Extracellular microvesicle microRNAs, along with imaging metrics, improve detection of aggressive prostate cancer

Kapil K Avasthi¹,
Jung Choi²,
Tetiana Glushko²,
Brandon J Manley³, M.D.,
Alice Yu³, M.D.,
Julio Pow-Sang³ M.D.,
Robert Gatenby², M.D.,
Liang Wang¹, * Ph.D.,
Yoganand Balagurunathan⁴, * Ph.D.

¹Tumor Microenvironment and Metastasis, H Lee Moffitt Cancer Center, Tampa, FL.
²Diagnostic & Interventional Radiology, H Lee Moffitt Cancer Center, Tampa, FL.
³Genitourinary Oncology, H Lee Moffitt Cancer Center, Tampa, FL.
⁴Machine Learning, H Lee Moffitt Cancer Center, Tampa, FL.

*Corresponding Authors
Liang Wang, Ph.D /Yoganand Balagurunathan, Ph.D

Tumor Microenvironment and Metastasis/Machine Learning,
H Lee Moffitt Cancer Center, Tampa, FL.
Liang.Wang@moffitt.org
Yoganand.balagurunathan@moffitt.org

Running title: miRNA and Imaging features in prostate cancer
Keywords: liquid biopsy, miRNA and imaging biomarkers, prostate cancer, radiomics of MRI

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Emails:

Avasthi, Kapil K  kapil.avasthi@moffitt.org
Choi, Jung W  Jung.Choi@moffitt.org
Glushko, Tetiana  Tetiana.Glushko@moffitt.org
Manley, Brandon J  Brandon.Manley@moffitt.org
Yu, Alice  Alice.Yu@moffitt.org
Pow-Sang, Julio M  Julio.Powsang@moffitt.org
Robert Gatenby,  Robert.ganteby@moffitt.org
Wang, Liang  Liang.Wang@moffitt.org
yoga balagurunathan  Yoganand.Balagurunathan@moffitt.org

FUNDING STATEMENT

We would like to acknowledge funding from Cohon’s family donation (to JPS, YB).) and NCI (U01 CA200464 and R37 CA229810 to YB) and NCI (R01CA250018 to LW) which partly supported research activity for the indicated authors.

ETHICS STATEMENT

This study was approved by the Moffitt Cancer Center IRB (protocol number mcc#18104).

Authors Contributions.

KKA, LW: miRNA assay and quantification
JC, TG: Radiological annotation and opinion
YB, JC: Imaging metrics and quantification
YB, KKA, LW: statistical analysis, inference and hypothesis
KKA, LW, YB, BJM, RG: manuscript writing/editing and approval.
ABSTRACT

Prostate cancer is the most commonly diagnosed cancer in men worldwide. Early diagnosis of the disease provides better treatment options for these patients. Magnetic resonance imaging (MRI) provides an overall assessment of prostate disease. Quantitative metrics (radiomics) from the MRI provide a better evaluation of the tumor and have been shown to improve disease detection. Recent studies have demonstrated that plasma extracellular vesicle microRNAs (miRNAs) are functionally linked to cancer progression, metastasis, and aggressiveness. In our study, we analyzed a matched cohort with baseline blood plasma and MRI to access tumor morphology using imaging-based radiomics and cellular characteristics using miRNAs-based transcriptomics. Our findings indicate that the univariate feature-based model with the highest Youden’s index achieved average areas under the receiver operating characteristic curve (AUC) of 0.76, 0.82, and 0.84 for miRNA, MR-T2W, and MR-ADC features, respectively, in identifying clinically aggressive (Gleason grade) disease. The multivariable feature-based model demonstrated an average AUC of 0.88 and 0.95 using combinations of miRNA markers with imaging features in MR-ADC and MR-T2W, respectively. Our study demonstrates combining miRNA markers with MRI-based radiomics improves predictability of clinically aggressive prostate cancer.
INTRODUCTION

