“PREDICTIVE AI-MODEL TO MANAGE TRANSCRIPTOMICS-BASED TRANSITION OF HORMONE-NAIVE TO CASTRATE-RESISTANT PROSTATE CANCER: AN OPPORTUNITY TOWARDS PRECISION MEDICINE”

Research Paper

Neetu Singh, King George's Medical University, Uttar Pradesh, India. neetaashi@yahoo.com, neetusingh@kgmcindia.edu

David Annan, Swiss School of Business and Management, Geneva, Switzerland.

david.annan@ssbm.ch, davidannan2000@gmail.com

“Abstract”

Precision medicine is a groundbreaking approach to the treatment of prostate cancer. It involves tailoring medical decisions and therapies to individual patients based on their unique genetic makeup, lifestyle, and environmental factors. By analysing a patient's genetic profile, doctors can identify specific genetic alterations that drive the development and progression of prostate cancer. This knowledge allows for the selection of targeted therapies that directly address the underlying molecular abnormalities, increasing treatment efficacy and minimizing side effects. Precision medicine in prostate cancer offers the potential for more personalized and effective treatments, ultimately improving patient outcomes and quality of life. Hence, the study developed a highly accurate Convolutional Neural Network-1Dimensional (CNN) -1D model to predict the different pathological stages of PCa and castrate-resistant prostate cancer/Hormone resistant using a dataset generated on Affymetrix GeneChip Human Transcriptome Array 2.0 with accuracy of 0.98.

Keywords: *Predictive Model, Transcriptomics, Hormone-Naive Prostate Cancer, Castrate-Resistant Prostate Cancer, Precision Medicine*

1. Introduction

The India oncology market is expected to grow by $ 734.18 million, accelerating at a compound annual growth rate (CAGR) of almost 13.02% during the forecast period of 5 years. Technavio’s research report traces the growth trajectory of the oncology market and provides a detailed analysis of the prevalent market forces, trends, and drivers that are likely to impact the market in focus. Prostate Cancer (PCa) has been a major malice ~1,414,259

fresh cases of PCa and 375,304 PCa-related mortality across the world shows its high prevalence (Sung et al., 2021). Local spread or borderline resectable PCa (>4cm cancer) presents a favorable prognosis, castration, chemo- and radiotherapy are traditionally used for patient survival. Although, many advanced PCa generally become resistant to available treatment and no definite cure exists (Thomas & Pachynski, 2018), Hormone-naive PCa (HNPC) is an androgen (testosterone)-dependent for their existence and progression, the customary therapy is androgen deprivation therapy (ADT) which gives very low levels of testosterone and effective initial tumor-reduction, relapse occurs and gives rise to castrate-resistant prostate cancer (CRPC)

Although many treatments are accessible for CRPCs, including a combination of androgen receptor pathway inhibitors (ARPIs, such as the first-generation anti-androgens abiraterone acetate, enzalutamide, apalutamide, or darolutamide) and chemotherapies (docetaxel and cabazitaxel), and radionuclides (Beer et al., 2017; Berthold et al., 2008; Chang et al., 2019; Fallara et al., 2020; Kellokumpu-Lehtinen et al., 2020; Pu et al., 2022; Ryan et al., 2015), and most give primary or acquired resistances in CRPC, all available therapies remain comforting and lead to progression-free survival of only for 9–30 months (Labriola et al., 2021; Rebello et al., 2021). Thus, more research efforts are warranted to determine new predictive biomarkers that may help in tailored therapeutic resolutions for advanced PCa. The study uses Deep Learning (DL) programs to identify novel prognostic biomarkers to help avoid resistance during the transition from Hormone-naive PCa to CRPC or may help to develop novel resistant biomarkers for which small molecule therapy can be developed. Identification of biomarkers plays a dynamic role in the field of Precision Medicine Market of Oncology which may help in tailored therapeutic resolutions for advanced PCa.

