

Deep learning–based histologic classifiers enable molecular subtyping of metastatic prostate cancer

Zhijun Chen, Erolcan Sayar, Daniela Guevara, Helen Richards, Haoyue Zhang, Radhika A. Patel, Agnes C. Gawne, Lucas J. Liu, Ilsa Coleman, Ruth Dumpit, Colm Morrissey, Michael T. Schweizer, Ruben Raychaudhuri, Laura S. Graham, Evan Y. Yu, Heather H. Cheng, Chien-Kuang C. Ding, Yuzhuo Wang, Peter Choyke, Baris Turkbey, Chantal Chaneel-Vos, Christina Fedorov, John R. Otilano III, Troy Kane, Jyothi Manohar, Michael Sigouros, Jones T. Nauseef, Ana Molina, David Nanus, Scott T. Tagawa, Juan Miguel Mosquera, Himisha P. Beltran, Ruth Etzioni, Peter S. Nelson, Rama Soundararajan, Ana M. Aparicio, Cora N. Sternberg, Michael C. Haffner, Stephanie A. Harmon

JCI Insight. 2026. <https://doi.org/10.1172/jci.insight.201872>.

Research In-Press Preview Clinical Research Oncology

Metastatic prostate cancer is a clinically and molecularly heterogeneous disease. Under the selective pressure of androgen receptor (AR)–directed therapies, resistant phenotypes frequently emerge, posing significant diagnostic and therapeutic challenges. Neuroendocrine prostate cancer (NEPC) is a clinically important phenotype characterized by lineage plasticity, neuroendocrine features, visceral metastases and poor prognosis. Accurately diagnosing NEPC remains difficult due to its histologic and molecular complexity but has high clinical relevance. In this study, we developed a deep learning model that leverages interpretable cellular features to improve feature extraction from H&E-stained tissue sections (NEURAL-PC). By incorporating a multiple instance learning (MIL) framework, NEURAL-PC enables robust NEPC classification solely from H&E tumor images, achieving an area under the receiver operating characteristic curve (AUROC) of 0.921 in independent external validation. In addition to its diagnostic utility, NEURAL-PC provides prognostic information that enables further subclassification of advanced prostate cancer across diverse datasets supporting its strong prognostic value and generalizability. Broadly, our work highlights a hybrid approach that integrates features across different domains, offering a promising strategy for developing reliable deep learning tools in pathology. Built on this framework, NEURAL-PC represents an extensively validated diagnostic and prognostic model for advanced prostate cancer.

Find the latest version:

<https://jci.me/201872/pdf>



Deep learning–based histologic classifiers enable molecular subtyping of metastatic prostate cancer

Zhijun Chen¹, Erolcan Sayar², Daniela Guevara³, Helen Richards², Haoyue Zhang¹, Radhika Patel², Agnes Gawne², Lucas Liu², Ilsa Coleman², Ruth Dumpit², Colm Morrissey⁴, Michael T. Schweizer^{2,4}, Ruben Raychaudhuri⁴, Laura Graham⁴, Evan Y. Yu^{2,4}, Heather H. Cheng^{2,4}, Chien-Kuang C. Ding⁵, Yuzhuo Wang⁶, Peter Choyke¹, Baris Turkbey¹, Chantal Chanel-Vos³, Christina Fedorov³, John Otilano³, Troy Kane³, Jyothi Manohar³, Michael Sigouros³, Jones T. Nauseef³, Ana Molina³, David Nanus³, Scott Tagawa³, Juan Miguel Mosquera³, Himisha P. Beltran³, Ruth Etzioni², Peter S. Nelson^{2,4}, Rama Soundararajan⁷, Ana Aparicio⁷, Cora N. Sternberg^{3#}, Michael C. Haffner^{2,4#}, Stephanie A. Harmon^{1#}

¹National Cancer Institute, NIH, Bethesda, MD, USA

²Fred Hutchinson Cancer Center, Seattle, WA, USA

³Englander Institute for Precision Medicine and Weill Cornell Medicine, New York, NY, USA

⁴University of Washington, Seattle, WA, USA

⁵University of California San Francisco, San Francisco, CA, USA

⁶University of British Columbia, Vancouver, Canada

⁷The University of Texas MD Anderson Cancer Center, Houston, TX, United States.

[#]Authors contributed equally

CORRESPONDING AUTHORS

Michael C Haffner, MD PhD

mhaffner@fredhutch.org

206-667-6769

Fred Hutchinson Cancer Center, 1100 Fairview Ave. N., E2-112, Seattle, WA 98109

Stephanie Harmon, PhD

stephanie.harmon@nih.gov

240-858-3067

9000 Rockville Pike, NIH, Building 41, Room A101-I, Bethesda, MD 20892

Cora N. Sternberg, MD

cns9006@med.cornell.edu

646-962-2072

Englander Institute for Precision Medicine, Weill Cornell Medicine, Belfer Research Bldg., 413 East 69th Street, Rm. 1412, New York, NY 10021

CONFLICT OF INTERESTS

C.C.D. received research funding support from Intuitive surgical. C.N.S. has served as consultant/advisory board member for Astellas Pharma, AstraZeneca, Bayer, Bristol-Myers Squibb/Medarex, Gilead, Merck, MSD, Pfizer, Janssen, Roche, UroToday, OncoLive, Duality Biologics and Tolmar. H.H.C. received Research Funds to Institution from Clovis Oncology, Color Genomics, Janssen, Medivation, Promontory Pharmaceuticals, Sanofi. M.T.S. received research funding support from Novartis, Zenith Epigenetics, Eli Lilly, BMS, Merck, Immunomedics, Tmunity, SignalOne Bio, Epigenetix, Xencor, Incyte, Ambrx, Oric Pharmaceuticals, AstraZeneca, Janssen, Pfizer. Serves on advisory board for Sanofi, Fibrogen, Daiichi Sankyo, Janssen, Pfizer. E.Y.Y. received research funding support (institution) from Dendreon, Merck, SeaGen/Pfizer, Blue Easter, Bayer, Lantheus, Tyra. Consulting for Astellas, Johnson & Johnson, AstraZeneca, Tolmar, Merck, Bayer, Lantheus, Bristol-Myers Squibb, Loxo. M.C.H. served as a paid consultant/received honoraria from Pfizer and Astra Zeneca and has received research funding from Merck, Novartis, Genentech, Promicell and Bristol Myers Squibb.

ABSTRACT

Metastatic prostate cancer is a clinically and molecularly heterogeneous disease. Under the selective pressure of androgen receptor (AR)-directed therapies, resistant phenotypes frequently emerge, posing significant diagnostic and therapeutic challenges. Neuroendocrine prostate cancer (NEPC) is a clinically important phenotype characterized by lineage plasticity, neuroendocrine features, visceral metastases and poor prognosis. Accurately diagnosing NEPC remains difficult due to its histologic and molecular complexity but has high clinical relevance. In this study, we developed a deep learning model that leverages interpretable cellular features to improve feature extraction from H&E-stained tissue sections (**NEURAL-PC**). By incorporating a multiple instance learning (MIL) framework, NEURAL-PC enables robust NEPC classification solely from H&E tumor images, achieving an area under the receiver operating characteristic curve (AUROC) of 0.921 in independent external validation. In addition to its diagnostic utility, NEURAL-PC provides prognostic information that enables further subclassification of advanced prostate cancer across diverse datasets supporting its strong prognostic value and generalizability. Broadly, our work highlights a hybrid approach that integrates features across different domains, offering a promising strategy for developing reliable deep learning tools in pathology. Built on this framework, NEURAL-PC represents an extensively validated diagnostic and prognostic model for advanced prostate cancer.

INTRODUCTION

Prostate cancer (PC) is the most common non-cutaneous malignancy in men in the United States (1). While localized PC is often curable, metastatic disease is associated with high mortality and remains challenging to treat (2). Importantly, metastatic PC is highly heterogeneous, and over the past decade, several clinically relevant phenotypes have been identified (3–13). However, accurately diagnosing these phenotypes remains difficult and represent a clinical challenge (4). Among the phenotypes encountered in patients with advanced disease, high-grade neuroendocrine prostate cancers (NEPC) is particularly notable. NEPC exhibits morphologic features overlapping with small cell carcinomas, though its histologic appearance is variable and difficult to interpret (14, 15).

Broadly, NEPCs with or without small-cell morphology, are increasingly observed in patients who progress on therapies targeting the androgen receptor (AR) signaling axis (4, 10, 16, 17). These treatment-emergent NEPCs are typically associated with aggressive clinical behavior, poor response to further AR-directed therapy, and a dismal prognosis (12, 15, 16, 18–20). Transcriptomic profiling distinguishes these tumors from conventional AR-driven prostate adenocarcinomas (12, 13, 21–24). They frequently express high levels of stem-cell and progenitor as well as neural and neuroendocrine markers, reflecting their profound lineage plasticity.

Despite these defining molecular features, accurate diagnosis of NEPC remains challenging—even for expert pathologists—owing to the marked histomorphologic diversity of these tumors. Notably, a subset of PCs displays clinical and molecular similarities to NEPC but lacks overt histologic features of high-grade neuroendocrine carcinoma (4, 11, 14, 16, 18, 19, 25–27). Ancillary studies, such as immunohistochemical staining for neuroendocrine markers, are not universally available and are often applied in a non-standardized manner (4, 15), resulting in inconsistent detection and reporting of NEPC. Moreover, neuroendocrine marker positivity alone does not always correlate with the aggressive clinical behavior characteristic of true NEPC (3, 19). This can lead to diagnostic misclassification and can complicate clinical interpretation.

These limitations underscore the urgent need for a robust, standardized, and reproducible diagnostic assay to accurately identify NEPC and enable consistent detection of this aggressive PC variant. From a clinical standpoint, such a test would be of substantial value, as it would streamline the diagnostic process, allowing for accurate and timely treatment decisions in patients who require more intensive therapeutic intervention, involving platinum-based chemotherapy—often within a narrow window of treatment opportunity (16, 18, 25, 26).

With the recent development of imaging devices and machine learning algorithms, tasks like tumor detection, segmentation and classification in digital pathology can be addressed by various deep learning models (28–30). However, the training of models for such applications usually requires a large number of manually annotated images. Collecting high quality data and annotations are usually difficult since it requires well-trained pathologists and is time-consuming. Thus, developing unsupervised or semi-supervised methods has become popular (31–33). A convolutional autoencoder (CAE) is a type of neural network composed of an encoder that uses convolutional layers to extract low-dimensional image features, and a decoder that employs transposed convolutional layers to reconstruct the original image from these features (34). Variants of CAEs have shown promising performance across a range of medical imaging tasks (33, 35–38). However, many CAE-based models face challenges related to feature interpretability and are susceptible to overfitting (39). Prior work has attempted to address these limitations by introducing sparsity constraints (40, 41) or modifying the objective function to regularize the latent space (42). While such strategies can enhance model performance, it remains difficult to directly relate latent features to human-interpretable characteristics.