Prostate cancer (PCa) is the second most common cause of cancer mortality among men worldwide and represents a significant health burden\(^1\). PCa is a heterogeneous disease that can manifest as either a low-risk, indolent tumor; about 42-66% of patients are estimated to be indolent Pca\(^2\) or as a high-risk, aggressive tumor that may eventually metastasize and become lethal if untreated. The widespread adoption of serum-based prostate-specific antigen (PSA) tests has significantly improved early detection of PCa\(^3\). However, the PSA test lacks specificity, which has led to a higher rate of false detection of tumors\(^4, 5\). Most localized PCa patients with higher Gleason grade tend to prophetically obtain radical prostatectomy (RP) as a curative option. In some of these patients, the disease progresses to present as biochemical recurrence (BCR) with an increased risk of metastasis\(^6-9\). It becomes critical to have reliable biomarkers capable of diagnosing clinically significant diseases early and can distinguish disease progression, which will greatly improve patient outcomes\(^10\).

Due to its ability to assess the whole prostate gland, magnetic resonance imaging (MRI) has been adopted as the primary modality to clinically stage prostate disease\(^11\). The prostate imaging reporting and data system (PIRADS) allows radiological assessment of prostate disease, which has improved standardized reporting but still suffers from inter-reader variability\(^12\). Radiomics has evolved as a methodology to characterize tumor morphology, and these quantitative metrics can prognosticate disease progression in oncological diseases\(^13-15\).

Extracellular vascular microRNAs (miRNAs) are short, non-coding RNA molecules that regulate gene expression post-transcriptionally\(^16\). They play critical roles in various cellular processes and are involved in the pathogenesis of numerous diseases, including cancer\(^16, 17\). miRNAs are typically found inside cells. However, some are shed into circulation in lipid-coated particles known as exosomes\(^18\). Circulatory exosomal miRNAs have been identified as possible disease biomarkers as they are relatively stable in blood and are protected from endogenous RNase activity. Recently, several miRNAs have been implicated as crucial regulators in PCa progression, some targeting oncogenes with an impact on cancer proliferation\(^19-22\). These miRNAs have been shown to target the most
common oncogene pathways like the mTOR pathway\textsuperscript{23} and cell cycle regulation\textsuperscript{24, 25}. These findings underscore the diverse roles of miRNAs in PCa pathogenesis and therapeutic responses.

In this study, we obtained radiomic characterization of abnormal regions using baseline biparametric MRI. We also quantified the exosomal miRNAs in blood plasma in the same cohort of patients. In combination, we showed that these non-invasive complementary assessments (imaging, miRNAs) can predict clinically significant prostate disease at the patient level. We also outline the role of multi-omic features using biparametric MRI (MR-T2w, MR-ADC) features and blood plasma (miRNAs) to provide improved predictability and allow better reproducibility across patients.

**MATERIAL AND METHODS**

**Patient Cohort and Plasma Sample Preparation**

The patient cohort for the study was retrospectively obtained from The Moffitt Cancer Center. Patients were enrolled in the institutional research protocol (Total Cancer Care), which waived additional informed consent for the research study. Our retrospective research protocol allows access to the presented study, approved by the Moffitt Cancer Center/University of South Florida’s Institutional Review Board (IRB). Diagnostic multiparametric Magnetic resonance imaging (MRI) was obtained before treatment or biopsy. We selected patients who had prostate MRIs for this study to get the best characterization of the gland region. The patients clinical record and pathological assessment of the biopsy specimen were obtained from the medical record.

Blood samples (5-10 mL) were collected in EDTA K2 vacutainers from patients diagnosed with PCa (n=48), prior to treatment. Plasma samples were processed by initial centrifugation at 1300× g for 10 minutes at room temperature (RT). The resulting plasma was then transferred to a fresh 1.5 mL centrifuge tube and subjected to a second centrifugation step at 5000× g for 10 minutes at RT to obtain platelet-poor plasma. Aliquots of 250 µL from the processed plasma were quickly preserved at -80°C until
subsequent processing for exosome isolation. We formed patient sub-cohorts with following groups: MR imaging, blood/plasma, matched imaging and blood plasma.