1.1. Machine learning method

Although various Machine Learning (ML) based methods used for the early clinical detection of PCa do not point towards a specific diagnosis, ML-based multiparametric magnetic resonance imaging (mpMRI) has been suggested mainly for clinically diagnosed PCa cases (Bi et al., 2019). PCa risk stratification as per ML-based mpMRI-derived radiomic features had better prediction accuracy and capability to avoid up to 50% unnecessary biopsies (Chiu et al., 2022; Varghese et al., 2019; Yu et al., 2021). However, patients may times neglect the tumor in its early phases and look for treatment when it has already developed. Therefore,

early manifestation of PCa is important for which different screening tests are available like prostate-specific antigen (PSA) testing, digital rectal examinations (DRE), and transrectal ultrasonography (TRUS) guided prostate system biopsies in clinics (Ilic et al., 2018; Vale et al., 2020). Hence to avoid the overdiagnosis of PCa of the existing diagnostic tools PSA testing, DRE, and TRUS-guided prostate system biopsies (Mottet et al., 2021), it is important to use a dynamic contrast-enhanced Magnetic Resonance Imaging (MRI), and diffusion-weighted imaging (DWI) to increase the early diagnostic efficiency and also helped treatment decision (Kim and Park, 2008).

1.2 Deep learning algorithm

Precision Medicine includes analysis of gene, transcript, and protein levels in individual patients to design their treatment plan, follow how the cancer treatment and cancer diagnostics progress and identify which small molecule will be able to handle the developed resistance. This supplements the segregated markets like hospitals, diagnostic laboratories, and biopharmaceutical companies. Gene expression profiles have been used as raw inputs into the layers of intermediate features for building a model. A DL algorithm of high predictive accuracy was developed to typify the samples as normal or tumor by using publicly available gene expression databases (Ahn et al., 2018; Gupta et al., 2022). Gene expression data and DL technique, a subset amalgamates and address various questions such as assessment of survival times of patients with cancer, establishment of biomarker genes or transcripts (Xie et al., 2021) biomarkers based effective therapeutics for cancer treatment (Zeng et al. 2021), and improve prognosis of cancer (S. Huang et al., 2020; Z. Huang et al., 2020).

Using the DL algorithm helps transcripts related to the DNA replication stress model to predict good or poor prognosis and accordingly tailor therapy in primary prostate cancer (Huang et al., 2023). All these studies show that combining transcriptomics data and the DL technique will provide important information about the diagnosis, prognosis, and targeted therapeutics of cancer. Hence, the study aims to train probe-based differential transcript expression (microarray) among various stages of PCa which may help in prognosis, management, and aggressiveness.

2. Methodology

To generate prognostics analysis of precision oncology, the study downloaded the PCa data deposited to the NCBI Gene Expression Omnibus (GSE200879_RAW.tar) which was made

public on May 18, 2022. A total of 137 CEL files were analyzed (Table 1). The PAIR prostate (PP) dataset which includes 9 normal peritumoral tissues, 116 hormone-naïve prostate cancer (HNPC) tissues, and 13 castrate-resistant prostate cancer (CRPC) tissues processed on Affymetrix Gene Chip HTA 2.0 conducted by Firlej et al., (2022).

Table 1.

Specimens	Gleason score	Stages based on TNM features	Individual number of samples	Total number of Samples
Normal				9
low risk	GS-6/7(3+ 4)	pT2a/b/c	7+2+27	36
Intermediate risk	GS-7 (4 + 3)	pT3a	46	46
high risk	GS-7 (4 + 3)	pT3b/4a	28+6	34
Hormone resistant				13

The PAIR prostate (PP) dataset

Table 1. Except for normal the tumor samples were classified into low-risk, intermediate-risk, and high-risk groups using an alteration in D'Amico's classification, which does not take into account the PSA rate but only the histologic data based on the Gleason Score (GS) and TNM features.

The above 137 CEL files were generated by processing total RNA on Affymetrix HTA 2.0 arrays according to the manufacturer's recommendations (Affymetrix, Santa Clara, Calif) and were scanned through GENECHIP Scanner-7G (Affymetrix, CA) (Singh et al., 2016) log2 signal expression of 16202 transcripts as shown in Supplementary Table1 as represented in figure 1.

1.1 Figure

Figure 1 Total, upregulated, and downregulated differentially expressed transcripts in each combination generated by Transcriptome Analysis Console v3.0. 1b. Represents the source of variation was 0.14.

The data was analyzed using the Affymetrix Human Transcriptome Array 2.0 has been designed to aid in human disease research and clinical translational medicine by providing the most comprehensive view of the transcriptome including >285,000 full-length transcripts covered (>245,000 coding transcripts; >40,000 non-coding transcripts >339,000 probe sets covering exon-exon junctions) and measures alternative splicing events/transcript variants. The PAIR prostate (PP) dataset which includes 9 normal peritumoral tissues, 116 hormone-naïve prostate cancer (HNPC) tissues, and 13 castrate-resistant prostate cancer (CRPC) tissues processed on Affymetrix Gene Chip HTA 2.0 conducted by Firlej et al., (2022).