This study, we test the hypothesis that image-based machine learning methods can robustly detect NEPC. To this end, we develop a CAE-based model to detect NEPC by improved feature interpretability and robustness, termed **NEURAL-PC** (NEuroendocrine Recognition Using Attentive Learning in Prostate Cancer). NEURAL-PC integrates biologically meaningful cellular features into the learning process through a hybrid architecture. Specifically, we introduce a projection module that maps the encoder’s output

to a predefined latent space informed by cellular features. This projection acts as a form of regularization and enhances the quality of image feature extraction. A downstream classifier built on top of our pretrained model outperforms benchmark approaches in detecting NEPC. Validated within external cohorts, NEURAL-PC shows strong generalizability and relevance to clinical diagnostics and decision-making.

RESULTS

Initial development of NEURAL-PC

NEURAL-PC extracts features from standard H&E images using a convolutional autoencoder which is regularized, i.e. penalized during the learning process, by handcrafted cellular features (**Figure 1**; see **Methods for additional details**). The model was first pre-trained in an unsupervised manner from a diverse development cohort of metastatic PC specimens obtained from the University of Washington Rapid Autopsy (UW RA) cohort (43, 44). This cohort comprises of 722 tissue-microarray (TMA) cores taken from 204 metastatic sites of 52 patients, representative of the molecular and morphological heterogeneity (including sarcomatoid and squamous features) of advanced PC (4, 43, 44). All samples have undergone comprehensive histologic and molecular characterization including expert histopathology review for reference standards in model development (45–47) (see **Methods** for additional study cohort details and NEPC classification reference standard guidelines). 502 cores (450 for training and 52 for validation) from 31 patients were reserved for training, while the remaining cores from 21 patients were reserved as a testing set.

Cores were pre-processed in tiles of size 250x250 at 40x magnification through an automated pipeline to identify cancer areas (see **Figure 1A** and **Methods**). Cellpose (48) was used to segment nuclei and extract a cellular feature vector (49) for each tile (**Figure 1B**) which contains information such as nuclear size, shape, texture, and distribution.. The objective of the pre-training step was to reconstruct images and corresponding handcrafted cellular features, as shown in **Figure 1C**, with further architecture details in **Supplemental Figure 1**. The model inherently optimizes image features to incorporate interpretable handcrafted cellular features information. Following unsupervised training, the pre-trained weights of the

encoder and the projector were used to initialize an image feature extractor for downstream classification of NEPC (see **Figure 1D**; Methods). A multiple instance learning (MIL) approach was employed to predict labels for cores in the training set (NEPC positive: 149, NEPC negative: 256) using pathologist and IHC positivity (see below) as a reference standard(4). The hyperparameter settings and the training strategies for all steps are discussed in the Methods section. The final output of NEURAL-PC model is a single sample-level output score and spatial map probability distribution (heatmap) from all tiles used during MIL inference.

We benchmarked our NEURAL-PC model against other commonly used image feature extractors including ResNet50(50), Swin Transformer(51) pre-trained on ImageNet and UNI(52) foundation model. Additionally, we evaluate a standard CAE model (image only) and a separate model evaluating handcrafted cellular features to understand the relative benefit of NEURAL-PC model. All models were evaluated using multiple 5-fold cross validation schemes. The area under the receiver operating characteristic (AUROC) and the area under the precision-recall curve (AUPRC) were calculated for each model (**Table 1**). Interestingly, the use of handcrafted cellular features alone achieved high AUROC/AUPRCs, outperforming CAE alone, but fusing information during pre-training resulted in the highest performance for NE detection (NEURAL-PC AUPRC 0.717, AUROC 0.883). UNI foundational model closely followed NEURAL-PC with AUPRC 0.704 and AUROC 0.900. Given the imbalance in frequency of NEPC and Adenocarcinoma, AUPRC is the preferred metric. Other standard models pre-trained from ImageNet (ResNet50, Swin Transformer) had poor AUPRCs, perhaps attributable to domain mismatch and dataset imbalance. It is important to note that NEURAL-PC has a simple structure (see **Methods**) with much fewer parameters than other transformer-based models making it efficient to run. When compared to pathologist morphology assessment, NEURAL-PC showed median NEURAL-PC score in poorly differentiated carcinomas was 0.214 (0.01-0.91), compared with 0.068 (0.01-0.66) in Adenocarcinoma and 0.847 (0.15-0.97) in NEPC morphology (**Supplemental Figure 2a**). Based on the IHC ground-truth panel for NEPC classification, majority of false negatives were in non-NEPC morphologies (**Supplemental Figure 2b**).

Independent validation of NEURAL-PC across different cohorts and sample types

To assess the performance of NEURAL-PC, model prediction scores were first compared to semi-quantitative IHC staining for Androgen Receptor [AR] and two well-established and clinically used neuroendocrine markers (synaptophysin [SYP], Insulinoma-associated protein 1 [INSM1])(4) in the UW-RA TMA testing cohorts (**Figure 2A-C**). NEURAL-PC scores correlated well with IHC H-scores: high NEURAL-PC probabilities were observed for strong SYP (median: 0.832; IQR: 0.545–0.865) and INSM1 (median: 0.826; IQR: 0.504–0.889) staining (H-score ≥ 70), while intermediate H-scores (20–70) yielded more variable correlation with SYP (**Figure 2B**) and INSM1 (**Figure 2C**). NEURAL-PC spatial prediction maps revealed uniform probability output in cases with consistent AR+/NE– or AR–/NE+ phenotypes (**Figure 3A**). In mixed cases, detection was concentrated in spatially distinct regions, reflecting heterogeneous cell states and biomarker expression and demonstrating that the NEURAL-PC can differentiate intra-tumoral cell state differences.

As a proof-of-concept to evaluate how well the model is able to discern the transition from an AR-dominant phenotype to an NE-dominant phenotype, we applied NEURAL-PC to a well-established patient-derived xenograft (PDX) model, LTL331, which undergoes reproducible transition from an adenocarcinoma to an AR-independent NEPC following castration (53). Samples from three timepoints representing different stages in the lineage transition cascade were evaluated and compared to IHC as reference standard (**Figure 3B**). After H&E normalization for mouse tissue, NEURAL-PC prediction maps matched IHC findings across timepoints. A uniformly negative NEURAL-PC probability map is observed at baseline, with strong AR expression at baseline and negative NEPC expression. In the post-castration setting (week 46), NEURAL-PC shows uniformly positive probability map within the emergent NEPC tumor showing strong NE marker expression and negative AR expression. Importantly, NEURAL-PC correctly identified the emergence of focal NE marker expression during the transitional timepoint (week 24), with NEURAL-PC spatial probability correctly identifying focal regions of NEPC features. This shows that even in admixed tumors, NEURAL-PC can correctly identify focal NEPC.

To assess the utility of NEURAL-PC in clinical biopsy samples, we evaluated its performance in a retrospective cohort of 63 clinical biopsy samples (UW-MATCH cohort: 24 NEPC, 38 adenocarcinoma, 1 amphicrine). These samples were characterized following published guidelines for IHC confirmation to determine NEPC(4). NEPC cases had higher NE scores (median: 0.877; IQR: 0.749–0.949) than adenocarcinomas (median: 0.243; IQR: 0.091–0.523) (**Figure 2D**). In this cohort, NEURAL-PC achieved an AUROC of 0.921, which was even higher than the mean AUC reported during cross-validation training in TMA cores (**Figure 2E**). Using an optimal NE threshold of 0.598 (Youden’s J statistic), the model achieved a sensitivity of 0.875 and PPV of 0.724. Classification performance is summarized in **Table 2**, with representative prediction maps shown in **Figure 3C**, highlighting the morphological consistency in high-attention regions for NEPC sample. NEURAL-PC demonstrated uniformly low probabilities in confirmed adenocarcinoma samples (**Figure 3D**).

As an additional real-world cohort with limited IHC data, we analyzed 92 slides from an external cohort of 74 patients diagnosed with NEPC-like features, of which 85 slides from 68 patients had sufficient tumor content (WCMC cohort). Pathology reports were queried to identify all patients with small cell prostate cancer, NEPC, or any overlap of adenocarcinoma and NEPC features consistent with amphicrine carcinomas (3, 22, 54). A summary of pathology report diagnosis and NEURAL-PC classification can be found in **Table 3**. NEPC validation through IHC was not completed in this cohort. Within these NEPC-like features, NEURAL-PC scores were lower than IHC-proven cases (median: 0.637; IQR: 0.245–0.920; **Supplemental Figure 3**).

To evaluate quality-related impacts on NEURAL-PC performance, we next reviewed low-scoring cases with an expert genitourinary pathologist to further assess performance. In cases where the cancer detection results in a high number of false positive regions (normal surrounding tissue), these regions are largely predicted as NE negative by the NEURAL-PC algorithm and therefore result in underestimation of NE proportion and overall prediction in the image (**Supplemental 4A**). Another notable cause of failures in this cohort were common artifacts in whole slide images (WSI) that may also cause detection failure,

including scanning artifacts (blur) or tissue artifacts (processing or mounting effects), as shown in **Supplemental Figure 4B**. Finally, NE detection may fail when there are small/focal NE components in a large sample. The slide shown **Supplemental Figure 4C** demonstrates the majority of the slide is represented by typical adenocarcinoma-like features.

Correlation of NEURAL-PC predictions and tumor transcriptomic features

To determine the association between NEURAL-PC outputs and tumor transcriptomic features, we leveraged 118 WSI of large tissue sections from 75 patients for which H&E tissue images as well as bulk RNA-seq data was available (43, 44). Each H&E slide was reviewed by an expert pathologist for detailed morphological characterization and NEPC determination was made from a combination of morphology and molecular features. NEURAL-PC's outputs were well-correlated with pathologist-defined morphology group and NEURAL-PC achieved an AUROC of 0.810 (**Supplemental Figure 5**).

When evaluating the association between NE prediction scores and transcriptomic profiling of tumors, specimens predicted as NEPC-positive by NEURAL-PC showed high level expression of NEPC-related genes such as *SYP*, *INSM1* and *ASCL1* while specimens predicted as NEPC-negative demonstrated higher expression of canonical AR+ luminal epithelial cells seen in classic prostate adenocarcinoma such as *KLK3*, *KLK2*, *AR*, *STEAP1* and *NKX3.1* (**Figure 4A**). Furthermore, expression in 10 NE-related genes was significantly higher for samples predicted as NEPC positive by NEURAL-PC ($p=0.00135$, **Figure 4B**), showing a dominant association to positive expression in NEPC genes and negative expression in AR-related genes (**Figure 4C**). Correspondingly, there was significantly lower expression in 10 AR-related genes for samples classified NEPC negative by NEURAL-PC ($p=3.92 \times 10^{-8}$, **Figure 4D**). In addition, in line with the overall higher proliferation rates of NEPC, samples with high NEURAL-PC scores showed higher expression of CCP.31 (**Figure 4E**), a 31-gene cell cycle progression signature which is tightly correlated with cellular proliferation rates and associated with poor outcomes in PC(55).

