Predicting Prostate Cancer Molecular Subtype With Artificial Intelligence

Eric Erak[1] ,Uttara Joshi[2], Chaith Kondragunta[2], Nitin Singhal[2] , Saikiran Bonthu[2], Nilanjan Chattopadhyay[2], Lia DePaula Oliveira[1], Adrianna A Mendes[1], Onur Ertunc[1], Tracy Jones[1], Jessica L Hicks[1], Stephanie Glavaris[1], Gunes Guner[4], Igor Damasceno[5], Ibrahim Kulac [6], Javier A. Baena-Del Valle[3] ,Misop Han[1] ,Mark Markowski[1], Bruce J Trock[1], Angelo M De Marzo[1],Tamara L. Lotan[1]

[1] Department of Pathology, Johns Hopkins School of Medicine, Baltimore, MD; [2] AIRA Matrix, India; [3] Fundación Santa Fe de Bogotá University Hospital; [4] Hacettepe University; [5] UAB Hospital ; Birmingham, AL

Background

- Unlike other genitourinary cancers, such as renal cell carcinoma, visual microscopic examination of prostate cancer has failed to reveal a reproducible association between tumor molecular subtype and morphologic features. However, deep learning-based algorithms trained on whole slide images (WSI) from large cohorts with known molecular classification may outperform the human eye and provide a cost-effective and rapid method to identify cases with clinically relevant genomic alterations. As proof of principle, we described two such algorithms to identify prostate tumors with underlying *ERG* fusions and/or *PTEN* deletion.

Methodology

- We created a Concatenated Feature Based Classification (CFBC) (**Figure 1**) system based on deep learning that extracts non-linear characteristics from histopathology imaging data.
- The system included four stages: A) tumor identification in histopathology images; B) patch generation; C) pre-training; D) classification.
- Pre-training used self-supervised learning via the Barlow Twin SSL ResNeXt50 architecture (**Figure 1**).
- For training, we utilized a single representative H&E WSI of the dominant nodule from each of 224 tumors from a previously published Johns Hopkins prostatectomy cohort (Ref 1) where ERG and PTEN status were determined by genetically validated IHC assays (Ref 3 and 4).
- The classification task employed supervised learning using ResNet18 architecture and was performed on multiple validation cohorts.
- For both biomarkers, the algorithm performance was assessed using 64 additional WSI held out from the pre-training cohort as well as additional cohort of radical prostatectomy samples and/or needle biopsies with known PTEN/ERG status from IHC.

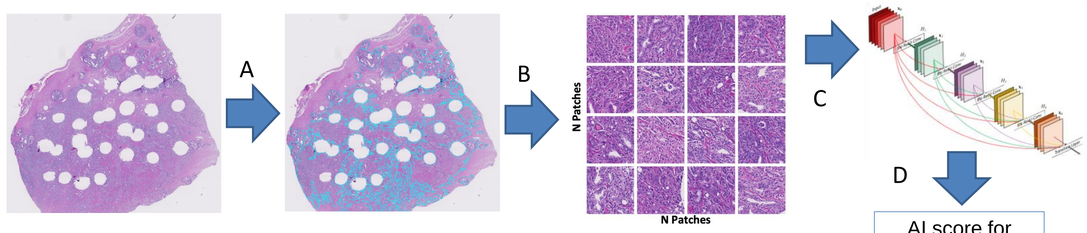

Figure 1. Schematic of deep learning algorithm to predict molecular status using H&E stained WSI. The Concatenated Feature Based Classification system included four stages: A) tumor identification in histopathology images; B) patch generation; C) pre-training; D) classification.

A

Training Dataset	# of patients
NHC (RP), Ref 1)	224 (104)

B Five-fold cross validation for ERG algorithm

Dataset	# of patients	AUC
NHC (RP, Ref 1)	64 (26)	0.91 ± 0.034
Case Cohort (RP, Ref 2)	248 (107)	0.86 ± 0.025
Race cohort (RP, Ref 3)	375 (127)	0.89 ± 0.007
Radiation Therapy Cohort (NB)	179 (70)	0.78 ± 0.017
GG1 cohort (NB)	158 (65)	0.79±0.027

Table 1: ERG algorithm training and performance. Table 1A represents the ERG training dataset obtained from the NHC cohort; 104/224 cases were ERG positive. **Table 1B** shows ERG algorithm performance using area under the receiver operator characteristic curve (AUC) for the ERG algorithm within the validation cohorts, both radical prostatectomy (RP) and needle biopsy (NB).