2.1. Model building

A large 1-dimensional data in the form of CSV or Excel Convolutional Neural Network - 1Dimension (CNN-1D) has been recommended. CNN-1D model was built by importing numpy as np, pandas as pd, and matplotlib. pyplot as plt and seaborn as sns packages were imported to read and process the raw data for Exploratory Data Analysis (EDA). Various steps are involved in the CNN-1D model a) Convolution means to create a convolved

feature/matrix; b) The activation factor ReLu is used to add non-linearity and to eliminate the negative values; c) Pooling of neurons which may occur through any of the three ways like average, minimum or maximum pooling; d) Flattening of neural network and creating a fully connected (Fc) neural network layer. A few other steps are also involved like Strides, Padding, and Kernel Matrix. Hence all the above steps create a convolved matrix which reduces the dimensionality of the 1D data and ensures important features are selected to create the matrix.

Figure 2-Dimensional model display count

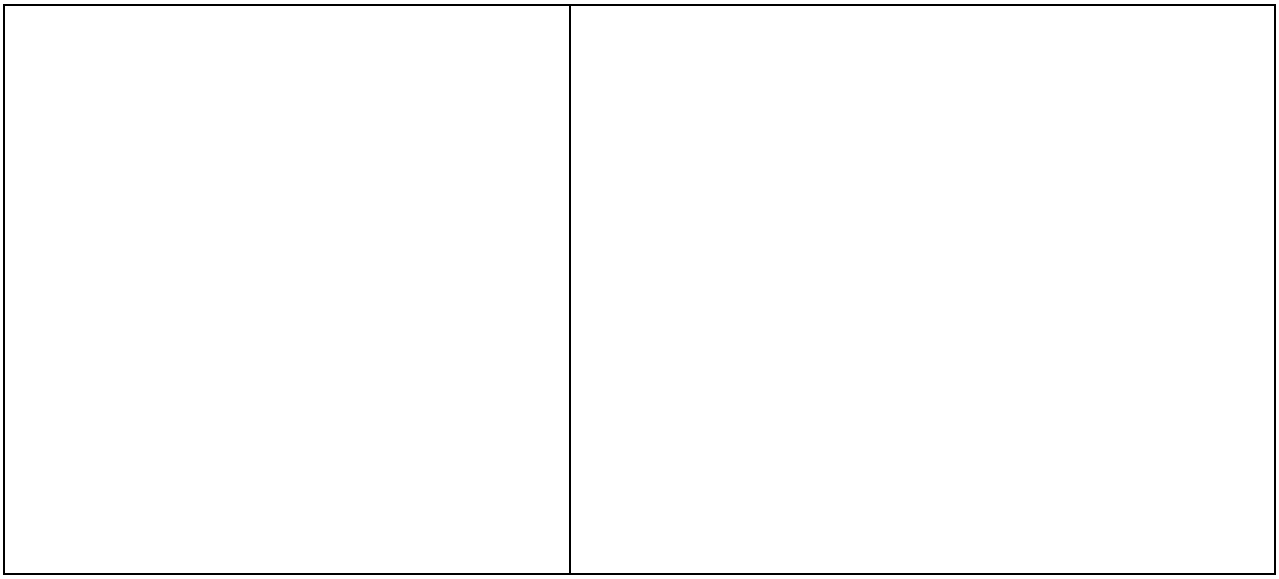

Figure 2. A box plot was generated to understand the distribution of log2fold signals in train_data and identified that the signals ranged between 4 to 12.

The display counts classes by sn's. plot was created using the label and per class and value counts were generated as 0:9, 1:36, 2:24, 3:24, 4:13 (Figure 2a). The value_counts of classes created by sns. plot using the label portrayed that data was imbalanced.

All the classes were annotated by using train_data. iloc [:,-1] which meant all rows except for the last column which was the outcome. The last column [outcome] showed an array of 5 classes [0. ,1. ,2.,3. ,4.]. As outcome has been considered as a label it was necessary to convert floating outcome terms into integers so .as type (‘int’). The last column of both train data and test data was converted to an integer. Lastly, the fitted model was used for predictions, and it returns a history callback object it validates the model in x_test and y_test datasets and keeps track of the accuracy, loss, and other training metrics. Then we predicted the model and five distributed probabilities (one is the lowest probability and the other one is the highest probability obtained by using np. argmax is converted to discrete values. The index with

maximum probability was designated as that class and saved under that, the predicted classes. The process then checks y_test (100 samples) through a confusion matrix by using a confusion matrix (y_hat, y_test). A plot was also created to give a better view of the confusion matrix using sns. heatmap (confusion_matrix(y_test, yhat), annot = True, fmt='0.0f'). Finally, the classification report showed that all classes have good precision, recall, and f-1 scores when compared between Y-predicted (yhat) and y-_test.