NEURAL-PC predicts outcome in cohorts with aggressive variant PC

To assess the relationship between NEURAL-PC -derived predictions and clinical outcomes, we trained a random survival forest (RSF) prognostic model using all model-derived features: slide-level NE prediction, proportion of NE-predicted tiles, average tile-level probabilities and logits (see **Methods**) in the UW-MATCH cohort. Out-of-bag predictions from the RSF effectively stratified patients by risk of death within two years, with median overall survival of 14.3 months in the high-risk group versus >24 months in the low-risk group ($p<0.0001$; **Figure 5A**). This demonstrates that both the overall slide-level likelihood of NEPC and the extent of NEPC features detected by NEURAL-PC within the tumor are prognostic of outcome

To test this further, we applied the NEURAL-PC prognostic model first to an independent cohort of NEPC, amphotericin, or small cell prostate cancer (WCMC cohort). Among 66 patients with sufficient tumor burden and outcome data, patients predicted as high risk for death within two years had a median survival of 8.62 months vs 19.6 months in those predicted as lower risk (**Figure 5B**; $p=0.075$). While not reaching the statistical significance threshold, this trend suggests even within NEPC cohorts NEURAL-PC features can provide additional prognostic information.

To further validate the model, we evaluated NEURAL-PC in the C-COLA trial (56, 57), a phase II clinical study (NCT03263650) which assessed the activity of chemotherapy followed by PARP inhibitor maintenance in patients with aggressive variant prostate cancer (AVPC) (19). H&E scans from 82 patients with sufficient tumor burden were evaluated regardless of treatment arm by NEURAL-PC. Here, the NEURAL-PC prognostic model features demonstrated statistically significant association with 2-year outcomes, with the high risk group demonstrating median survival of 12.3 months vs. 20.4 months in patients classified as lower risk ($p=0.01$; **Figure 5C**). The association to outcome was stronger for the RSF model constructed from all NEURAL-PC features (HR 2.27 [95%CI 1.19-4.31; $p=0.01$]) compared to overall NEURAL-PC classification alone (HR 2.15 [95%CI 0.29-15.8; $p=0.45$]), demonstrating the prognostic value of incorporating NE prediction burden and heterogeneity across tiles.

DISCUSSION

Given the biological complexity of metastatic PC, accurate tumor detection and subclassification are of significant clinical importance. The overall objective of this study was to develop a deep learning-based model for accurate detection and quantification of NEPC from H&E slides. To this end, we have introduced an approach, NEURAL-PC, which utilizes so-called “handcrafted” features of nuclear size, shape, and appearance during training of the deep-learning based CAE to regularize and provide attention to important characteristics of the NEPC phenotype. This methodology outperforms common encoders used in the digital pathology domain and demonstrated significant association to NEPC-related gene expression signatures. NEURAL-PC was trained on diverse specimens reflecting the heterogeneity of metastatic PC and generalized well to multiple validation cohorts spanning PDX to clinical biopsy samples.

Accurate detection of NEPC is critical for effective patient management. However, establishing a definitive diagnosis of NEPC remains challenging in standard pathology practice. Although IHC markers can assist in diagnosis, isolated positivity for a single neuroendocrine marker, for example, can be misinterpreted as evidence of NEPC, leading to tumor misclassification and potentially inappropriate or excessive treatment(4). Conversely, failure to recognize NEPC—or delays in diagnosis—can prevent patients from receiving timely, aggressive therapy for this rapidly progressive disease.

These challenges highlight the urgent clinical need for robust diagnostic tools that can be applied broadly to all metastatic prostate cancer biopsies to evaluate the presence and extent of NEPC. Notably, our studies demonstrate that the NEURAL-PC assay can detect even small clusters of emerging NEPC, underscoring its potential utility for identifying even early or incipient NEPC lesions.

In addition to its diagnostic capability, we demonstrate that NEURAL-PC provides prognostic information, enabling the stratification of clinically aggressive variants of PC further into prognostic subgroups. By combining overall classification results with other tumor metrics including the proportion of cancer predicted as NEPC, and the spatial distribution of model outputs, NEURAL-PC can separate patients into prognostic groups. This enhanced prognostic value suggests that the heterogeneity in morphological

patterns contains clinically relevant information. Broadly, this observation aligns with prior reports highlighting the phenotypic heterogeneity of NEPC(20). For instance, tumors with “mixed” histologies have been shown to have more favorable outcomes compared to pure NEPC(3, 19).

While deep learning excels at capturing complex and hierarchical histologic patterns, such models are often criticized for their lack of interpretability. In contrast, handcrafted cellular features—based on established biological knowledge—have demonstrated prognostic value across multiple cancer types(49, 58–60). These features offer interpretability and can be engineered for robustness to technical variation, enhancing generalizability.

Prior studies have explored feature fusion strategies(61–63), typically treating deep and handcrafted features as independent inputs fused by concatenation or attention mechanisms. Our approach differs by integrating cellular features directly into the feature extraction process. Specifically, we regularize a CAE by projecting the encoder’s output into a fixed latent space for cellular features reconstruction. This hybrid strategy improves both image reconstruction and model interpretability. Notably, our model requires only a single feature extraction backbone at inference time, thereby eliminating the need for separate handcrafted feature computation while retaining the benefits of interpretable guidance during training. The same architecture holds potential for multi-modal fusion, transfer learning, and knowledge distillation in future applications.

NEURAL-PC was trained and validated across a wide range of clinically relevant specimens including samples from different metastatic sites, specimen types and biopsied obtained for a variety of clinical indications. Despite this diversity, we observed stable and reproducible performance, demonstrating the robustness of NEURAL-PC. To achieve this, we implemented a cancer detection module to pre-select tumor regions, reducing bias from varying metastatic sites and surrounding tissue microenvironments. This ensured sample type independent performance. Further, our user-in-the-loop workflow and interpretable outputs also allow for refinement of annotations by expert pathologists when necessary. Nonetheless, our end-to-end evaluation revealed that rare NEPC misclassifications stemmed from tumor detection errors and

tissue-related crush artifacts. Further evaluation of slide-related tissue artifacts and quality issues including image blur are an area for future improvement.

In summary, we developed a convolutional autoencoder regularized by handcrafted nuclear features NEURAL-PC to enable robust detection and characterization of NEPC on routine H&E-stained slides. By leveraging cellular morphology, the NEURAL-PC model outperformed multiple state-of-the-art classifiers, including foundation models. It demonstrated strong generalizability across diverse validation cohorts, including clinical biopsies. Given the molecular heterogeneity of late-stage prostate cancer and reliance on IHC for NEPC diagnosis, this H&E-based AI model offers substantial potential to enhance diagnostic precision and streamline pathology workflows.

Limitations of the study

This study has several limitations. While model development was performed on a large number of tissue samples, these ultimately came from a relatively small cohort of rapid autopsy patients which may not be fully representative of all disease phenotypes. However, careful validation in three settings (PDX, clinical biopsy, external clinical samples) was performed including an external site and digital scanners. Future prospective work and validation at additional clinical sites are warranted. Incomplete records or inter-reader variability may have an impact on results. Ground truth NEPC determination was made from a panel of IHC markers in the training set. Discrepancies between NEPC-related IHC markers and model performance was not assessed due to small incidence in the test set. Similarly, a small set of amphicrine tumors were available for rigorous. The impact of variable IHC expression and amphicrine-like tumors is a topic for future consideration. Our model was challenged against several state-of-the-art encoders for digital pathology; however, as foundation models continue to emerge future evaluation is warranted. In addition, future correlation of our NEURAL-PC model in addition to clinical features of prognostic relevance and its use as a predictive biomarker in various treatment settings is warranted.

STAR★METHODS

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

University of Washington Rapid Autopsy Cohort

All rapid autopsy tissues were collected from patients enrolled in the Prostate Cancer Donor Program at the University of Washington. Formalin-fixed, paraffin-embedded (FFPE) tissues from various metastatic sites within 52 patients were used to construct tissue microarrays, as described previously(43, 44). IHC of various molecular targets were available from prior published studies (45–47). For all IHC assays, a semi-quantitative measure of staining level was reported, H-score, with a range of 0–200 (47). Transcriptional profiling was completed for a sub-set of tumors within this cohort, as described previously(24) and available through (GSE147250, GSE228283). Corresponding H&E images from tumor sections undergoing transcriptional analysis were collected. All slides were digitized using a Ventana DP 200 Slide Scanner (Roche) at 0.25 $\mu\text{m}/\text{pixel}$ (effective 40x magnification) and deidentified images were evaluated as part of this study.

University of Washington Biopsy Cohort

A retrospective cohort of 63 patients with metastatic PC undergoing routine clinical metastatic biopsies at the University of Washington were identified after a review of pathological and clinical data. IHC studies, when performed as part of routine diagnostic workup, were included in the evaluation. Samples were evaluated for the presence of neuroendocrine-like features, based on morphological and/or molecular features by an expert genitourinary pathologist (4). All slides were digitized using a Ventana DP 200 Slide Scanner (Roche) at 0.25 $\mu\text{m}/\text{pixel}$ (effective 40x magnification) and deidentified images were evaluated as part of this study. Full clinical history including initial diagnosis, all prior treatments, molecular assays, and clinical follow-up were collected. Specifically, outcomes related to metastatic prostate cancer, including date and cause of death or last follow-up were collected.

Weill Cornell Medical Center Cohort

A retrospective cohort of 74 patients with metastatic PC who were clinically identified as having neuroendocrine-like differentiation and any specimens obtained throughout the course of care (biopsy) or death (autopsy) were reviewed. All slides were digitized using Aperio scanner at 0.25 $\mu\text{m}/\text{pixel}$ (effective 40x magnification) and deidentified images were evaluated as part of this study. Clinical history including diagnosis and treatment of prostate cancer, emergence of neuroendocrine phenotype, and date from biopsy to death or last follow-up were collected.

MD Anderson Cancer Center Cohort

A clinical trial cohort of 82 patients with metastatic PC enrolled in C-COLA trial (56, 57), a phase II clinical study which assessed the activity of chemotherapy followed by PARP inhibitor maintenance in patients with aggressive variant prostate cancer (AVPC) (19). All slides were digitized using Aperio scanner at 0.5 $\mu\text{m}/\text{pixel}$ (effective 20x magnification) and deidentified images were evaluated as part of this study. Image tiles were resampled to 40x magnification for analysis. Response to therapy and date from biopsy to death or last follow-up were collected per protocol.

LTL331 PDX model

Cryopreserved pieces of LTL 331 PDX(53) (generous gift from Dr. Yuzhuo Wang, PhD, FCAHS; UBC; Vancouver, British Columbia, Canada to Dr. Peter Nelson, MD; FH; Seattle, Washington, USA) were engrafted into the sub-renal capsules (SRC) of intact immunodeficient host mice [NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ; JAX#005557]. When tumors reached approximately 150 mm^3 – 200 mm^3 , as determined by MRI, tumors were excised and a portion was implanted subcutaneously (SQ) into the flanks of recipient mice and designated LTL331 SQ UBC. For this experiment, LTL331 SQ UBC tumors were implanted SQ into intact male mice, allowed to grow for approximately 2 months followed by castration and tumor resection at specific timepoints post-castration. Tumors were harvested at the assigned post-castration timepoint or when they reached the maximum allowable size of $\sim 2000\text{mm}^3$. Portions of the tumors were

fixed in 10% neutral buffered formaldehyde, transferred to 70% ethanol and sent to the Fred Hutch Experimental Histopathology Core for paraffin processing.

