Predicting Metastatic Risk at Diagnosis from Digital Pathology Slides in the VA Health Care System

Man Minh Ho[1], Ben J. Brintz[2], Yosep Chong[3], Arkadiusz Gertych[4], Candace Haraldson[5], Ilsa Garraway[6,7], Beatrice Knudsen[8,9]

Affiliations:
1. Institute of Scientific Computing and Imaging, University of Utah, Salt Lake City, UT
2. Department of Internal Medicine, Division of Epidemiology, University of Utah, SLC, UT
3. The Catholic University of Seoul, St. Mary's Hospital, South Korea
4. Cedars Sinai Medical Center, Los Angeles, CA
5. Salt Lake City Veteran Administration Health Care System, Salt Lake City, UT
6. Department of Urology, UCLA, Los Angeles, CA
7. Greater Los Angeles VA Medical Center, Los Angeles, CA
8. Department of Pathology, University of Utah, Salt Lake City, UT
9. Huntsman Cancer Institute, Salt Lake City, Utah

Background. Men diagnosed with high-grade prostate cancer harbor increased rates of occult metastatic disease. The clinical stage of prostate cancer represents an important factor for treatment decisions at the time of prostate cancer diagnosis. PSMA imaging improves the early diagnosis of metastatic cancer, but it is expensive and may not be readily available. Genomic assays predict patient outcomes, but are not suited to predict the clinical stage at diagnosis. We propose that features of the cancer in H&E stained tissue sections from prostate needle biopsies can be used to predict metastatic disease at diagnosis.

Methods. We analyzed 168 high-risk cases from the Greater Los Angeles VA (GLA VA), equally divided into non-metastatic with at least five years of follow-up after diagnosis (M0) or synchronous metastatic cases (M1), and 38 cases with metachronous metastases (M0-P). From each case, we analyzed all available images of H&E biopsies. Tiles encompassing the whole biopsy core were first passed through a pre-trained feature extractor convolutional neural network, Transformer, or Foundation model. The output of the feature extractor served as the input to a fully connected network or a multiple instance learning (MIL) aggregator model that was trained de-novo to distinguish M0 from M1 cases and outputs a metastasis risk (MR) score. The features highlighted by the models were visualized in the digital slide. The MR scores were combined with race, age and PSA clinical variables as a covariate in a linear regression or scores were combined using a super learner approach to examine the contribution of each variable for the M1 prediction. Missing clinical data were imputed. Generalized performances of the models were assessed using 5-fold nested cross-validation.

Results. The two feature extractor models, GigaPath and PathDINO, performed the best for feature extraction together with a Vision Transformer (vanilla ViT) to classify M0 versus M1. The average AUC achieved by this combination of models was 0.89. The addition of PSA and race further increased the AUC to 0.91. The embeddings generated by the feature extractor model were backpropagated to the H&E image and explained by a pathologist. Features learned by the aggregator model were further examined by analyzing the model's attention on epithelium and cancer regions.

Conclusions. We demonstrate that digital pathology slides from diagnostic prostate needle biopsies can provide information for prostate cancer staging at diagnosis. Further, using state-of-the-art foundation models as feature extractors improves the accuracy of the stage predictions, particularly for small cohort sizes. The explanation of features that the model uses for its predictions increases the confidence in the results generated by AI models. Together, this data demonstrates that H&E images have the potential to improve the accuracy of staging of prostate cancer patients in a point-of-care setting.

Funding Acknowledgements: CDMRP Prostate Cancer Awards W81XWH-21-1-0725 and HT94252410186, VA Merit Award I01CS002622, Department of Pathology at University of Utah

Conflicts of Interest Disclosure Statement. The authors have no conflicts of interest related to this research.