2.2. EDA log2 fold change

All 137 rma-gene-ful. chp and. rma-alt-splice-dabg. chp files were analyzed through TAC software v_4.0 and using cut off value more than >4 or<-4, 85 transcripts log2 signal value and log2 fold change (FC) were selected. PP Normal (control) was used as baseline differential log2 FC among various groups like C_vs_T2a_b_c_FC, C_vs_T3a_FC, C_vs_T3b_4FC and C_vs_R_FC were generated, and 85 genes were narrowed down as shown in ‘PCA_fourfold_gene_expression.xlsx' (Supplementary Table 2.). Similar to the CNN-1D model raw data and numpy, pandas, matplotlib. pyplot and seaborn packages were imported for exploratory data analysis EDA as specified in Supplementary data 2). A heatmap was generated using sns.heatmap(train_data.corr() command (Figure 3)

Figure 3 A heatmap was generated using sns. heatmap(train_data.corr() command

Figure 3. A heatmap was created among all four subgroups. ‘C_vs_T2a_b_c_FC ' versus ' C_vs_T3a_FC ' were highly correlated with the values close to +1. While ‘C_vs_R_FC ' was

least correlated with C_vs_T2a_b_c_FC ' (0.096) and ‘C_vs_T3a_FC’ (0.15) using. However, ‘C_vs_T3b_4FC ‘showed a correlation of 0.51. This correlation data was generated by narrowing down 85 genes after analysis of 137 rma-gene-ful. chp and. rma-alt-splice-dabg. chp files through TAC v_4.0 software using cut off of >4/<-4 fold change.

CNN-1D Model was finalized with an accuracy of 0.98 and was saved as Prostrate_Model.hd5 and Prostrate_Model_logs. All 137 samples without labels were considered as test data (Supplementary Table 2) and were processed using the saved Prostrate_Model.hd5 and Prostrate_Model_logs.

3 Results

3.1. Classification of pathological stages using similarity

The study attempted to develop an automated Convolutional Neural Network-1Dimensional (CNN) -1D model for predicting the pathological stages of PCa i.e., pT2a/b/ pT2c, pT3a, pT3b/pT4a, and CRPCs/Hormone resistant (R) as well as prostatectomy Gleason score (GS), the GS-6 and -7(3+4), GS-7(4+3), GS-8/9 by evaluating the log2signal values of 16201 transcripts among tumor samples including both HNPC and CRPCs/R cases. The developed model can predict pT2a/b/ pT2c, pT3a, pT3b/pT4a, and CRPCs/ R with an accuracy of 0.98. Hence, data generated by processing RNA from fresh tumor tissue during prostatectomy and hybridizing on Affymetrix GeneChip HTA v_2.0 can be classified into pT2a/b/ pT2c, pT3a, pT3b/pT4a, and CRPCs/ R with accuracy of 0.98.

3.2. A highly upregulated transcripts in tumors compared to control identified through Transcriptomics Analysis Software v_4.0

All the three genes Alpha-methyl acyl-CoA racemase (AMACR), Prostate cancer gene 3 (PCA3), and Tudor Domain Containing 1 (TDRD1) were highly upregulated in pT2a/b/c and subsequently the upregulation declined in pT3a, pT3b_4 and CRPC/R group (Table 7b). Importantly, constant overexpression of AMACR has been reported in PCa epithelium. Its beta-oxidizes the peroxisomal dietary branched-chained fatty acids (Lin et al., 2012).

PCA3. a spliced, long non-coding RNA explicitly expressed in the PCa gland, and significant over-expression has been observed in cancerous PCa tissues compared to benign prostate tissues (Bussemakers et al., 1999). We also identified that both coding and noncoding transcript levels increased in pT2a_b_c and pT3a groups up to 539 to 332-fold respectively. However, the levels of PCA3 diminished to 35.98 and 9.58 in the pT3b_4 and Resistant

group. Further, overexpression of TDRD1 has been reported in primary PCa tumors (Kacprzyk et al., 2013; Xiao et al., 2016) which corroborates with our study. TDRD1 overexpression is conserved independent of nodal metastasis. TDRD1 is established in PCa cells. TDRD1 also regulates PCa cancer proliferation by associating with the snRNP biogenesis (Kim et al., 2023).