Formalin-fixed, paraffin processed tumors were embedded into composite paraffin blocks, sectioned at 4 μ m and mounted on positively charged slides. Sections were baked at 65°C for 1 hour to remove excess paraffin, then loaded onto the VENTANA Discovery Ultra automated staining platform (Ventana Medical Systems). Slides underwent onboard deparaffinization in Discovery Wash Buffer (Roche Diagnostics, 950-510), followed by heat-induced epitope retrieval (HIER) using Discovery CC1 solution (Roche Diagnostics, 950-224). Sections were stained with primary antibodies against AR (rabbit, D6F11, 1:100, Cell Signaling Technologies), INSM1 (mouse, A8, 1:100, Santa Cruz) or SYP (rabbit, SP11, 1:100, Genetex) diluted in Ventana Antibody Diluent with Casein (Roche Diagnostics, 760-219). For INSM1 staining, slides were further incubated with a rabbit anti-mouse IgG secondary antibody (1:400, Abcam) prior to detection. Detection was carried out using the DISCOVERY anti-rabbit HQ secondary antibody (Roche Diagnostics, 760-4815) and anti-HQ HRP enzyme conjugate (Roche Diagnostics, 760-4820), followed by visualization with ChromoMap DAB (Roche Diagnostics, 760-159). Sections were counterstained with Hematoxylin II (Roche Diagnostics, 790-2208) and Bluing Reagent (Roche Diagnostics, 760-2037), then scanned at 40x magnification using the Ventana DP 200 digital slide scanner.

METHOD DETAILS

Sex as a biological variable

All subjects in this research were biologically male, owing to the prostate as an organ and prostate cancer as an entity within only males.

Image Processing

All WSI undergo evaluation by a cancer detection algorithm to limit the influence of non-tumor associated pathology features, such as site of metastasis and surrounding microenvironment. Briefly, the cancer detection algorithm is trained from 165 WSI from metastatic biopsy (n=95), TMA core (n=58), and normal

tissue images (n=12; lymph node / adrenal controls). Initially, a subset of images were annotated by two pathologists (M.H., E.S) using QuPath software. A DeepLabV3 segmentation model was trained at 20x magnification with patch size 250x250. As visually depicted in **Figure 1(A)**, the model was iteratively applied to new WSI, converted to QuPath-friendly format(64) manually adjusted by pathologists, and re-trained from adjusted annotations until no major refinements were needed, using user-in-the-loop process defined previously(65). The final model achieved Dice 0.761 in a validation set. This model is then applied to all images in the study and was used to select tiles with more than 10% cancer area of size 250x250 at 40x magnification. Tiles with more than 45% white space were also discarded, resulting in 247,203 tiles (215,104 for training and 32,099 for validation). Cellpose(48) was used to segment nuclei and extract a cellular feature vector(49) for each tile. All cellular features are normalized using z-score transformation using mean and standard deviation calculated across all training TMA cores.

Stain normalization

For all experiments with external validation dataset, we apply the Macenko Stain Normalization(66) to ensure the consistence of the model inference. Specifically, the Macenko method estimates the stain vectors of a giving tile image using singular value decomposition. Then these stain vectors are normalized to some given references and these new stain vectors are used to construct the normalized images. The code we used in our experiments for stain normalization is available in https://github.com/zjchen10/stainnorm_pytorch.

NEURAL-PC Model architecture

The proposed NEURAL-PC model consists of an encoder f , a feature projection module p , an image decoder g_I , a cellular feature decoder g_C and a classifier module h . Briefly, each tile was first transformed by an encoder f and a projector p into two latent vectors z_m and z_c . The image docoder, g_I uses the sum $z = z_m + z_c$ to reconstruct the input tile image for regularization, which the feature decoder g_C uses only z_c to reconstruct cellular feature vector.

- **Encoder f .** Illustration of f is shown in **Supplemental Figure 1A**. It is a convolutional network which consists of several Residual Blocks, where the overall architecture of f is similar to ResNet34 (see **Supplemental Figure 1B**). To better preserve the spatial information in the image, we replace the pooling operation with stride 2 convolution. Let $X \in \mathbb{R}^{M^2 \times 3}$ be an input image of size M , the encoder f transforms it into a 3-d pre-feature map $I = f(X) \in \mathbb{R}^{N^2 \times q}$.
- **Feature projector p .** This module projects the pre-feature map I into a low dimensional latent space and split the projection into two parts. A linear layer followed by non-linear tanh activation first transforms the pre-feature map into a coefficient vector $y \in \mathbb{R}^n$, then the image feature z is obtained by matrix-vector product $z = Dy$, where $D \in \mathbb{R}^{l \times n}$ is a fixed and non-trainable dictionary matrix whose columns are i.i.d. sampled from uniform distribution over the unit sphere(40). This design allows us to decompose z into two “almost orthogonal” vectors. Specifically, if we view y as the concatenation of two vectors $y = [y_m^T, y_c^T]^T$ and D as the concatenation of two submatrices $D = [D_m, D_c]$, then

$$z = Dy = D_m y_m + D_c y_c = z_m + z_c$$

Since columns of D are i.i.d. random vectors, the inner product $|z_m^T z_c|$ is close to 0 with high probability. z_m is referred to as the main part which contains general visual information, and z_c is referred to as the complimentary part which will be used to reconstruct cellular features. The process is shown in **Supplemental Figure 1C**.

- **Image decoder g_I and cellular features decoder g_C .** These two decoder networks use feature from p to reconstruct images and corresponding cellular features. g_I is a deconvolutional network whose structure is symmetric to that of the encoder f , while g_C is a simple feed forward network. For image reconstruction, g_I uses the feature z . For cellular features reconstruction, g_C uses the complementary z_c , i.e., only a part of the image feature is used. Illustration of g_I is shown in **Figure 1C**.

- **Classifier h .** For downstream supervised classification tasks, we connect a classifier module h to the feature projector p (**Figure 1D**). Since we only have weak label for the cores/slides, we adopt the multiple instance learning (MIL) scheme. The module h contains an attention block(67) that congregates a collection of image features into a single output. Let $Z = [z_1, \dots, z_L]^T$ be a bag (or a collection) of all tile features from one core/slide image, the classifier h outputs the core-level feature as follows:

$$Y = \text{MultiHeadAttention}(Q, K, V)$$

Where $Q \in \mathbb{R}^{n_{class} \times d}$ contains trainable class tokens as its rows, $K = W^K Z$ and $V = W^V Z$. Note that this attention block can take any number of tile features. A feed forward neural network with sigmoid activation then takes this core/slide-level feature Y and outputs a probability/score. The architecture of h is shown in **Supplemental Figure 1D**.

The sizes of the pre-feature map I , coefficient vector y_m , coefficient vector y_c and the feature vector z were set to $16 \times 16 \times 16$, 1600, 300 and 1600 respectively in our experiments.

Loss functions

During pre-training phase, we train f , g_I and g_C for image and cellular features reconstruction. The following loss terms are used in this stage.

- **Image MSE loss.** It is defined as the mean-square difference between the input image X and the reconstructed image $\tilde{X}$,

$$\ell_{MSE} = \frac{1}{3M^2} \sum_{i=1}^M \sum_{j=1}^M \sum_{k=1}^3 (X_{ijk} - \tilde{X}_{ijk})^2$$

where X_{ijk} is the value of k -th channel of pixel at position (i, j) , M is the size of the image.

- **Image SSIM loss.** Optimizing ℓ_{MSE} alone will produce blurry reconstructed images and cause structural information loss. To better preserve this information, we use the Structural Similarity

Index Measure (SSIM) loss ℓ_{SSIM} . The SSIM between two images is defined as $SSIM(X, \tilde{X}) = l(X, \tilde{X}) \cdot c(X, \tilde{X}) \cdot s(X, \tilde{X})$, where l , c and s are luminance, contrast and structural comparison functions respectively. The SSIM loss is then defined as

$$\ell_{SSIM} = 1 - SSIM(X, \tilde{X})$$

- **Cellular features MSE loss.** Let c be the cellular feature vector corresponds to the input image, and $\tilde{c} = g_c(z_c)$ be the reconstructed cellular feature vector. Then the cellular feature MSE loss ℓ_{CMSE} is defined as the mean-square difference between c and $\tilde{c}$:

$$\ell_{CMSE} = \frac{1}{N} \sum_{i=1}^N (c_i - \tilde{c}_i)^2$$

The loss function we minimize is the weighted sum $\ell = \lambda_1 \cdot \ell_{MSE} + \lambda_2 \cdot \ell_{SSIM} + \lambda_3 \cdot \ell_{CMSE}$. The weights used in our experiment are $\lambda_1 = 1$, $\lambda_2 = 0.3$ and $\lambda_3 = 1$.

For the downstream NEPC classification task training, the binary cross-entropy is used as the loss function. Note that only parameters of the projector p and the MIL module h are updated at this step.

Multiple Instance Learning Framework

Here the encoder is fixed but the weights of the projector will be optimized. The encoder and the projector first transformed the bag of tile images into the bag of tile features. An attention-based MIL module then congregated the bag of tile features and output prediction for the input core/slide (see **Figure 1**). Binary cross entropy is used as the loss function for training.

Reference Standard Label for NEPC

The gold reference standard for determining NEPC status based on IHC criteria from Haffner et al(4). Briefly, NEPC markers from Synaptophysin [SYP], and Insulinoma-associated protein 1 [INSM1] were used in combined with androgen receptor [AR] pathway / prostate lineage markers from AR and NKX3.1. Following the expert criteria, each marker is scored for the percentage of positive cancer cells and overall

intensity. IHC markers were semi-quantitatively evaluated by H-Score, reflecting the positivity of protein expression in a sample. H-score >20 was considered positive as described previously(7). In rare cases where initial origin of metastasis was unknown, prostatic origin was confirmed by IHC for HOXB13(68). Samples with morphology reflecting a suspected secondary or non-prostate origin were excluded. These molecular markers are reported in addition to morphology classification into the following criteria: small cell carcinoma, adenocarcinoma, poorly differentiated carcinoma, and other. These gold standard labels were available in training (UW-RA) and biopsy testing cohorts (UW-MATCH).

).

Visualizing model performance via prediction heatmap

To visualize the model’s performance, we computed and stored the unnormalized scores (logits) for all individual tiles extracted from the core/slide by applying the classifier head to those tile features. These individual tile scores were then mapped linearly onto the indices in the colormap RdBu_r with center set to be 0, and spatially aggregated to form the detection heatmap for the input core/slide. The tiles with positive predictions (positive scores, contribute to positive prediction of the input core/slide) were displayed in red and tiles with negative predictions (negative scores, contribute to negative prediction of the input core/slide) were displayed in blue.

Experimental setup

For semi-supervised pretraining, module f , p , g_I and g_C were trained by optimizing the loss function ℓ with the Adam(69) optimizer using learning rate $1e-4$, exponential decay rate 0.975 and batch-size 100. Data augmentation such as random flip and rotations were used. We also randomly corrupted half of the images with Gaussian noise (mean 0, variance 0.04) during training. The model was trained for a maximum of 50 epochs. The training was stopped when the image MSE loss $\ell_{MSE} < 0.008$, $\ell_{SSIM} < 0.4$ and $\ell_{CMSE} < 0.3$.

For NEPC classification training on rapid autopsy TMA cohort, we used multiple ($n = 5$) 5-fold cross-validation (split at patient-level) to evaluate the performance of each model. All models were trained for a maximum of 40 epochs with Adam optimizer using learning rate $1e-5$, exponential decay rate 0.95 and batch-size 1. Training was stopped when validation did not decrease for 5 epochs in a row and weights that achieved the lowest validation error were stored for performance evaluation.