**Prostate Lesion delineation**

The multiparametric MR imaging was assessed by our clinical radiologists (JC and TG), who provided consensus reading for the most aggressive (pathological grade) disease regions, along with glandular boundaries on the prostate MRI (T2W). The annotations were digitally recorded as an RT (radiotherapy format) referenced to T2W and stored on our research PACS (MIM software®). The clinical reports (radiology, pathology) were available for the clinical radiological assessment. We used bi-parametric MR modalities for the study, MR-ADC modality was semi-automatically registered to T2W using intensity-based image registration in Matlab® and the image resolution was remapped to T2W, used as a reference.

**Prostate Imaging and Quantification.**

Baseline bi-parametric MRI on patients with blood specimens were collected before treatment for PCa. The imaging cohort was restricted to the patients who followed the prostate MR image protocol to have a better characterization of prostate glandular anatomy. Our institutional clinical radiologists (JC, TK) reviewed the patient’s imaging (T2W, diffusion-weighted imaging or DWI/ Apparent diffusion coefficient or ADC) and identified the abnormalities across the prostate glandular anatomy. The abnormal regions were digitally recorded spans across the glandular volume (3 Dimensions) in RT (radiation therapy) format and stored on our research Picture Archive Communication System (PACS) (MIM Software Inc.). Institutional in-house radiomics feature extraction (306 features) was used to quantify abnormal regions of interest on the bi-parametric MR imaging (T2, ADC), which complies with the recommendations of the Image Biomarkers Standardization Initiative (IBSI)26. The radiomics toolbox has 306 quantitative features spanning three major categories (size, shape and texture), that was independently extracted in each of the biMRI modalities (MR-T2W, MR-ADC).
Plasma Exosome Isolation

The exosome isolation process began with the thawing of 250 µl plasma aliquots at room temperature (RT), followed by their transfer to 1.5 mL microcentrifuge tubes. These tubes were then centrifuged at 10,000× g for 15 minutes at 4°C to eliminate large vesicles and cellular debris, yielding a supernatant utilized for exosome isolation. The SBI SmartSEC™ Single for EV Isolation™ (System Biosciences, Palo Alto, CA, USA; cat# SSEC200A-1) was employed as a size exclusion chromatography-based approach. The exosomes were eluted using phosphate buffer saline (PBS) and stored at -80°C until miRNA extraction. The plasma samples were randomized before processing, which will mitigate batch effects in our analysis.

miRNA Extraction

The extraction of exosome miRNA was conducted using the miRNeasy (Micro) Kit (Qiagen, Valencia, CA, USA, cat# 217084). Initially, 200 µL of the exosome sample was mixed with 1 mL of QIAzol Lysis Reagent, followed by chloroform addition and centrifugation at 12,000 × g for 15 minutes at 4°C to isolate the RNA-containing aqueous phase. The extracted RNA underwent purification using the RNeasy Mini Elute spin column, involving ethanol washes and specific buffers, and elution with 15 µl of RNase-free water. The concentration of RNA was measured using the QuantiFluor® RNA System (Promega, Madison, WI, USA, Cat#E3310) with Quantus equipment, and the eluted RNA was subsequently stored at -80°C.

Library preparation

miRNA libraries were generated employing the QIAsseq miRNA library kit (Qiagen, Valencia, CA, USA cat#331502) with 5 µl of total RNA utilized for library preparation. The process involved initial ligation of 3’ and 5’ adapters to the miRNAs. Complementary DNA (cDNA) libraries were then constructed via reverse transcription, followed by 22 cycles of PCR amplification and subsequent cleanup of cDNA using QMN beads. The
concentration of the prepared libraries was quantified using the Qubit 2.0 Fluorometer with the Qubit™ dsDNA Quantification Assay Kits (Thermo Fisher Scientific, Middletown, VA, USA, cat#Q32851). Additionally, library quality was assessed using the Agilent High Sensitivity DS1000 method and the Agilent 2200 TapeStation (Agilent, Santa Clara, CA, USA, cat#5067-5585). Subsequently, libraries were pooled in an equimolar ratio based on their molarity, and the weight-to-moles conversion ratio for nucleic acids was determined.