Five-fold cross validation for PTEN algorithm

Dataset	# of patients	AUC
Training: NHC (RP, Ref 1)	205 (56)	NA
NHC (RP, Ref 1)	50 (13)	0.81±0.043
Case cohort (RP, Ref 2)	201 (58)	0.72±0.014
Race cohort (RP, Ref 3)	337 (56)	0.80±0.015
Radiation Therapy Cohort (NB)	151 (28)	0.75±0.028

Table 2. PTEN algorithm training and performance. PTEN algorithm training and validation was performed excluding cases with heterogeneous (subclonal) PTEN loss by IHC. Performance is shown for three radical prostatectomy (RP) cohorts and two needle biopsy (NB) cohorts. In the NB cohort, the PTEN heterogeneous cases are not excluded, likely leading to the lower AUC.

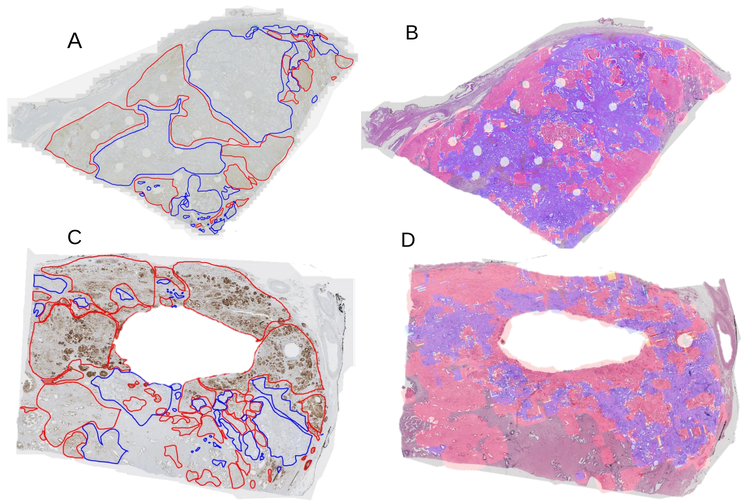

Figure 2. PTEN algorithm performance on excluded cases from training and validation cohort with heterogeneous (subclonal) PTEN loss by IHC. A and C. PTEN IHC staining with manual annotations for PTEN loss (blue) or PTEN intact (red) tumor areas, performed by a pathologist on representative WSI. **B and D.** Heatmap annotations performed by a deep learning model using a nearby H&E-stained level from the same block, with areas predicted to have PTEN loss (blue) or intact PTEN (red).

Results and Discussion

- ERG algorithm performance was assessed using three radical prostatectomy (RP) cohorts (**Table 1**), including WSI held out from the pre-training cohort (**AUC: 0.91**) and WSI from two independent RP cohorts (**AUC: 0.86**, **AUC: 0.89**) In addition, we tested ERG algorithm performance in WSI from two needle biopsy (NB) cohorts (**AUC: 0.78** and **AUC: 0.79**).
- PTEN algorithm performance (excluding cases with heterogeneous loss) was assessed using additional WSI held out from the pre-training cohort, as well as two additional RP cohorts (**AUC: 0.72** and **AUC: 0.80**) and one NB cohort (**AUC: 0.75**) (**Table 2**).

Conclusions

- Deep learning algorithms to predict ERG or PTEN status on H&E stained WSI images from prostate carcinoma showed excellent performance across independent validation cohorts, including radical prostatectomy and needle biopsy tissues. Additional validation in independent cohorts is ongoing (**Figure 2**).
- For *PTEN*, additional training is needed in order to account for frequent heterogeneous and subclonal loss.

References

1. Ross AE, Johnson MH, Yousefi K*, et al.* Tissue-based Genomics Augments Post-prostatectomy Risk Stratification in a Natural History Cohort of Intermediate- and High-Risk Men. *Eur Urol* 2016; **69**: 157-165.
2. Trock BJ, Jing Y, Mabey B*, et al.* Cell Cycle Progression Score, but Not Phosphatase and Tensin Homolog Loss, Is an Independent Prognostic Factor for Metastasis in Intermediate- and High-risk Prostate Cancer in Men Treated With and Without Salvage Radiotherapy. *J Urol* 2022; **208**: 1182-1193.
3. Kaur HB, Guedes LB, Lu J*, et al.* Association of tumor-infiltrating T-cell density with molecular subtype, racial ancestry and clinical outcomes in prostate cancer. *Mod Pathol* 2018; **31**: 1539-1552.
4. Guedes LB, Tosoian JJ, Hicks J*, et al.* PTEN Loss in Gleason Score 3 + 4 = 7 Prostate Biopsies is Associated with Non-organ Confined Disease at Radical Prostatectomy. *J Urol* 2017; **197**: 1054-1059.