All experiments were performed on NIH HPC Biowulf cluster (<https://hpc.nih.gov/>)

Random Survival Forest

To evaluate the clinical utility of NE prediction model in biopsy samples, all model output features were considered for input to a Random Survival Forest model to predict time to death. Variables considered for model input included: overall slide-level NE prediction (value 0-1) from NEURAL-PC, total number of tiles from the sample, proportion of individual tiles predicted as NE positive by NEURAL-PC, proportion of individual tiles predicted as NE negative by NEURAL-PC, , average model logits from all tiles, average model attention from all tiles, and average model prediction probability from all tiles. The individual tile-level predictions were determined through same process as described for visualization and heatmap generation above. The model was trained from the UW-MATCH cohort, with outcomes censored to death within 2 years, using the Fast unified random forests for survival, regression, and classification (RF-SRC) package in R using logrank splitting during model fit(70). Out-of-bag predictions were reported for the training set and the model was applied to Weill Cornell cohort as external validation using Explainable Machine Learning in Survival Analysis "survex" library(71).

STATISTICAL ANALYSIS

The main evaluation metrics for NEPC classification include the area under the receiver operating characteristic (AUROC) and the area under the precision-recall curve (AUPRC). The threshold that maximizes Youden's J statistic is chosen as the threshold for NEPC classification. Sensitivity and positive predictive value (PPV) were calculated for NEPC classification performance evaluation. For predicted NE

score distributions, we presented the median as well as the interquartile range (IQR). Statistical evaluation of NEURAL-PC classification compared to gene expression was completed using the wilcox rank sum test. These statistical analyses were conducted using Python (version 3.10), PyTorch (version 1.12.1), Scikit-learn (version 1.3.0), Seaborn (version 0.12.2) and R (version 4.3.3).

STUDY APPROVAL

The study was approved by the University of Washington (IRB#2341) and Fred Hutchinson Cancer Center (IRB#10706) Institutional Review Boards in Seattle, Washington (Clinical trial number: not applicable). All participating men provided written informed consent for a rapid research autopsy and tissue procurement. The Institutional Review Board of Weill Cornell in New York City, New York approved this study (IRB# 20-06022185) and consent was obtained (Clinical trial number: not applicable). The Institutional Review Board of MD Anderson Cancer Center in Houston, Texas approved this study (IRB# 2017-0133) and consent was obtained (Clinical trial number: NCT03263650). This research was approved by the National Institutes of Health in Bethesda, Maryland (IRB 001088), consent was not required due to determination of non-human subjects research involving only anonymized images.

DATA AVAILABILITY STATEMENT

All values for graphical figures are summarized and available in supporting data values are available within the supplemental files. Links to data is provided within the manuscript or supplementary information files, when applicable. Code used to train and generate model predictions and probability maps is available at <https://github.com/NIH-MIP/NEURAL-PC> and model weights are available upon request. H&E WSI imaging data is not publicly available; readers are encouraged to contact corresponding author(s) for reasonable requests.

AUTHOR CONTRIBUTION

Z.C., E.S., H.R., H.Z., L.L., P.C., B.T., M.C.H., S.A.H. contributed to model development, design, computational resources, performance review, and validation. M.T.S., R.R., L.G., E.Y.Y, H.H.C, J.T.N., A.M., D.N., S.T., H.P.B., P.S.N., A.A., C.N.S., M.C.H. contributed to patient accrual. E.S., D.G., R.P., A.G., I.C., R.D., C.M., Y.W., C.C.V., C.F., J.O., T.K., J.M., M.S., P.S.N., R.S., C.N.S., M.C.H. contributed to data collection, including clinical information and/or clinical specimens. E.S., C.C.D., J.M.M., M.C.H. contributed to pathological review. R.D., Y.W., P.S.N. contributed to PDX generation and/or tissue preparation. Z.C., H.R., R.E., M.C.H., S.A.H. contributed to statistical analysis. Z.C., I.C., R.D., C.M., M.T.S., E.Y.Y, H.H.C, C.K.C.D., H.P.B., P.S.N., A.A., C.N.S., M.C.H., S.A.H. contributed to manuscript writing. All authors reviewed the manuscript.

FUNDING

This work is the result of NIH funding, in whole or in part, and is subject to the NIH Public Access Policy. Through acceptance of this federal funding, the NIH has been given a right to make the work publicly available in PubMed Central.

- This research was supported in part by the Center for Cancer Research, National Cancer Institute, National Institutes of Health Intramural Research Program project number ZIABC012163 (to: SAH). The contributions of the NIH author(s) were made as part of their official duties as NIH federal employees, are in compliance with agency policy requirements, and are considered Works of the United States Government. However, the findings and conclusions presented in this paper are those of the author(s) and do not necessarily reflect the views of the NIH or the U.S. Department of Health and Human Services.
- NIH/NCI 1P01CA298991-01 Grant; to: MCH, PSN
- NIH/NCI R37CA286450 Grant; to: MCH
- NIH/NCI P30CA15704 (Fred Hutch/University of Washington Cancer Center Support Grant) Grant; to: PSN (Cancer Center leadership)
- NIH/NCI P50CA097186 (Pacific Northwest Prostate Cancer SPORE) Grant; to: PSN, HHC
- NIH/NCI P01CA163227 Grant; to: PSN, MCH
- NIH/NCI R01CA234715 Grant; to: HPB
- NIH/NCI R01CA266452 Grant; to: HPB
- NIH/NCI R01CA280056 Grant; to: MCH
- NIH Office of Research Infrastructure Programs (ORIP) S10OD028685 Grant; to: IC
- U.S. DoD Prostate Cancer Research Program W81XWH-20-1-0111 Grant; to: PSN
- U.S. DoD Prostate Cancer Research Program W81XWH-21-1-0229 Grant; to: HHC
- U.S. DoD Prostate Cancer Research Program W81XWH-22-1-0278 / W81XWH-22-1-0279 Grant; to: MCH, SAH

- U.S. DoD Prostate Cancer Research Program W81XWH-18-1-0347 Grant; to: PSN
- U.S. DoD Prostate Cancer Research Program W81XWH-18-1-0689 Grant; to: HHC
- U.S. DoD Prostate Cancer Research Program W81XWH-21-1-0264 Grant; to: MTS
- U.S. DoD Prostate Cancer Research Program W81XWH-14-2-0183 Grant; to: PSN
- U.S. DoD Prostate Cancer Research Program W81XWH-17-2-0043 Grant; to: HHC
- U.S. DoD Prostate Cancer Research Program PC230420 Grant; to: MCH
- U.S. DoD Prostate Cancer Research Program PC230582 Grant; to: HHC
- Doris Duke Charitable Foundation (2021184) Grant; to: HHC
- V Foundation Grant; to: MCH
- PCF Felix Feng PC-SYNERGY Award Grant; to: PSN, MCH
- Safeway Foundation Grant; to: MCH
- Richard M. Lucas Foundation Grant; to: PSN
- Brotman Baty Institute for Precision Medicine Grant; to: MCH
- UW/FHCC Institute for Prostate Cancer Research Grant; to: MCH

ACKNOWLEDGEMENT

We are grateful to the patients and their families, and the rapid autopsy teams for their contributions to the University of Washington Medical Center Prostate Cancer Donor Rapid Autopsy Program. We also thank the members of the Haffner and Nelson laboratories at Fred Hutch and the Artificial Intelligence Resource at NCI for their constructive suggestions. From the Nelson lab, we thank Yuzhuo Wang, Lisa Ang, Talina Nunez, Tarana Arman, Canan Dirican, Helen Akushie, and Fred Hutch TRMS Core for their contributions in PDX generation, propagation, and harvesting. From the Nelson lab, we thank Emery Boehnke, Hannah Meade, and Ruth Dumpit for tissue fixation and IHC. All computational experiments were performed on NIH HPC Biowulf cluster (<https://hpc.nih.gov/>)

REFERENCES

1. Sherman RL, et al. Annual Report to the Nation on the Status of Cancer, featuring state-level statistics after the onset of the COVID-19 pandemic. *Cancer*. 2025;131(9):e35833.
2. Raychaudhuri R, Lin DW, Montgomery RB. Prostate Cancer: A Review. *JAMA*. 2025;333(16):1433–1446.
3. Graham LS, et al. Clinical, pathologic, and molecular features of ampicrine prostate cancer. *Prostate*. 2023;83(7):641–648.
4. Haffner MC, et al. Framework for the Pathology Workup of Metastatic Castration-Resistant Prostate Cancer Biopsies. *Clin Cancer Res*. 2025;31(3):466–478.
5. Haffner MC, et al. Genomic and phenotypic heterogeneity in prostate cancer. *Nat Rev Urol*. 2021;18(2):79–92.
6. De Sarkar N, et al. Nucleosome Patterns in Circulating Tumor DNA Reveal Transcriptional Regulation of Advanced Prostate Cancer Phenotypes. *Cancer Discov*. 2023;13(3):632–653.
7. Roudier MP, et al. Patterns of intra- and intertumor phenotypic heterogeneity in lethal prostate cancer. *J Clin Invest*. 2025;135(15). <https://doi.org/10.1172/JCI186599>.
8. Zaidi S, et al. Single-cell analysis of treatment-resistant prostate cancer: Implications of cell state changes for cell surface antigen-targeted therapies. *Proc Natl Acad Sci U S A*. 2024;121(28):e2322203121.
9. Cejas P, et al. Subtype heterogeneity and epigenetic convergence in neuroendocrine prostate cancer. *Nat Commun*. 2021;12(1):5775.

10. Davies AH, Beltran H, Zoubeidi A. Cellular plasticity and the neuroendocrine phenotype in prostate cancer. *Nat Rev Urol*. 2018;15(5):271–286.
11. Aggarwal R, et al. Clinical and Genomic Characterization of Treatment-Emergent Small-Cell Neuroendocrine Prostate Cancer: A Multi-institutional Prospective Study. *J Clin Oncol*. 2018;36(24):2492–2503.
12. Beltran H, et al. Divergent clonal evolution of castration-resistant neuroendocrine prostate cancer. *Nat Med*. 2016;22(3):298–305.
13. Bluemn EG, et al. Androgen Receptor Pathway-Independent Prostate Cancer Is Sustained through FGF Signaling. *Cancer Cell*. 2017;32(4):474-489.e6.
14. Hu J, Han B, Huang J. Morphologic Spectrum of Neuroendocrine Tumors of the Prostate: An Updated Review. *Arch Pathol Lab Med*. 2020;144(3):320–325.
15. Epstein JI, et al. Proposed morphologic classification of prostate cancer with neuroendocrine differentiation. *Am J Surg Pathol*. 2014;38(6):756–767.
16. Beltran H, et al. The Role of Lineage Plasticity in Prostate Cancer Therapy Resistance. *Clin Cancer Res*. 2019;25(23):6916–6924.
17. Davies A, et al. Biological Evolution of Castration-resistant Prostate Cancer. *Eur Urol Focus*. 2019;5(2):147–154.
18. Beltran H, Demichelis F. Therapy considerations in neuroendocrine prostate cancer: what next? *Endocr Relat Cancer*. 2021;28(8):T67–T78.