Table 2a

Gene Symbol	**Condition FDR F-Test**	**C_vs_T2a_b_c_FC**	**C_vs_T3a_FC**	**C_vs_T3b_4FC**	**C_vs_R_FC**
AMACR	4.48E-07	-31.48	-29.09	-13.74	-9.07
PCA3	6.55E-14	-272.21	-178.35	-28.2	8.42
PCA3	4.17E-14	-539.89	-332.5	-35.98	9.58
TDRD1	0.0031	-42.72	-21.34	-22.67	-10.27

Table 2b-Highly downregulated transcripts in various stages of PCa

Gene Symbol	Condition FDR F-Test	C_vs_T2a_b_c FC	C_vs_T3a_FC	C_vs_T3b_4FC	C_vs_R_FC
SFN	3.87E-05	5.46	4.9	4.25	23.22
MIR205HG; MIR205	9.85E-09	5.56	10.19	10.77	49.93
MIR205HG	3.95E-08	5.78	11.29	11.51	59.54
TGM4	0.0355	12.2	47.14	49.13	58.87
MYLK	2.31E-08	6.22	9.7	15.65	56.28
PCAT4	0.0131	10.25	13.7	53.87	275.55
OLFM4	7.71E-05	5.23	10.47	21.14	173.25
SLC14A1	2.30E-10	24.31	37.19	37.17	48.51
CD177P1	1.45E-07	7.87	15.13	38.48	89.24
NEFH	0.0002	9	14.23	12.72	66.98
HSD17B13	1.18E-06	14.4	15.42	10.91	18.16

The expected results of the above tables indicate that ‘C_vs_T3b_4FC’ has a correlation of 0.51. This correlation data was generated by narrowing down 85 genes after analysis of 137 rma-gene-ful. chp and. rma-alt-splice-dabg. chp files through TAC v_4.0 software using cut off of >4/<-4-fold change.

CNN-1D Model was finalized with an accuracy of 0.98 and was saved as Prostrate_Model.hd5 and Prostrate_Model_logs. All 137 samples without labels were considered as test data in (Table 2) and were processed using the saved Prostrate_Model.hd5 and Prostrate_Model_logsS.

4. Discussion

The study aimed to build a model to predict the pathological stages of PCa i.e., pT2a/b/c, pT3a, pT3b/pT4a, and CRPCs/R using a dataset generated on Affymetrix GeneChip Human Transcriptome Array 2.0 by Firlej group ((Firlej et al., 2022). Through the analysis, the study developed an automated Convolutional Neural Network-1Dimensional (CNN) -a 1D model which predicts pT2a/b/ pT2c, pT3a, pT3b/pT4a, and CRPCs/ R with an accuracy of 0.98. The only requirement is to isolate 500 ng of RNA from fresh tumor tissue during the prostatectomy process on Affymetrix GeneChip Human Transcriptome Array 2.0 (Singh et al., 2016) and they will be classified into any of these classes pT2a/b/ pT2c, pT3a, pT3b/pT4a, and CRPCs/ R with accuracy of 0.98.

Furthermore, the study also analyzed 137 rma-gene-ful. chp and. rma-alt-splice-dabg. chp files through TAC v_4.0 software. To identify highly differentially expressed transcripts among all pT2a/b/c, pT3a, pT3b/4, and CRPCs/ R stages of PCa 85 genes were narrowed down using cut-off values of more than fourfold and less than minus fourfold differentially expressed transcripts keeping control as a baseline to understand the progression, therapeutic management and aggressiveness of the disease which will lead a path towards Precision Medicine in PCa. The generated heatmap showed a high correlation with the values close to +1 between pT2a_b_c and pT3a. While CRPC/R showed no correlation with pT2a_b_c_FC and ‘pT3a but 50 percent correlation was observed with pT3b_4.

Lastly, during the analysis of more than plus/minus 50-fold of individual differentially regulated transcripts, we were able to clearly understand the therapeutic management of different stages of PCa. The downregulation of TGM4, CD177, and MYLK genes was cold towards immunotherapy and their combinations with other drugs have shown limited

responses in mCRPC and toxicity to normal PCa cells which possess these transcripts (Lopez-Bujanda et al., 2021, Kim et al., 2021, Wang et al., 2022).

5. Summary implications

Using RNA isolated from fresh tumor tissue during prostatectomy of PCa patients treated on Affymetrix GeneChip HTA 2.0 can be processed on our developed automated CNN -1D model which classifies PCa patients into any of these classes pT2a/b/c, pT3a, pT3b/4, and CRPCs/ R with accuracy of 0.97.