**miRNA-seq and Data Analysis**

The miRNA-seq procedure began with pooling 20 to 24 libraries, following the guidelines outlined in the NextSeq System - Denature and Dilute Libraries Guide. To ensure quality control, 1% PhiX Control v3 was incorporated into all pools as an internal standard. Single-read sequencing was performed with a 75 bp read length using the NextSeq 500 Sequencing System and the NextSeq 500/550 High Output v2.5 kit (75 cycles) (Illumina, San Diego, CA, USA cat#20024906).

Prior to alignment, the sequencing data’s quality control (QC) was executed using FastQC (version 0.11.9). Subsequently, adaptor removal was carried out using cutadapt (version 3.3). The adapter-trimmed small RNA sequencing reads were then mapped against the miRBase database (version 21) utilizing the DNAStar tool (version 3.2). All statistical analyses were conducted within the R environment (version R4.0.3).

**Biological Pathway Enrichment related to miRNAs**

Regulatory targets and functional annotations of microRNAs were identified using TargetScan 27 and miRDB 28. The Database for Annotation Visualization, and Integrated Discovery (DAVID V 6.7) was used to identify functional biological pathways for top miRNAs identified by our analysis. Furthermore, miRanda software was utilized for target prediction of the putative novel microRNA sequences 29, 30.

**Redundancy reduction and Statistical methods**

Coefficient of discrimination (R²) between the features was computed to quantify dependency across the patient samples in our cohort. the metric (R²) was iteratively
computed between all possible features and highly dependent features ($R^2 \geq 0.99$) were flagged. In this dependent group, a representative feature with the highest variability across the patient population was selected, and others were removed. This process was repeated across each sub-cohort and modalities (miRNA, MR-T2W, MR-ADC)\textsuperscript{31}. The process allowed forming a feature set that was uncorrelated (see Table 1). The level of dependency threshold needs to be balanced between removing correlated features and leaving behind those with information.

A logistic regression-based classifier model in univariate and multivariable (up to three dimensions) was then built using uncorrelated features identified in our cohorts. All possible combinations of features were evaluated to find the best feature combination in each cohort, and this was repeated independently across the modalities. In our study, over 4.45 million possible pairs in the miRNA’s cohort, over 708 thousand pairs in the MR-T2W, and 971 thousand in the MR-ADC cohort were evaluated, respectively. The feature pair was sorted based on hold-out (80/20, train/test) test classification accuracy, and estimates were randomly repeated (over 200 times). Combination mixed multimodal features were then formed by selecting the top candidates from each combination (1, 2, and 3 pairs). Sensitivity, specificity, positive predictive value, negative predictive value, and area under receiver operator characteristics were estimated using cross-validation method with average estimates reported. The feature-based models were ranked based on Youden’s index (Sensitivity + Specificity -1) and receiver operator characteristics area under the curve (ROC AUC or AUC)\textsuperscript{32}. A hold-out cross validation approach (80% train, 20% test) was used to estimate the model performance, which was averaged over multiple repeats (over 200), and ensemble test statistics reported.

**RESULTS**

1. **Patient characteristics.**

The study included 48 primary PCa patients with pretreatment blood plasma samples and MR imaging using mixed protocol (pelvic, prostate, abdomen). Of the samples, we converged on 13 patients (18 biopsies) who had prostate MR imaging that followed
standardized prostate imaging protocol (see Table 1). We assessed patients imaging in each of the bi-parametric modalities (miRNA, MRI-T2/ADC) that were matched with plasma-based markers to create subcohorts (miRNA with MR-T2W and miRNA with MR-ADC). We carried out statistical analysis to identify features that discriminate clinically significant PCAs defined by Gleason scores (GS≥ 3+4) across these subgroups, considering them independently.