19. Beltran H, et al. Aggressive Variants of Castration-Resistant Prostate Cancer. *Clinical Cancer Research*. 2014;20(11):2846–2850.
20. Conteduca V, et al. Clinical features of neuroendocrine prostate cancer. *European Journal of Cancer*. 2019;121:7–18.
21. Nyquist MD, et al. Combined TP53 and RB1 Loss Promotes Prostate Cancer Resistance to a Spectrum of Therapeutics and Confers Vulnerability to Replication Stress. *Cell Rep*. 2020;31(8):107669.
22. Labrecque MP, et al. Molecular profiling stratifies diverse phenotypes of treatment-refractory metastatic castration-resistant prostate cancer. *J Clin Invest*. 2019;129(10):4492–4505.
23. Coleman IM, et al. Therapeutic Implications for Intrinsic Phenotype Classification of Metastatic Castration-Resistant Prostate Cancer. *Clin Cancer Res*. 2022;28(14):3127–3140.
24. Brady L, et al. Inter- and intra-tumor heterogeneity of metastatic prostate cancer determined by digital spatial gene expression profiling. *Nat Commun*. 2021;12(1):1426.
25. Aparicio AM, et al. Platinum-based chemotherapy for variant castrate-resistant prostate cancer. *Clin Cancer Res*. 2013;19(13):3621–3630.
26. Aparicio AM, et al. Combined Tumor Suppressor Defects Characterize Clinically Defined Aggressive Variant Prostate Cancers. *Clin Cancer Res*. 2016;22(6):1520–1530.
27. Butler W, Huang J. Neuroendocrine cells of the prostate: Histology, biological functions, and molecular mechanisms. *Precis Clin Med*. 2021;4(1):25–34.
28. Goldenberg SL, Nir G, Salcudean SE. A new era: artificial intelligence and machine learning in prostate cancer. *Nat Rev Urol*. 2019;16(7):391–403.

29. Cuocolo R, et al. Machine learning applications in prostate cancer magnetic resonance imaging. *Eur Radiol Exp*. 2019;3(1):35.
30. Riaz IB, et al. Applications of Artificial Intelligence in Prostate Cancer Care: A Path to Enhanced Efficiency and Outcomes. *Am Soc Clin Oncol Educ Book*. 2024;44(3):e438516.
31. Rubinstein E, et al. Unsupervised tumor detection in Dynamic PET/CT imaging of the prostate. *Med Image Anal*. 2019;55:27–40.
32. Yu G, et al. Accurate recognition of colorectal cancer with semi-supervised deep learning on pathological images. *Nat Commun*. 2021;12(1):6311.
33. Uzunova H, et al. Unsupervised pathology detection in medical images using conditional variational autoencoders. *Int J Comput Assist Radiol Surg*. 2019;14(3):451–461.
34. Masci J, et al. Stacked convolutional auto-encoders for hierarchical feature extraction. *International conference on artificial neural networks*. Springer; 2011:52–59.
35. Antonio VAA, et al. Classification of lung adenocarcinoma transcriptome subtypes from pathological images using deep convolutional networks. *Int J Comput Assist Radiol Surg*. 2018;13(12):1905–1913.
36. X. Li, et al. Discriminative Pattern Mining for Breast Cancer Histopathology Image Classification via Fully Convolutional Autoencoder. *IEEE Access*. 2019;7:36433–36445.
37. Zhou X, et al. Whole slide image based prognosis prediction in rectal cancer using unsupervised artificial intelligence. *BMC Cancer*. 2024;24(1):1523.

38. A. Lomacenkova, O. Arandjelović. Whole Slide Pathology Image Patch Based Deep Classification: An Investigation of the Effects of the Latent Autoencoder Representation and the Loss Function Form. 2021 *IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*. 2021:1–4.
39. Baldi P. Autoencoders, Unsupervised Learning, and Deep Architectures. In: Guyon I, et al., eds. *Proceedings of ICML Workshop on Unsupervised and Transfer Learning*. Bellevue, Washington, USA: PMLR; 2012:37–49.
40. Chen Z, et al. A sparse convolutional autoencoder for joint feature extraction and clustering of metastatic prostate cancer images. *International Conference on Artificial Intelligence in Medicine*. Springer; 2024:9–20.
41. Hou L, et al. Sparse Autoencoder for Unsupervised Nucleus Detection and Representation in Histopathology Images. *Pattern Recognit*. 2019;86:188–200.
42. Roy M, et al. Convolutional autoencoder based model HistoCAE for segmentation of viable tumor regions in liver whole-slide images. *Sci Rep*. 2021;11(1):139.
43. Roudier MP, et al. Phenotypic heterogeneity of end-stage prostate carcinoma metastatic to bone. *Hum Pathol*. 2003;34(7):646–653.
44. Roudier MP, et al. Characterizing the molecular features of ERG-positive tumors in primary and castration resistant prostate cancer. *Prostate*. 2016;76(9):810–822.
45. Ajkunic A, et al. Assessment of TROP2, CEACAM5 and DLL3 in metastatic prostate cancer: Expression landscape and molecular correlates. *NPJ Precis Oncol*. 2024;8(1):104.
46. Kumar A, et al. Substantial interindividual and limited intraindividual genomic diversity among tumors from men with metastatic prostate cancer. *Nat Med*. 2016;22(4):369–378.

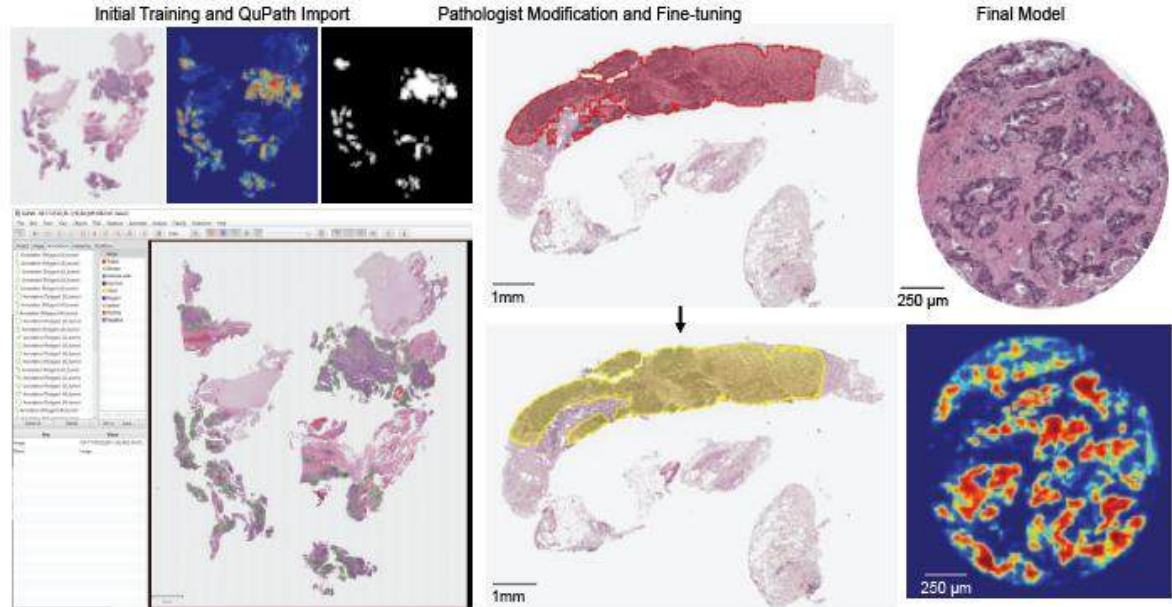
47. Sayar E, et al. Reversible epigenetic alterations mediate PSMA expression heterogeneity in advanced metastatic prostate cancer. *JCI Insight*. 2023;8(7). <https://doi.org/10.1172/jci.insight.162907>.
48. Stringer C, et al. Cellpose: a generalist algorithm for cellular segmentation. *Nat Methods*. 2021;18(1):100–106.
49. Lu C, et al. A prognostic model for overall survival of patients with early-stage non-small cell lung cancer: a multicentre, retrospective study. *Lancet Digit Health*. 2020;2(11):e594–e606.
50. He K, et al. Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016:770–778.
51. Liu Z, et al. Swin transformer: Hierarchical vision transformer using shifted windows. *Proceedings of the IEEE/CVF international conference on computer vision*. 2021:10012–10022.
52. Chen RJ, et al. Towards a general-purpose foundation model for computational pathology. *Nat Med*. 2024;30(3):850–862.
53. Lin D, et al. High fidelity patient-derived xenografts for accelerating prostate cancer discovery and drug development. *Cancer Res*. 2014;74(4):1272–1283.
54. Prendeville S, et al. Prostate carcinoma with amphicrine features: further refining the spectrum of neuroendocrine differentiation in tumours of primary prostatic origin? *Histopathology*. 2017;71(6):926–933.
55. Cuzick J, et al. Prognostic value of an RNA expression signature derived from cell cycle proliferation genes in patients with prostate cancer: a retrospective study. *Lancet Oncol*. 2011;12(3):245–255.

56. Pilié PG, et al. *Randomized phase II study of olaparib (Ola) maintenance following cabazitaxel-carboplatin induction chemotherapy (CabCarb) in men with aggressive variant prostate cancer (AVPC)*. American Society of Clinical Oncology; 2023.
57. Subramani E, et al. Targeting arginine metabolism overcomes chemotherapy resistance in aggressive-variant prostate cancers. *iScience*. 2026;29(6):116184.
58. Amgad M, et al. A population-level digital histologic biomarker for enhanced prognosis of invasive breast cancer. *Nat Med*. 2024;30(1):85–97.
59. Lu C, et al. Feature-driven local cell graph (FLoCK): New computational pathology-based descriptors for prognosis of lung cancer and HPV status of oropharyngeal cancers. *Med Image Anal*. 2021;68:101903.
60. Kapse S, et al. Subtype-specific spatial descriptors of tumor-immune microenvironment are prognostic of survival in lung adenocarcinoma. *2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI)*. IEEE; 2022:1–5.
61. Huang X, et al. Fusing hand-crafted and deep-learning features in a convolutional neural network model to identify prostate cancer in pathology images. *Front Oncol*. 2022;12:994950.
62. Goel P, et al. Coca-Mil: Attention-Based Handcrafted-Deep Feature Fusion in Computational Pathology. *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*. IEEE; 2024:1–5.
63. Zheng K, et al. Deep learning model with pathological knowledge for detection of colorectal neuroendocrine tumor. *Cell Rep Med*. 2024;5(10):101785.
64. Arlova A, et al. Artificial Intelligence-based Tumor Segmentation in Mouse Models of Lung Adenocarcinoma. *J Pathol Inform*. 2022;13:100007.

