Detection of Pancreatic Ductal Adenocarcinoma Basal-Like and Classical Subtypes from H&E Whole Slide Images

Rohan Joshi[1], Stephane Wenric[1] , Irvin Ho[1], Aicha BenTaieb[1], and Martin C Stumpe[1]

[1]Tempus Labs, Inc., Chicago, IL // **Disclosure:** All authors are employees of Tempus Labs, a for-profit company // **Correspondence:** rohan.joshi@tempus.com

TEMPUS

INTRODUCTION

FOLFIRINOX and gemcitabine/abraxane chemotherapy regimens are first-line treatments for pancreatic ductal adenocarcinoma (PDAC). While FOLFIRINOX generally has superior efficacy, it is associated with severe side effects that often make treatment intolerable. Moffitt molecular subtypes of PDAC using RNA expression data have identified a basal-like subtype that is less responsive to FOLFIRINOX than a classical subtype. Patients with a basal-like subtype therefore could be better candidates for gemcitabine/abraxane. Here, we developed a proof-of-concept predictor of basal-like subtype from H&E whole slide images (WSIs) that could be used to rapidly identify and prioritize cases for further RNA profiling.

METHODS

WSIs and whole transcriptome RNA-seq data were collected from pancreatic cancer specimens (N=3,331, Table 1). Basal-like or classical labels were assigned by applying the Purity Independent Subtyping of Tumors (PurIST) algorithm to RNA-seq input data. An attention-based convolutional neural network was trained to predict the subtype label from each WSI in the training set (60% of data). Hyperparameters were selected using the optimization set (20%) of data. Performance is reported on a holdout set (20% of data). The entire process of training, optimization, and evaluation was repeated using 5-fold cross-validation.

Characteristic	Value	Classical (n)	Basal-like (n)	p-value*
WSI location	In pancreas	1160	277	<0.01
	Out of pancreas	1376	508	
Specimen type	Small specimen	1458	532	<0.01
	Large specimen	1051	246	
Scanner type	Leica GT450	1308	376	0.08
	Philips UFS	1236	411	

Table 1. Cohort characteristics used for development. *p-value for association of subgroups with PurIST label as determined by Fisher Exact Test.

SUMMARY

- Molecular **subtypes of PDAC** are associated with features that can be captured from **H&E whole-slide images** using a deep learning model.
- Identification of patients likely to have basal-like PDAC could be used to **rapidly identify** patients less likely to benefit from FOLFIRINOX therapy and **prompt follow-up RNA expression testing** for subtype confirmation.

RESULTS

Illustration of deep learning pipeline for predicting PDAC subtype from WSIs

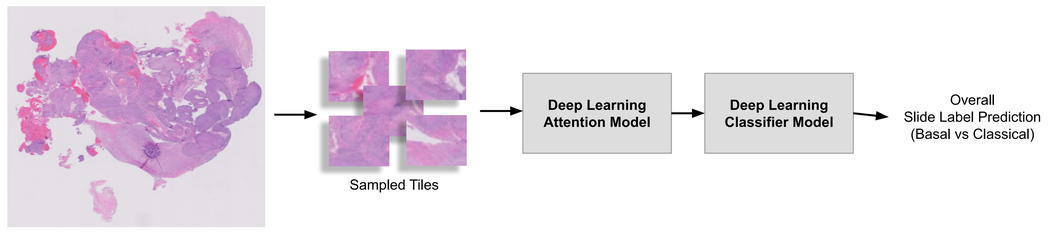

Figure 1. Whole slide images (WSIs) are broken up into tiles. Tile data are passed to deep learning modules to create a prediction. The weights of the deep learning models are iteratively updated until the area under the receiver operating characteristic curve (ROC-AUC) for the validation set no longer improves.

Trained model assigns high relevance to histologies associated with classical and basal subtypes

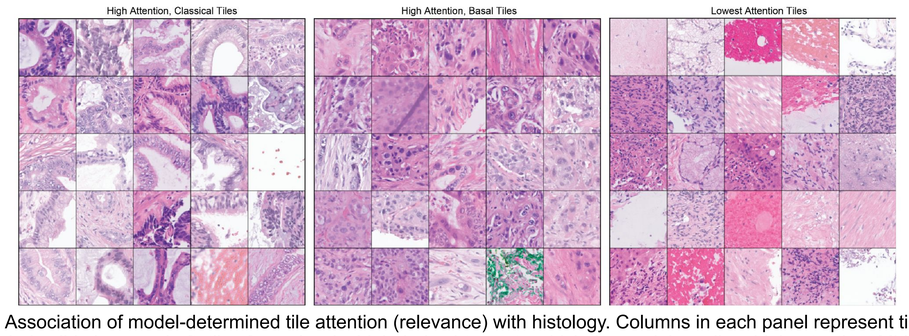

Figure 2. Association of model-determined tile attention (relevance) with histology. Columns in each panel represent tiles from different whole slide images. Left: tiles from classical-labeled subtypes to which the model ascribes high attention. Middle: tiles from basal-labeled subtypes to which the model ascribes high attention. Right: tiles where the model ascribes low attention (irrelevance) for the purposes of classical versus basal subtype classification.

Model performance is similar across magnification levels, slide location, specimen type, and scanner type

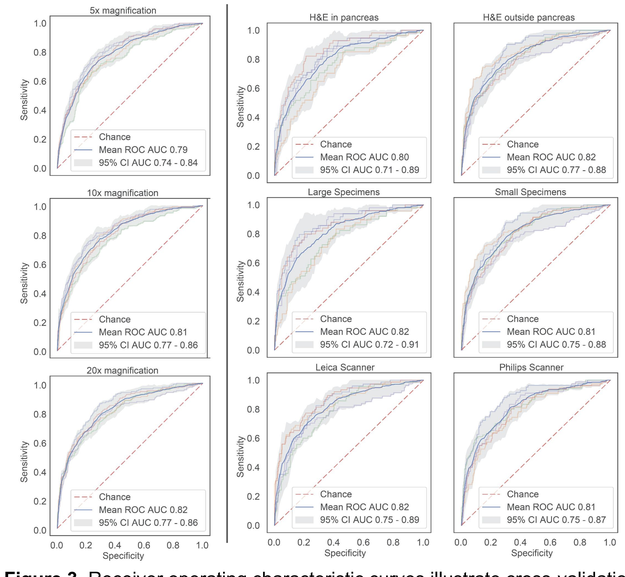

Figure 3. Receiver operating characteristic curves illustrate cross-validation model performance on held-out set of data. Left: performance of models trained and tested using tiles from 5x, 10x or 20x magnification. Right: performance of a model trained and tested on 20x tiles stratified by H&E location, size of specimen, and type of digitization scanner used.

Acknowledgements: We thank Josh Och, Sun Hae Hong and Yoni Muller for useful discussions and follow up, and the Tempus Scientific Communications team for visualization and poster review.