2. Modality base classifiers.
To identify individual miRNAs and image features that were associated with aggressive PCa, we first performed correlation analysis and removed correlated features (R^2>=0.99) across all possible features in a modality (miRNA, n=48; MR-T2W/ADC, n=18 biopsies). This step removed 6.8%, 46.7%, and 40.8% of the metrics, leaving us with 300, 163, and 181 uncorrelated features for miRNA, MR-T2W, and MR-ADC, modalities respectively. While in the matched cohort (imaging & miRNA, n=13), we had 285 (removed 11.4%), 143 (removed 53.2%), and 166 (removed 40.8%) uncorrelated features for miRNA, MR-T2W, MR-ADC modalities, respectively. We then performed non-parametric test and identified individual features that were statistically significant across indolent and clinically significant patients (Table 1 and Figure 1). We then built classifier models using logistic regression with univariate and multi-variable (2 and 3) features. Predictive ability of these models was assessed based on area under the receiver operator characteristics (AUC) using a cross-validation (hold out) approach. For univariate feature-based model using either miRNA or imaging modalities, we found that miRNAs (R193: miR-151a-5p, R46: miR-93-5p) based model had an average AUC in the range of 0.66-0.76. MR-T2W radiomic features (Laws-features) had an average AUC range from 0.78 to 0.87, while MR-ADC radiomic features (Co-Occurrence, volume, wavelet) showed an average AUC range from 0.78 to 0.84 (see Table 2). An example univariate feature-based classifiers are shown in Figure 3.

3. Multimodal classifier model.
To evaluate if combination of multi-modality features could improve performance of detecting aggressive disease, we built multi-modal predictors by combining miRNAs and
bi-parametric MR features (MR-T2W / ADC) from respective modalities in the matched cohort (n=13). In multimodal feature analysis, we selected the best univariate miRNA’s that had functional relevance to prostate oncology. Using miRNA and MR-T2W radiomics, univariate features had an average AUC of 0.65 to 0.71 and 0.77 to 0.90, respectively. Combination of these two features (miRNA & MR-T2W radiomics) had an average AUC of 0.73 to 0.86. Using two feature-based models (miRNA: miR-7704, miR-151a-5p, T2W: COV, Co-Occurrence) from each of the modalities, seems to moderately complement AUC (average range from 0.79 to 0.96). While using a single feature from miRNA and MR-ADC modalities, a combination (miRNA & MR-ADC) had an average AUC range from 0.73 to 0.75, 0.87 to 0.90 and 0.77 to 0.88, respectively. While using two features from each of the modalities (miRNA: miR-151a-5p, miR-338-3p & MR-ADC: Co-occurrence, Laws features) improved AUC (average range from 0.76 to 0.88) in comparison on using them individually. Importantly, the sensitivity/specificity was higher compared to individual modality-based models (see Table 3 and Figure 4).

4. Gene Ontology and regulatory pathways
After identifying top miRNAs (miR-151a-5p, miR-338-3p, miR-7704, miR-93-5p, and miR-190b-5p) that were predictors of aggressive PCa we used these markers to link regulatory pathways associated with their predicted targets using the following curated databases: TargetScan (targetsan.org) and miRDB target computational prediction software. We also evaluated other open-source pathway miner tools (KEGG33, PANTHER34, and Database for Annotation, Visualization, and Integrated Discovery, DAVID35). The data-mining analysis identified the most relevant pathways using miRNAs as seeds that were most common between TargetScan and miRDB. We found a significant enrichment in the following gene-pathway associations: Pathways in cancer, PI3K, Akt signaling, FoxO signaling and Wnt signaling pathway genes, reported by PANTHER (See Figure 5A). In addition, ras signaling, angiogenesis, FGF signaling, wnt signaling, and PDGF (Platelet-derived growth factor) signaling pathway were the most significant pathways obtained using the KEGG pathway (see Figure 5B).
DISCUSSION

