

Development of Predictive Models for Expression of a Tumor Specific Biomarker and CD3 on H&E Stained Digital Slides



K Bairavi¹, A Shah¹, A Muthuswamy², M Bonham¹, C W Lee², B Higgs², T W Chittenden¹, A Jerusalmi¹, ¹BioAl Health Inc, Manchester, NH, USA; ²Genmab Inc, Princeton, NJ, USA

INTRODUCTION

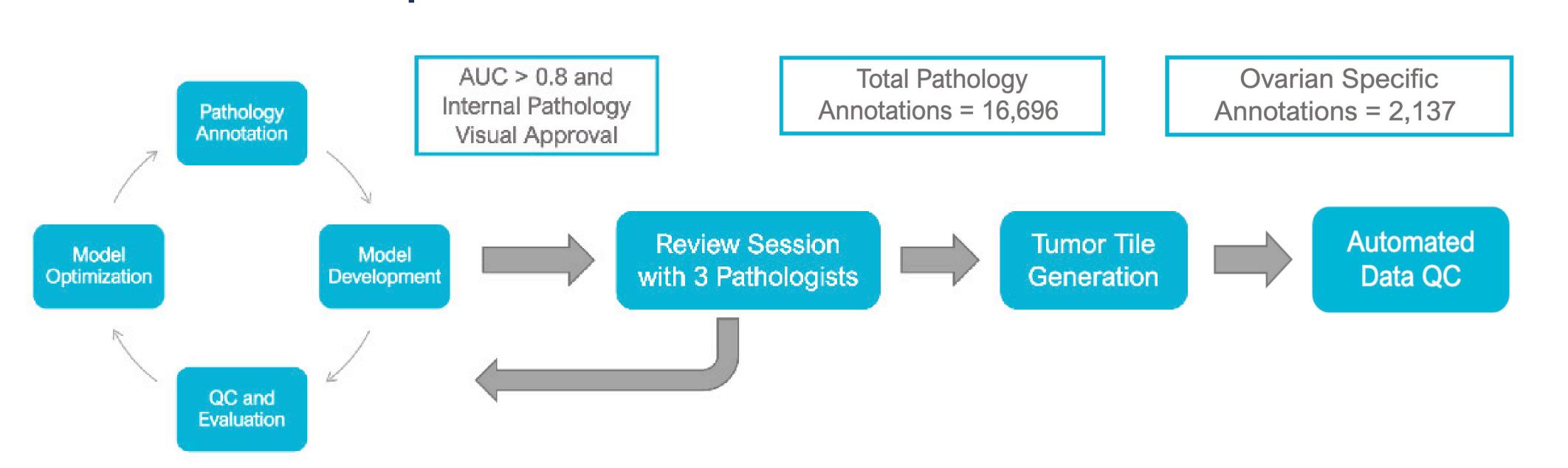
Characterization of a patient's tumor microenvironment is fundamental to advancing translational strategies in immuno-oncology. Histopathological evaluation of tissue slides from patients can provide this invaluable information, informing disease heterogeneity, tumor contexture, and target expression - all important to identifying patients more likely to benefit from therapy. However, availability of sufficient tissue is often a challenge to produce such a comprehensive tumor characterization. Measurement of target expression using an H&E slide could greatly reduce the usage of tissue for IHC making sections available for other investigative purposes. Here, an Al-driven predictive model was developed on H&E digital slides to simulate the expression of a tumor specific biomarker and CD3 across 4 different tumor types.

MATERIALS & METHODS

Data Used:

- 100 cases for each of the four epithelial tumor types (ovarian and three others), were used in this study.
- Each case consisted of three serial sections: H&E, IHC CD3 and a tumor specific IHC biomarker.
- The Bio-Al Predict-X platform was utilized for the optimization of separate tumor models using pathology annotations as basis for development.

Tumor Model Development:



IHC Quantification:

The BioAl quantification algorithm counts number of CD3 positive cells in all tiles from selected tumor region, including adjacent normal tissue; this results in a score of # CD3 positive cells per tile. For the tumor specific biomarker the BioAl quantification algorithm counts the number of positive and the number of negative cells in all tiles within tumor regions, including also adjacent normal tissue; this allows to compute a % positive score.

