0.91 [0.87 – 0.94]	0.92 [0.88 – 0.95]
HCC	0.94 [0.90 – 0.96]	0.96 [0.94 – 0.98]

For each cancer type, we measured: 1) the correlation between pathologists and the ground truth consensus score and 2) the correlation between our model and the ground truth consensus score. Pearson correlation coefficients and 95% confidence intervals are shown.

RESULTS



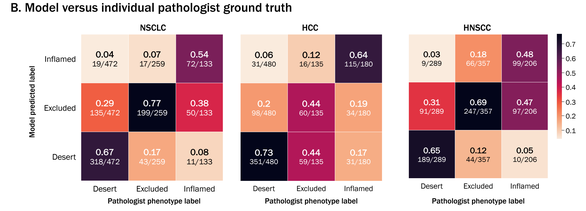


Figure 4. Topology and confusion matrices. A) Topology matrices of phenotype model predictions in transformed polar coordinates space, with data points representing pathologist consensus ground truth labels and shading representing model predictions. B) Confusion matrix measuring recall (true positive/combined true positive and false negative) for model predictions versus consensus ground truth.

Table 2. Quantitative feasibility assessment of CD8 topology labeling.

Cancer type	Inter-pathologist (Pairwise) [Kappa, 95% CI]	PathAI Model vs. Individual Pathologist (Pairwise) [Kappa, 95% CI]
UC	0.47 [0.39, 0.56]	0.48 [0.41, 0.56]
HSNCC	0.37 [0.32, 0.42]	0.41 [0.35, 0.46]
NSCLC	0.39 [0.33, 0.47]	0.47 [0.41, 0.53]
GC	0.45 [0.37, 0.53]	0.41 [0.31, 0.51]
CRC	0.35 [0.28, 0.42]	0.38 [0.29, 0.46]
PC	0.30 [0.18, 0.43]	0.25 [0.14, 0.36]
HCC	0.33 [0.25, 0.41]	0.39 [0.31, 0.46]

Inter-rater reliability of topology assignments across pathologists, and between our model and individual pathologists. Cohen's Kappa and 95% confidence intervals are shown for each cancer type.

CONCLUSIONS

ML model-predicted CD8+ cell counts are highly correlated with pathologist-generated counts across seven solid tumor types, and tumor CD8 topologies were predicted with a simple and highly interpretable two-parameter classifier. These data demonstrate the power of AI-powered digital pathology for accurate and reproducible quantitation of CD8+ lymphocytes and automated immunophenotyping in clinical samples, further confirming the potential for AI-based biomarker measurements in immuno-oncology.

AFFILIATIONS

[1] PathAI, Boston, MA
[2] Bristol Myers Squibb, Cambridge, MA

*Primary authors
^ Corresponding authors

CONTACT

George Lee: george.lee@bms.com
Sergine Brutus: sergine.brutus@pathai.com

REFERENCES

1. Waldman AD, et al. *Nat Rev Immunol*. 2020; 20:651-668.
2. Peske JD, et al. *Adv Cancer Res*. 2015; 128:263-307.
3. Ji RR, et al. *Cancer Immunol Immunother*. 2012;61:1019-1031.
4. Van Bockstal MR, et al. *Mod Pathol*. 2021; 34:2130-2140.
5. Trimm T, et al. *Acta Oncol*. 2018; 57:90-94.
6. Glass, B, et al. *J Immunother Cancer*. 2021; 9(Suppl 2):A859.
7. Lee, G, et al. *J Immunother Cancer*. 2021; 9(Suppl 2):A420
8. Beck, AH, et al. *J Immunother Cancer*. 2019; 7(Suppl 1):120.

ACKNOWLEDGMENTS

We thank Michael Drage, Mevlana Gemici, and Aryan Pedawi for their contributions to this study and BioScience Communications for assistance with figure design.

We thank the patients, families, and participating investigators in the clinical trials used in this study. All trials were sponsored and funded by Bristol Myers Squibb, and data analysis was performed by PathAI. This work was funded by PathAI and Bristol Myers Squibb.

This poster template was developed by SciStories LLC. https://scistories.com/

Society for Immunotherapy of Cancer, November 8-12, 2022, Boston, MA