Downregulation of TGM4, CD177, and MYLK genes were cold towards immunotherapy and their combinations with other drugs will show limited responses in CRPC and may be toxic to normal PCa cells that possess these transcripts.

Downregulation of PCAT4 may make HDAC inhibitors toxic and non-responsive in PCa.

Highly mutated Tp53 (due to a decrease in SFN expression), OLFM4, and NEFH after restoration might be novel candidate biomarkers for PCa.

overexpressed-TDRD1 may also play as novel candidate biomarkers for PCa

Drug Repurposing: N-methylthiocarbamate, an inhibitor of overexpressed AMACR and docetaxel treatment may reduce PCa cell proliferation.

In conclusion, a Nanostring/Taqman-based 85 gene transcripts panel including TGM4, CD177, MYLK, PCAT4, SFN, OLFM4, NEFH, AMACR, TDRD1 can be made which will be good enough for precision oncology of PCa patients with no overtreatment and toxicity in PCa patients.

This will be a good contribution to the Oncology market worldwide which is accelerating at a CAGR of almost 13.02% with simultaneous 9.2% growth of the Precision Medicine market focused on individual-specific cancer patient treatment over the next 5 years.

References

Ahn, T., Goo, T., Lee, C.-h., Kim, S., Han, K., Park, S., and Park, T. (2018). 'Deep Learning-based Identification of Cancer or Normal Tissue using Gene Expression Data.' *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 1748-1752.

Beer, T. M., Armstrong, A. J., Rathkopf, D., Loriot, Y., Sternberg, C. N., Higano, C. S., Iversen, P., Evans, C. P., Kim, C. S., Kimura, G., Miller, K., Saad, F., Bjartell, A. S., Borre, M., Mulders, P., Tammela, T. L., Parli, T., Sari, S., van Os, Tombal, B (2017). 'Enzalutamide in Men with Chemotherapy-naive Metastatic Castration-resistant Prostate Cancer: Extended Analysis of the Phase 3 PREVAIL Study.' *Eur Urol*, 71(2), 151-154.

Bi, W. L., Hosny, A., Schabath, M. B., Giger, M. L., Birkbak, N. J., Mehrtash, A., Allison, T., Arnaout, O., Abbosh, C., Dunn, I. F., Mak, R. H., Tamimi, R. M., Tempany, C. M.,

Swanton, C., Hoffmann, U., Schwartz, L. H., Gillies, R. J., Huang, R. Y., and Aerts, H. (2019). 'Artificial intelligence in cancer imaging: Clinical challenges and applications.' *CA Cancer J Clin*, 69(2), 127-157.

Brown, L. C., Lu, C., Antonarakis, E. S., Luo, J., and Armstrong, A. J. (2020). 'Androgen receptor variant-driven prostate cancer II: advances in clinical investigation.' *Prostate Cancer Prostatic Dis*, 23(3), 367-380.

Chang, L. W., Hung, S. C., Wang, S. S., Li, J. R., Yang, C. K., Chen, C. S., Ho, H. C., Cheng, C. L., Ou, Y. C., and Chiu, K. Y. (2019). 'Abiraterone Acetate and Enzalutamide: Similar Efficacy in Treating Post Docetaxel Metastatic Castration-resistant Prostate Cancer: Single Center Experience.' *Anticancer Res*, 39(7), 3901-3908.

Chiu, P. K., Shen, X., Wang, G., Ho, C. L., Leung, C. H., Ng, C. F., Choi, K. S., and Teoh, J. Y. (2022). 'Enhancement of prostate cancer diagnosis by machine learning techniques: an algorithm development and validation study.' *Prostate Cancer Prostatic Dis*, 25(4), 672-676.

Fallara, G., Lissbrant, I. F., Styrke, J., Montorsi, F., Garmo, H., and Stattin, P. (2020). *'Observational study on time on treatment with abiraterone and enzalutamide.'* PLoS One, 15(12), e0244462.

Firlej, V., Soyeux, P., Nourieh, M., Huet, E., Semprez, F., Allory, Y., Londono-Vallejo, A., de la Taille, A., Vacherot, F., and Destouches, D. (2022). 'Overexpression of Nucleolin and Associated Genes in Prostate Cancer.' *Int J Mol Sci*, 23(9).

Gupta, S., Gupta, M. K., Shabaz, M., and Sharma, A. (2022). 'Deep learning techniques for cancer classification using microarray gene expression data.' *Front Physio* l, 13, 952709.