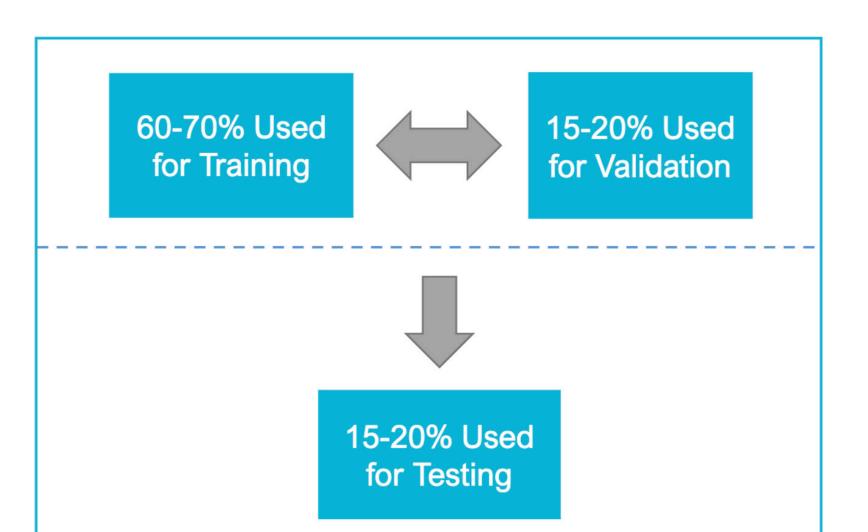
3-Step Predictive Model Development:

- Initial work focuses on finding optimal Deep Learning architecture for the data set and intended use
- Multiple marker thresholds are tested for best signal differentiation between positive / negative
 This allows for future model optimization using
- Models run on multiple cuts of data to account for observed data heterogeneity and biological variation

thresholds that match clinical intended use

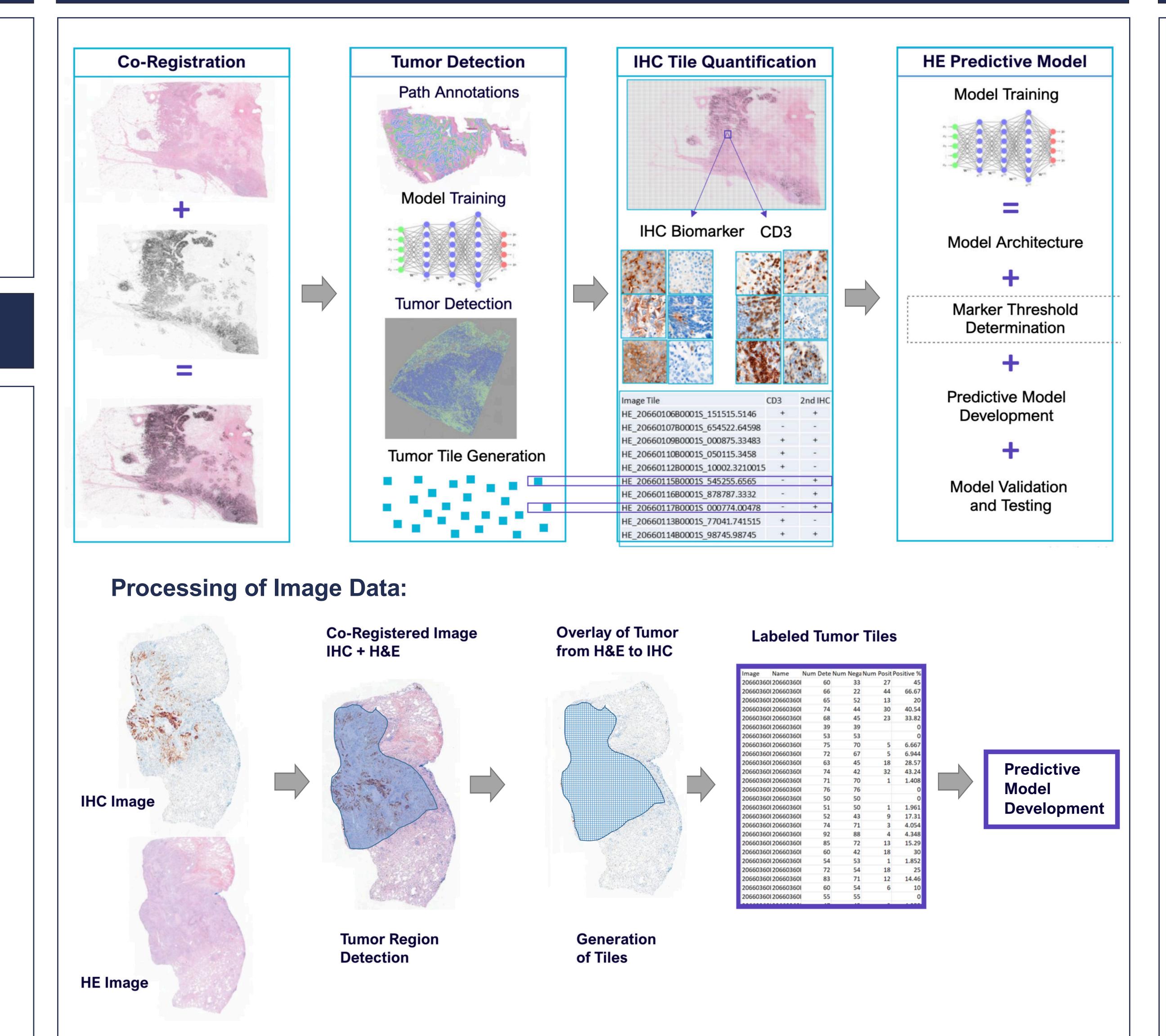
IHC Predictive Model Development:

(Tiled data set randomly split across groups)



Run 5 cuts • Repeat best cut-off across multiple cuts of data to ensure model performance

ANALYSIS WORKFLOW



TUMOR MODEL AND IHC QUANTIFICATION

Tumor Model Performance:

Ovarian Tumor Model Overall Accuracy: 0.958
Breast Tumor Model Overall Accuracy: 0.897
Cervical Tumor Model Overall Accuracy: 0.967
Lung Tumor Model Overall Accuracy: 0.836

IHC Quantification and Correlation:

CD3	Tumor Specific Biomarker							

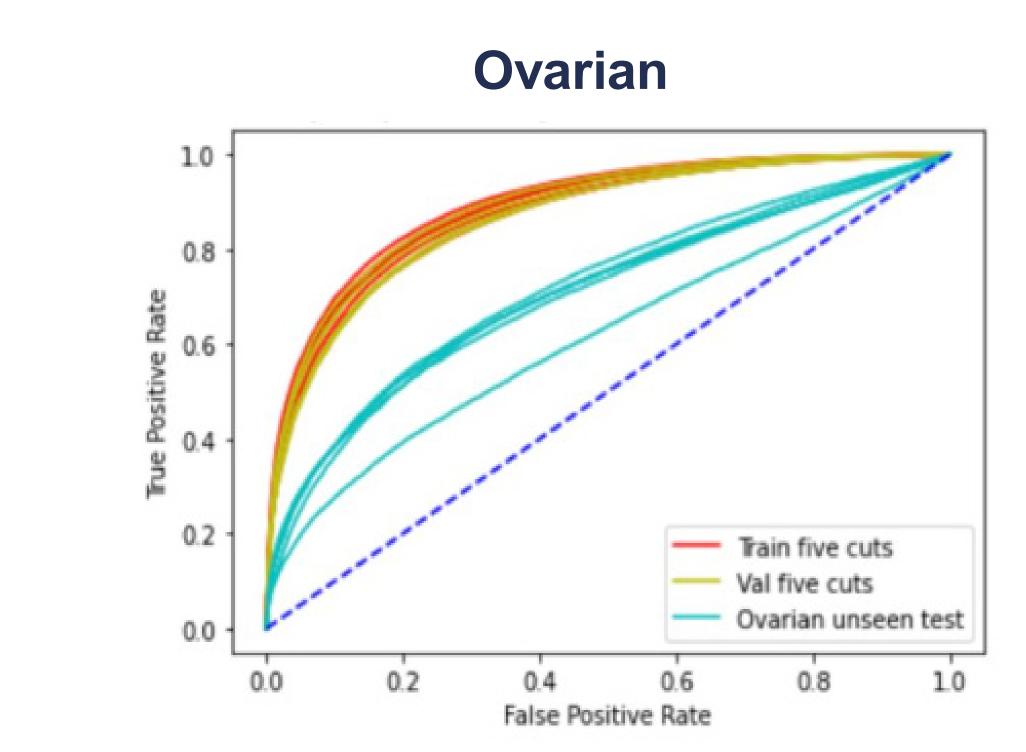
P-Value	<0.0001							
95% Confidence Interval	0.9552 to 0.9871							
Tumor Specific Bio	marker Correlation to Annotation							
Spearman R	0.8444							
P-Value	<0.0001							
95% Confidence Interval	0.7246 to 0.9147							
Total CD3 cells	100- 80- 80- 80- 80- 40- 20- 150000 150000 150							