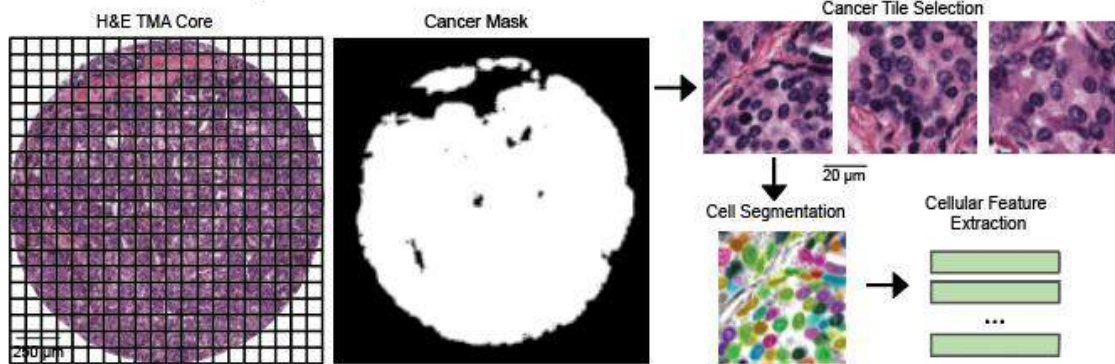
65. Patkar S, et al. A selective CutMix approach improves generalizability of deep learning-based grading and risk assessment of prostate cancer. *J Pathol Inform.* 2024;15:100381.
66. Macenko M, et al. A method for normalizing histology slides for quantitative analysis. *2009 IEEE international symposium on biomedical imaging: from nano to macro.* IEEE; 2009:1107–1110.
67. Vaswani A, et al. Attention is all you need. *Advances in neural information processing systems.* 2017;30.
68. Patel RA, et al. Characterization of HOXB13 expression patterns in localized and metastatic castration-resistant prostate cancer. *J Pathol.* 2024;262(1):105–120.
69. Kingma DP. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980.* 2014.
70. Ishwaran H, Kogalur UB. Fast unified random forests for survival, regression, and classification (RF-SRC). *R package version.* 2019;2(1).
71. Spytek M, et al. survex: an R package for explaining machine learning survival models. *Bioinformatics.* 2023;39(12):btad723.

FIGURES

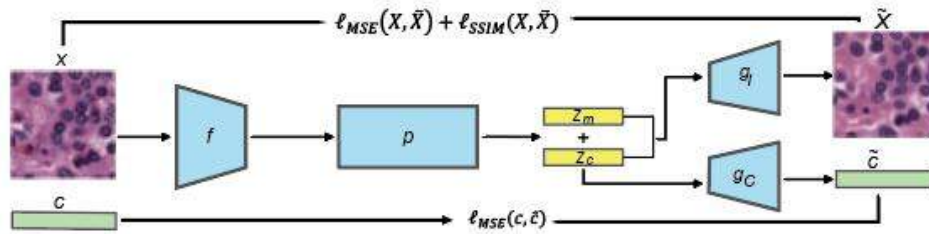
A Cancer Detection Model



B Tile Selection and Processing



C Unsupervised Model Pre-training



D Supervised NEPC Classification

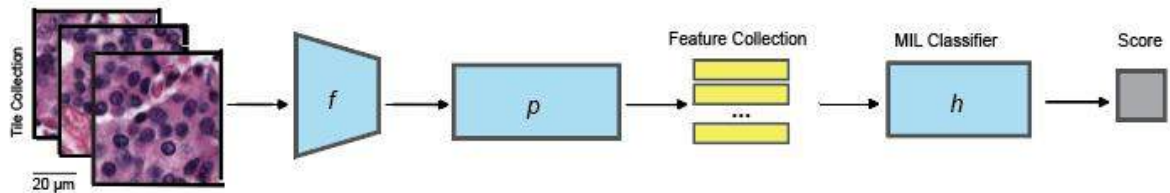


Figure 1. The workflow of the proposed NEURAL-PC model. (A) Following initial training from a subset of manually annotated images, a preliminary cancer detection algorithm is applied to each WSI, resulting in a probability map output and binary segmentation mask, which is then converted to geojson format for import to QuPath and manual refinement by pathologists. General modifications to reduce false positives were then used for re-training the cancer detection model. The final model results in highly specific segmentation of cancerous regions which performs well on various tissue sources. (B) Tiles of size 250x250 at 40x magnification with more than 10% cancer area were selected, nuclei were segmented using CellPose(48) algorithm, and the cellular features(59) vectors correspond to these tiles were extracted. (C) In the unsupervised pre-training step, the model was trained to reconstruct tiles as well as the cellular features. Tiles transformed by an encoder, f , and a projector, p , into latent vectors z_m and z_c , then a decoder g_c uses only z_c to reconstruct cellular feature vector correspond to the input tile and another decoder g_l uses the sum $z = z_m + z_c$ to reconstruct the input tile image. The objective is to minimize the differences between the input tile image and the reconstructed tile image, as well as the difference between the input cellular features and the reconstructed cellular features (see Methods). (D) The trained encoder f and projector p were combined with a multiple instance learning classifier h to aggregate all tile features corresponding to an input image to produce a core/slide-level NEPC prediction. Digital WSI were originally acquired at $0.25\mu\text{m}/\text{pixel}$ (40x magnification), scale bars for each image are shown.

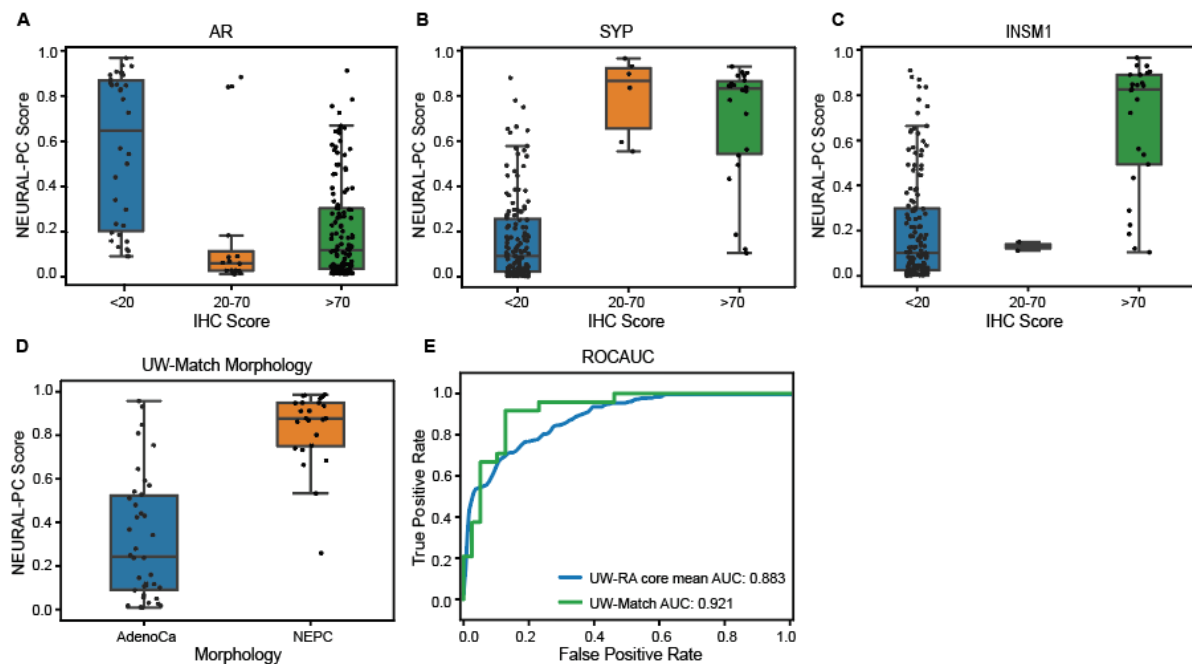


Figure 2. Model performances on internal and external validation cohorts. (A-C) NEURAL-PC predicted NE score distributions for UW-RA TMA testing cores in different IHC-score groupings for AR (A), SYP (B), and INSM1 (C). Generally, ≤ 20 is considered low/weak expression, 20-70 moderate expression, and ≥ 70 strong expression. (D) NEURAL-PC predicted NE score distributions for UW-Match cohort. AdenoCa = adenocarcinoma, NEPC= neuroendocrine prostate cancer with strict IHC guideline adherence. (E) ROC curves for UW-RA TMA core dataset (mean AUC from CV) and UW MATCH WSI dataset, respectively.

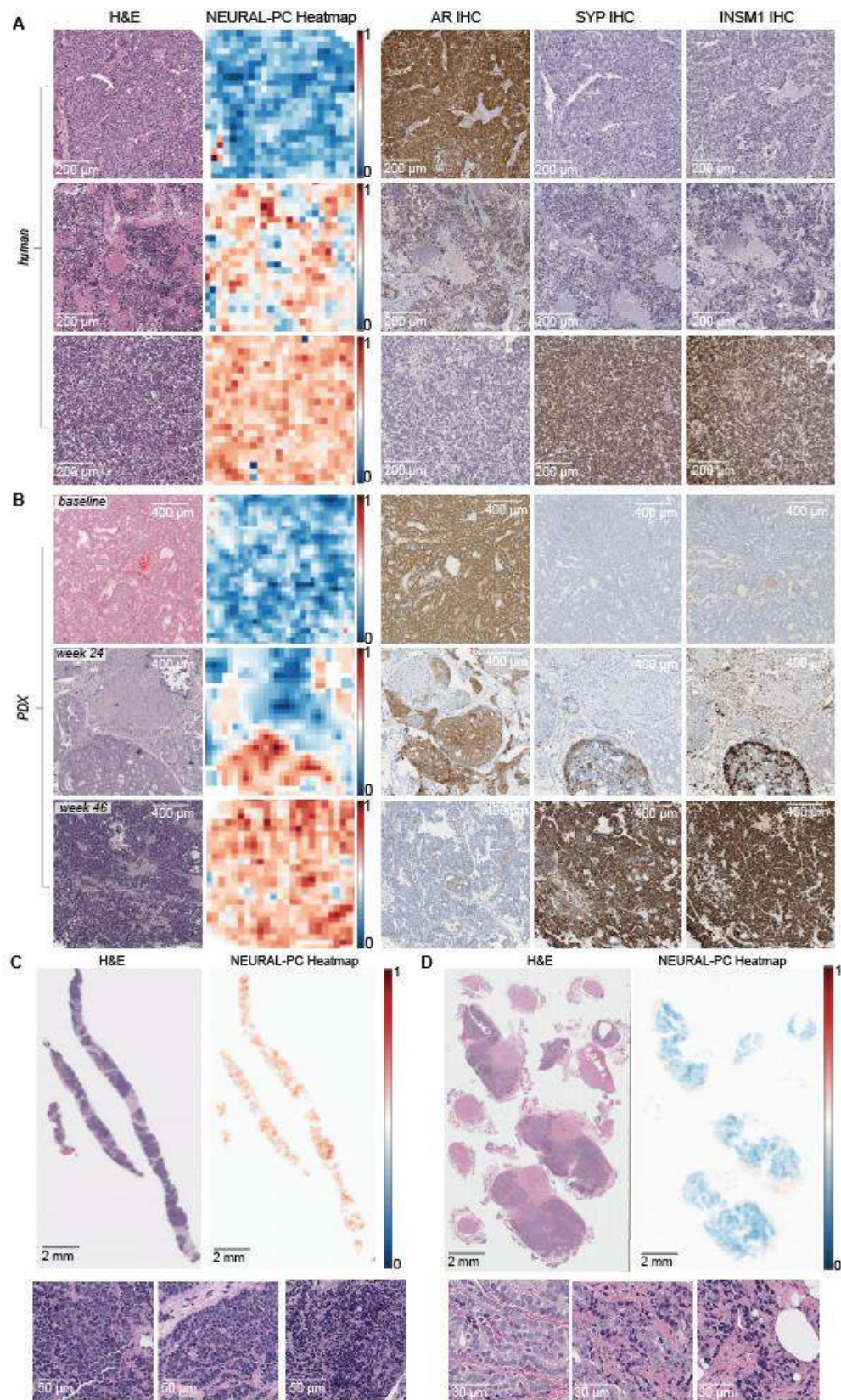


Figure 3. Representative visual NEURAL-PC predictions and spatial IHC association.