Huang, R.H., Hong, Y.K., Du, H., Ke, W.Q., Lin, B.B., Li, Y.L. (2023). 'A machine learning framework develops a DNA replication stress model for predicting clinical outcomes and therapeutic vulnerability in primary prostate cancer.' *J Transl Med*, 21(1), 20.

Huang, S., Yang, J., Fong, S., and Zhao, Q. (2020). 'Artificial intelligence in cancer diagnosis and prognosis: Opportunities and challenges.' *Cancer Lett*, 471, 61-71.

Huang, Z., Johnson, T. S., Han, Z., Helm, B., Cao, S., Zhang, C., Salama, P., Rizkalla, M., Yu, C. Y., Cheng, J., Xiang, S., Zhan, X., Zhang, J., and Huang, K. (2020). 'Deep learning-based cancer survival prognosis from RNA-seq data: approaches and evaluations.' *BMC Med Genomics*, 13(Suppl 5), 41.

Kim, H., Barua, A., Huang, L., Zhou, T., Bolaji, M., Zachariah, S., Mitra, A., Jung, S. Y., He, B., and Feng, Q. (2023). '*The cancer testis antigen TDRD1 regulates prostate cancer proliferation by associating with the snRNP biogenesis machinery.' Oncogene*.

Kim, M. C., Borcherding, N., Ahmed, K. K., Voigt, A. P., Vishwakarma, A., Kolb, R., Kluz, P. N., Pandey, G., De, U., Drashansky, T., Helm, E. Y., Zhang, X., Gibson-Corley, K. N., Klesney-Tait, J., Zhu, Y., Lu, J., Lu, J., Huang, X., Xiang, H., . . . Zhang, W. (2021). '*CD177 modulates the function and homeostasis of tumor-infiltrating regulatory T cells.*' Nat Commun, 12(1), 5764.

Labriola, M. K., Atiq, S., Hirshman, N., and Bitting, R. L. (2021). 'Management of men with metastatic castration-resistant prostate cancer following potent androgen receptor inhibition.' *A review of novel investigational therapies.* Prostate Cancer Prostatic Dis, 24(2), 301-309.

Li, H., Kim, C., Liu, W., Zhu, J., Chin, K., Rodriguez-Canales, J., and Rodgers, G. P. (2020). 'Olfactomedin 4 downregulation is associated with tumor initiation, growth, and progression in human prostate cancer.' *Int J Cancer*, 146(5), 1346-1358.

Lin, P. Y., Cheng, K. L., McGuffin-Cawley, J. D., Shieu, F. S., Samia, A. C., Gupta, S., Cooney, M., Thompson, C. L., and Liu, C. C. (2012). 'Detection of Alpha-Methylacyl-

CoA Racemase (AMACR). A Biomarker of Prostate Cancer, in Patient Blood Samples Using a Nanoparticle Electrochemical Biosensor.' *Biosensors*, 2(4), 377-387.

Liu, W., Lee, H. W., Liu, Y., Wang, R., and Rodgers, G. P. (2010). 'Olfactomedin 4 is a novel target gene of retinoic acids and 5-aza-2'-deoxycytidine involved in human myeloid leukemia cell growth, differentiation, and apoptosis.' Blood, 116(23), 4938-4947.

Lopez-Bujanda, Z. A., Obradovic, A., Nirschl, T. R., Crowley, L., Macedo, R., Papachristodoulou, A., O'Donnell, T., Laserson, U., Zarif, J. C., Reshef, R., Yuan, T., Soni, M. K., Antonarakis, E. S., Haffner, M. C., Larman, H. B., Shen, M. M., Muranski, P., and Drake, C. G. (2021). 'TGM4: an immunogenic prostate-restricted antigen.' J Immuno Cancer, 9(6).

Mottet, N., van den Bergh, R. C. N., Briers, E., Van den Broeck, T., Cumberbatch, M. G., De Santis, M., Fanti, S., Fossati, N., Gandaglia, G., Gillessen, S., Grivas, N., Grummet, J., Henry, A. M., van der Kwast, T. H., Lam, T. B., Lardas, M., Liew, M., Mason, M. D., Moris, L., . . . Cornford, P. (2021). 'EAU-EANM-ESTRO-ESUR-SIOG Guidelines on Prostate Cancer-2020 Update. Part 1. Screening, Diagnosis, and Local Treatment with Curative Intent.' *Eur Urol*, 79(2), 243-262.