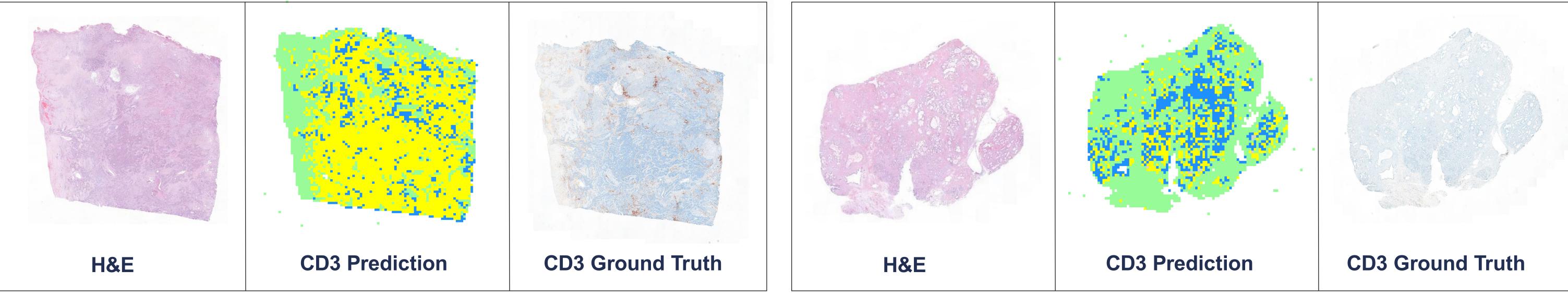
CD3 IHC Ovarian Tumor Correlation to Manual Counts

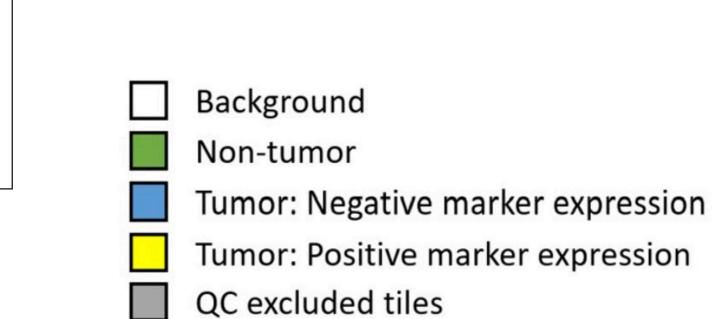
PREDICTIVE MODEL RESULTS

CD3 - Pan Model Results:

	Validation							Test Set	Other Indications AUC				
CD3	MCC	Balanced Accuracy	Sensitivity	Specificity	AUC	MCC	Balanced Accuracy	Sensitivity	Specificity	AUC	Lung	Cervical	Breast
Mean	0.578	0.789	0.821	0.758	0.875	0.285	0.642	0.613	0.669	0.689	0.790	0.677	0.707
Max	0.603	0.801	0.864	0.822	0.890	0.329	0.663	0.716	0.730	0.722	0.806	0.695	0.721
Min	.553	0.774	0.774	0.693	0.862	0.172	0.585	0.520	0.590	0.615	0.753	0.664	0.693