(A) Representative test core images from UW-RA TMA set. High-power view of H&E morphology, NEURAL-PC probability maps and corresponding AR, SYP, and INSM1 IHC stains. Top: a true negative from AR-positive (AR H-score: 160, SYP H-score: 0, INSM1 H-score: 10) with NEURAL-PC prediction <0.01 and uniform negative/low heatmap. Middle: a mixed phenotype true positive (AR H-score: 40, SYP H-score: 40, INSM1 H-score: 0) with NEURAL-PC prediction 0.83 and focal positivity in prediction heatmap. Bottom: true positive from NE-positive (AR H-score: 0, SYP H-score: 200, INSM1 H-score: 150) with NEURAL-PC prediction 0.93 and uniform positive prediction heatmap. **(B)** Representative images from PDX validation set. NE prediction heatmaps from NEURAL-PC model are shown separately, along with AR, SYP, and INSM1 IHC stains. Top: AR-positive sample from baseline/untreated PC, correctly predicted as negative by NEURAL-PC (probability score 0.04) with uniform low probability. Middle: mixed phenotype from transitionary phase (week 24), predicted as negative by NEURAL-PC (0.154) with focal AR and NE IHC marker positivity. Bottom: NE-positive sample from post-castration (week 46) PDX model, correctly predicted as positive by NEURAL-PC (probability score 0.851) with uniform positive prediction map. **(C)** Representative true positive classification from UW-Match cohort. NEURAL-PC probability map reflects uniformly positive predictions across the tissue sample, with overall slide classification score 0.948. Homogeneous morphology is observed across the sample (tile regions shown below). **(D)** Representative true negative classification from UW-Match cohort, with overall NEURAL-PC classification score of 0.055. Majority of the cancer burden shows adenocarcinoma-like morphology (left and middle tiles). Of note, a small proportion of focal areas demonstrate higher tile-level probabilities for containing NEPC in the NEURAL-PC heatmap (right tile region). Digital WSI were originally acquired at $0.25\mu\text{m}/\text{pixel}$ (40x magnification), scale bars for each image are shown.

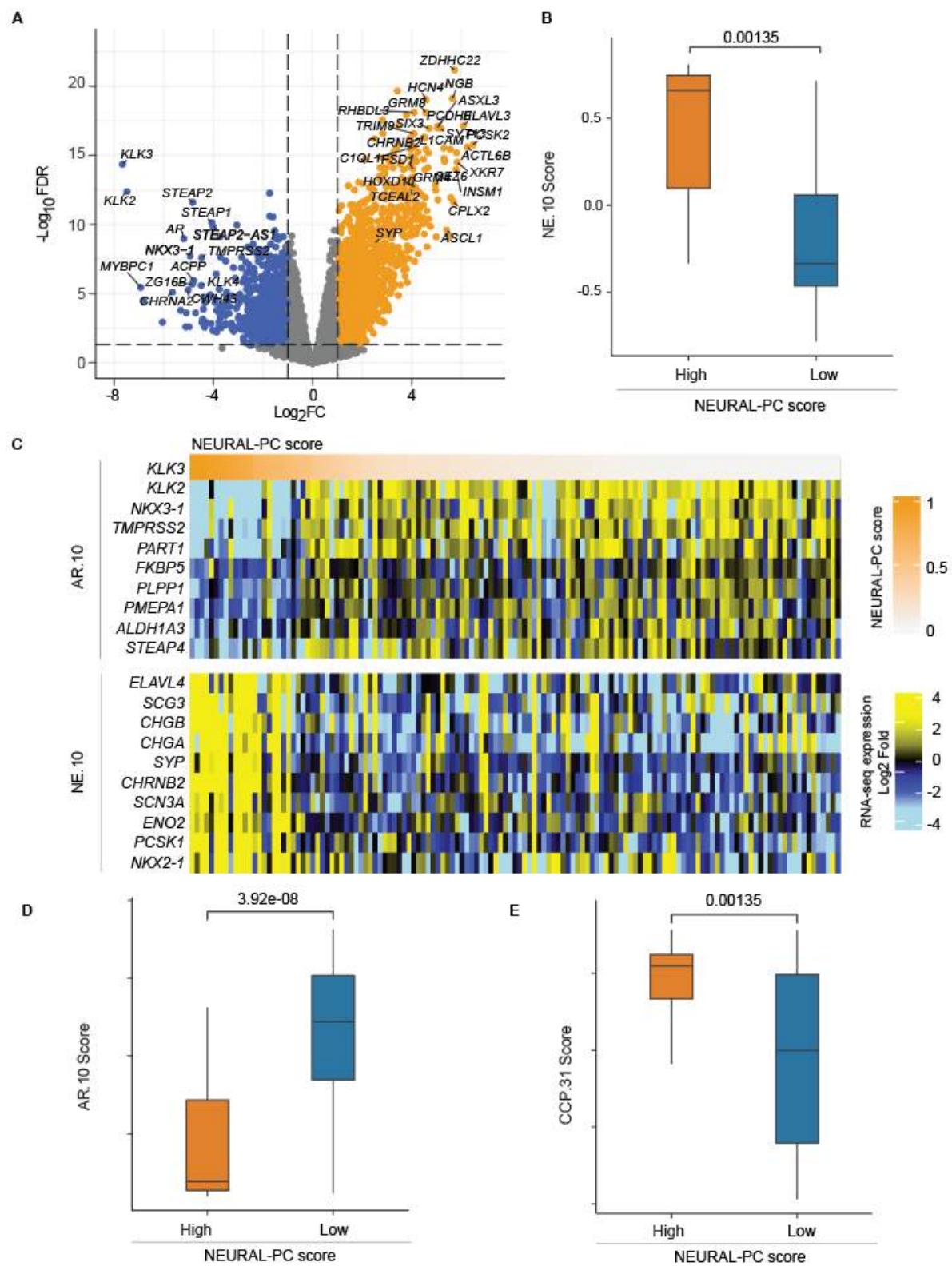


Figure 4. Association with RNAseq expression in UW-RA WSI cohort. (A) Differential Expression analysis for WSI called as positive for NEPC by NEURAL-PC algorithm (upregulated on right) vs those called as negative for NEPC. **(B)** NE.10 expression score vs. NEURAL-PC binary classification. **(C)** Association to gene expression profiles in 10 NE-related genes and 10 AR-related genes, shown in rank-order from NEURAL-PC classification probability output **(D)** AR.10 expression score vs. NEURAL-PC binary classification. **(E)** CCP.31 expression score vs. NEURAL-PC binary classification. P-values determined from wilcox rank sum test.

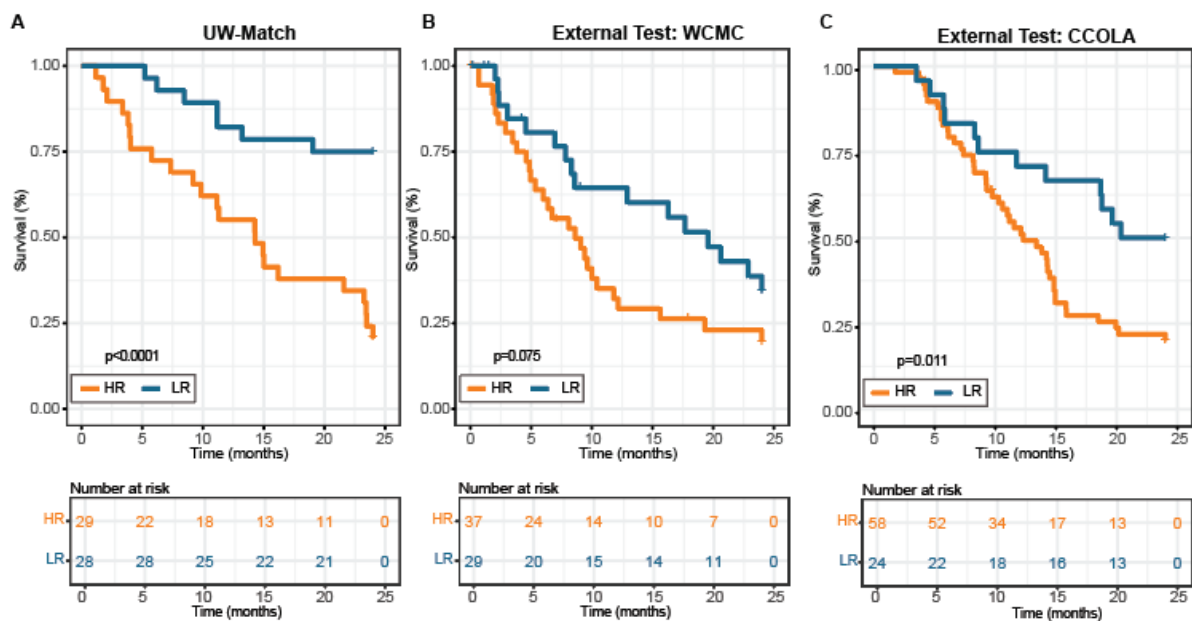


Figure 5. Kaplan-Meier curves for RSF model predicting likelihood of death within 2 years from biopsy. (A) out-of-bag predictions in training cohort (UW-MATCH), **(B)** external WCMC NEPC cohort, and **(C)** external CCOLA AVPC cohort. P-values from log-rank test comparing groups predicted not to reach 2 year survival (HR) and groups predicted to reach 2 year survival (LR).

TABLES

Table 1. Performance comparison on the internal UW-RA TMA test cohort.

Model	AUROC (NE)	AUPRC (NE)
NEURAL-PC	0.883 ± 0.016	0.717 ± 0.031
ResNet50*	0.825 ± 0.055	0.536 ± 0.060
Swin Transformer*	0.857 ± 0.022	0.541 ± 0.079
UNI	0.900 ± 0.026	0.704 ± 0.110
CAE only	0.652 ± 0.049	0.396 ± 0.050
HCF only	0.864 ± 0.022	0.623 ± 0.113

*indicates models pre-trained on ImageNet. CAE = Convolutional Autoencoder, HCF = Handcrafted Cellular Features.

Table 2. NEURAL-PC classification results for UW-MATCH cohort

NEURAL-PC prediction		
	AI positive	AI negative
AdenoCA ($n = 38$)	8	30
NEPC ($n = 24$)	21	3
Amphicrine ($n = 1$)	0	1

. Overall Sensitivity = 0.875 and Positive Predictive Value (PPV) = 0.724. NEPC = Neuroendocrine

Prostate Cancer, AdenoCA = Adenocarcinoma

Table 3. NE classification results for Weill Cornell cohort

Type	NEURAL-PC positive	NEURAL-PC negative
NEPC (<i>n</i> = 58)	30	28
Amphicrine (<i>n</i> = 27)	11	16

. Threshold for classification determined from UW-MATCH cohort (0.598). NEPC = Neuroendocrine

Prostate Cancer