Pu, Y. S., Ahn, H., Han, W., Huang, S. P., Wu, H. C., Ma, L., Yamada, S., Suga, K., and Xie, L. P. (2022). 'Enzalutamide in Chemotherapy-Naive Metastatic Castration-Resistant Prostate Cancer*': An Asian Multiregional, Randomized Study. Adv Ther*, 39(6), 2641-2656.

Rebello, R. J., Oing, C., Knudsen, K. E., Loeb, S., Johnson, D. C., Reiter, R. E., Gillessen, S., Van der Kwast, T., and Bristow, R. G. (2021).' *Prostate cancer'*. Nat Rev Dis Primers, 7(1), 9.

Ryan, C. J., Smith, M. R., Fizazi, K., Saad, F., Mulders, P. F., Sternberg, C. N., Miller, K., Logothetis, C. J., Shore, N. D., Small, E. J., Carles, J., Flaig, T. W., Taplin, M. E., Higano, C. S., de Souza, P., de Bono, J. S., Griffin, T. W., De Porre, P., Yu, M. K., . . . Investigators, C.-A.-. (2015). 'Abiraterone acetate plus prednisone versus placebo plus prednisone in chemotherapy-naive men with metastatic castration-resistant prostate cancer (COU-AA-302): final overall survival analysis of a randomized, double-blind, placebo-controlled phase 3 study.' *Lancet Oncol*, 16(2), 152-160.

Singh, N., Sahu, D. K., Mishra, A., Agarwal, P., Goel, M. M., Chandra, A., Singh, S. K., Srivastava, C., Ojha, B. K., Gupta, D. K., and Kant, R. (2016). 'Multiomics approach showing genome-wide copy number alterations and differential gene expression in different types of North-Indian pediatric brain tumors.' *Gene*, 576(2 Pt 2), 734-742.

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., and Bray, F. (2021). '*Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries*.' CA Cancer J Clin, 71(3), 209-249.

Thomas, T. S., and Pachynski, R. K. (2018). 'Treatment of Advanced Prostate Cancer.*' Mo Med*, 115(2), 156-161.

Varghese, B., Chen, F., Hwang, D., Palmer, S. L., De Castro Abreu, A. L., Ukimura, O., Aron, M., Aron, M., Gill, I., Duddalwar, V., and Pandey, G. (2019). 'Objective risk stratification of prostate cancer using machine learning and radiomics applied to multiparametric magnetic resonance images.' Sci Rep, 9(1), 1570.

Varghese, B., Chen, F., Hwang, D., Palmer, S. L., De Castro Abreu, A. L., Ukimura, O., Aron, M., Aron, M., Gill, I., Duddalwar, V., and Pandey, G. (2019). 'Objective risk stratification of prostate cancer using machine learning and radiomics applied to multiparametric magnetic resonance images.' *Sci Rep*, 9(1), 1570

Wang, I., Song, L., Wang, B. Y., Rezazadeh Kalebasty, A., Uchio, E., and Zi, X. (2022). *Prostate cancer immunotherapy. 'A review of recent advancements with novel treatment methods and efficacy.'* Am J Clin Exp Urol, 10(4), 210-233.

Xiao, L., Lanz, R. B., Frolov, A., Castro, P. D., Zhang, Z., Dong, B., Xue, W., Jung, S. Y., Lydon, J. P., Edwards, D. P., Mancini, M. A., Feng, Q., Ittmann, M. M., and He, B. (2016). '*The Germ Cell Gene TDRD1 as an ERG Target Gene and a Novel Prostate Cancer Biomarker'*. Prostate, 76(14), 1271-1284.

Xie, Y., Meng, W. Y., Li, R. Z., Wang, Y. W., Qian, X., Chan, C., Yu, Z. F., Fan, X. X., Pan, H. D., Xie, C., Wu, Q. B., Yan, P. Y., Liu, L., Tang, Y. J., Yao, X. J., Wang, M. F., and Leung, E. L. (2021). 'Early lung cancer diagnostic biomarker discovery by machine learning methods.' *Transl Oncol*, 14(1), 100907.

Yu, S., Tao, J., Dong, B., Fan, Y., Du, H., Deng, H., Cui, J., Hong, G., and Zhang, X. (2021). 'Development and head-to-head comparison of machine-learning models to identify patients requiring prostate biopsy.' *BMC Urol*, 21(1), 80.

Zeng, R., Dong, J. (2021). 'The Hippo Signalling Pathway in Drug Resistance in Cancer.' *Cancers (Basel)*, 16;13(2):318.