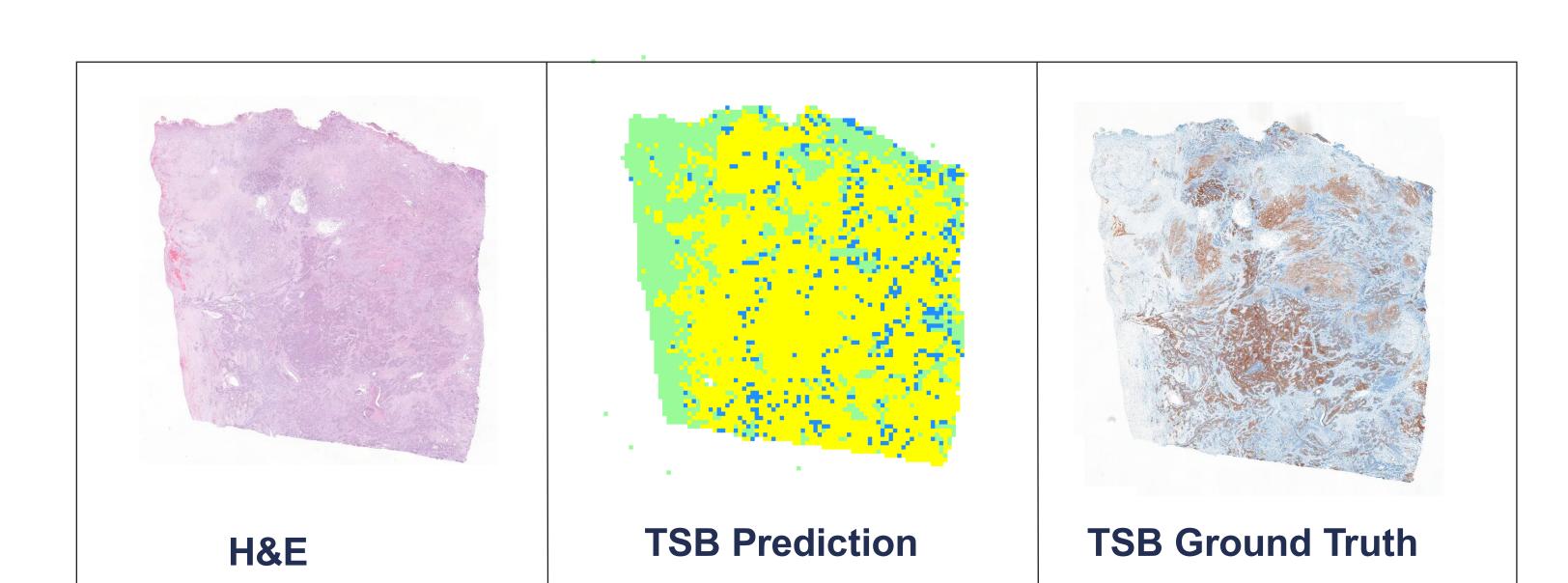


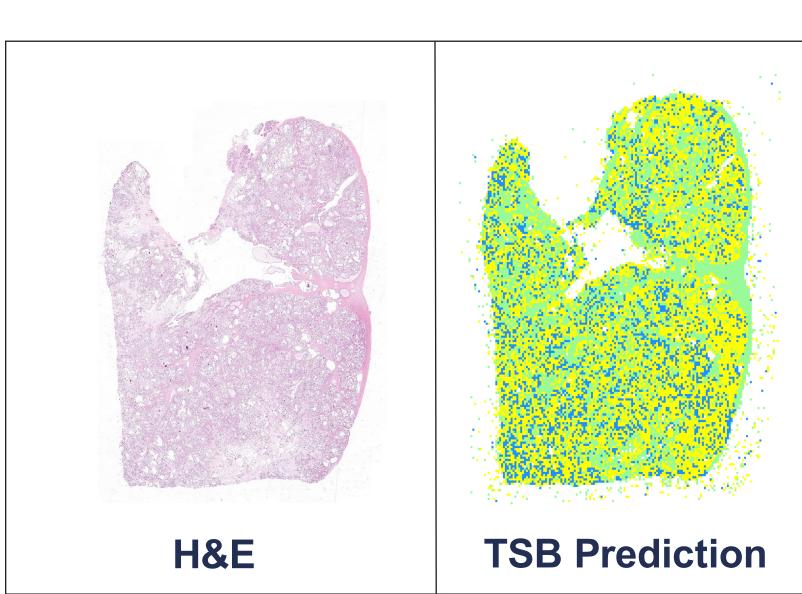


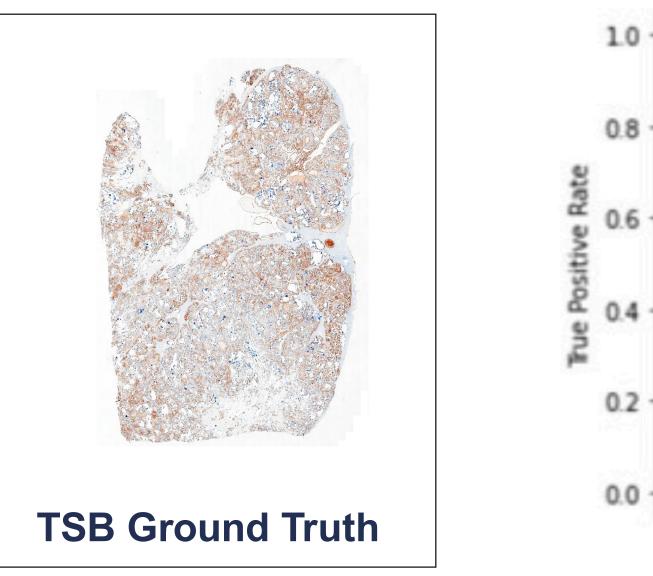


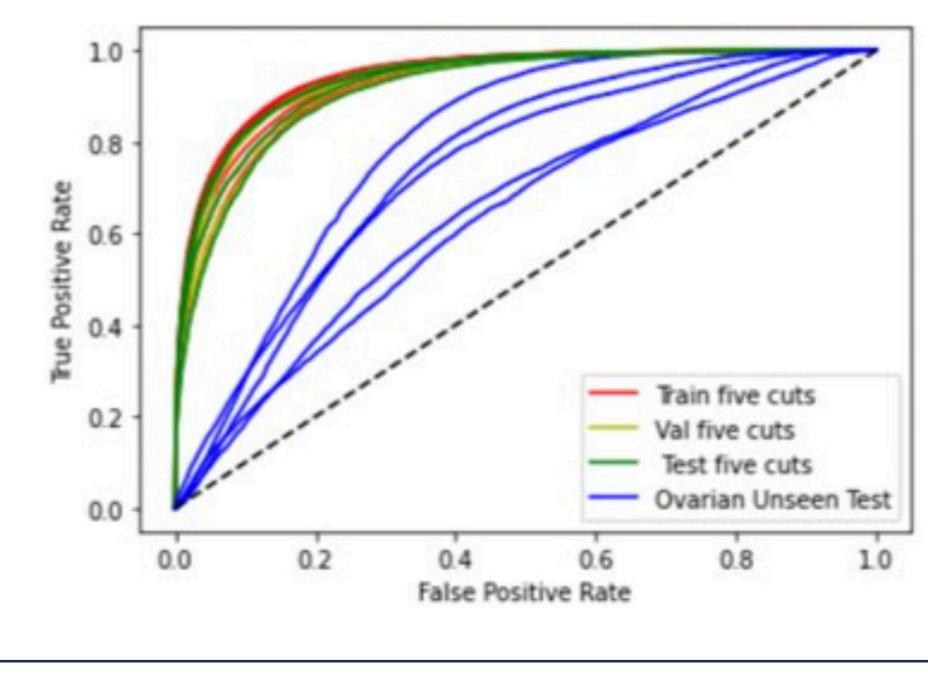
Tumor Specific Biomarker (TSB) - Model Results:

	Validation						Test Set					Other Indications AUC		
TSB	MCC	Balanced Accuracy	Sensitivity	Specificity	AUC	MCC	Balanced Accuracy	Sensitivity	Specificity	AUC	AUC	Lung	Cervical	Breast
Mean	0.682	0.846	0.922	0.770	0.933	0.685	0.847	0.923	0.772	0.934	0.714	0.866	0.878	0.865
Max	0.704	0.863	0.950	0.808	0.942	0.707	0.866	0.952	0.818	0.945	0.795	0.881	0.889	0.893
Min	0.647	0.824	0.903	0.697	0.918	0.643	0.826	0.904	0.701	0.914	0.634	0.846	0.869	0.840









Ovarian

CONCLUSION

These findings suggest that deep learning can be used as a complementary method to prescreen H&E-stained images to enhance the detection rate of the tumor specific biomarker and CD3 positivity in patient tumors. Additional work will involve the optimization of these models for implementation in a clinical setting.

ACKNOWLEDGEMENTS

Genmab:Suzana Couto

BioAl Health:

Robert Klopfleisch, Anand Datta, Armen Mkhitaryan, Thomas Colarusso, Michael Vergato