PCa diagnosis and treatment strategies have been improved in the last two decades\(^{36-39}\). Despite these changes, early detection of clinically significant cancers remains challenging\(^{39}\). Since the use of PSA-based tests has resulted in a high level of false diagnosis\(^ {3, 40}\), genomic-based technologies are believed to provide a promising tool in identifying aggressive disease\(^ {37, 41}\). Clinical use of genomic markers to assess metastasis of the disease has improved the management of the disease\(^ {42}\). For example, the use of extracellular miRNAs has evolved in the assay development for disease detection, including PCa\(^ {25, 43}\). Prior studies have shown radiomic features related to histogram intensity and cooccurrences were predictive of aggressive prostate disease, and these metrics have been related to biochemical recurrence\(^ {44, 45}\). Recent work has implicated radiomic metrics related to first-order statistics, texture (laws features, Haralick/cooccurrence) features extracted in MR-T2w, and radiomics features related to texture, edge descriptors (Laws, gradient, Sobel) computed in MR-ADC were associated to aggressive disease grades\(^ {45}\). In comparison, our study finds several imaging features related to intensity and texture-based features (co-occurrence, wavelets, Laws) in T2W/ADC modalities were predictive of aggressive prostate disease (see Tables 2 &3).

Numerous studies have demonstrated that specific miRNAs are differentially expressed in PCa, making them valuable for early diagnosis and disease monitoring. Our study, utilizing a univariate feature-based model with miRNA and imaging modalities, identified miR-151a-5p and miR-93-5p as having the highest AUC (see Table 2). Other possible combinations of miRNA and imaging features are deferred to supplemental section (See Supp. Table ST.1). miR-151a-5p is differentially expressed in PCa, indicating its role in tumor aggressiveness\(^ {46}\). This miRNA is well-known as an oncogene, particularly in colorectal cancer, and is also overexpressed in lung cancer and lymphoblastic leukemia\(^ {47-49}\). Our KEGG analysis predicted that miR-151a-5p indirectly targets the Neuregulin 1 (NRG1) gene through P53 and c-Myc. Recent studies have also shown that the NRG1 gene promotes antiandrogen resistance in PCa\(^ {50}\). The consistent dysregulation of miR-
151a-5p across various cancers highlights its universal role in oncogenic processes. The overexpression of miR-93-5p has been linked to increased migration and invasion in squamous cell carcinoma of the head and neck, suggesting an oncogenic role\textsuperscript{51, 52}. While the exact mechanism of action needs further investigation, elevated levels of miR-93-5p have also been associated with epithelial-mesenchymal transition (EMT), radiotherapy response, and poor prognosis\textsuperscript{52-54}. Additionally, miR-93-5p has been shown to be upregulated in oral cancer\textsuperscript{55, 56}. These findings reinforce the role of miRNAs in regulating PCa, a foundation for further research into their mechanistic roles and therapeutic potential (See Table 4).

In our multi-modal miRNA and bi-parametric MR feature-based analysis, we identified miR-7704, miR-3136-3p, miR-151a-5p, and miR-338-3p as having high AUC values. Based on other studies, miR-7704 emerges as a potential target for aggressive PCa, consistent with its reported roles in ovarian and breast cancers\textsuperscript{57, 58}. In ovarian cancer, miR-7704 is part of a feedback loop with IL2RB and AKT, influencing tumorigenesis and chemoresistance\textsuperscript{57}. This suggests that miR-7704 may play a critical role in cancer progression and underscores its potential as a therapeutic target and prognostic biomarker across different cancer types. In another study, miR-3136-3p was significantly upregulated in high-grade cervical intraepithelial neoplasia in liquid biopsy samples\textsuperscript{59}. miR-338-3p is downregulated in several cancers, including gastric, ovarian, and breast cancers\textsuperscript{60-63}. In PCa cells, overexpression of the miR-338-3p suppresses cell migration and invasion\textsuperscript{64-66}. These miRNA with high AUC in multi-modal miRNA and bi-parametric MR feature-based analysis, indicating their strong diagnostic potential for aggressive prostate cancer (PCa). These miRNAs may serve as valuable therapeutic targets and prognostic biomarkers across various cancer types (see Table 3 & 4 and Figures 3&4).

Although single omics analysis has shown promising in identifying aggressive disease, the PCa is known to be highly heterogenous. One omics data may not capture the complete landscape of PCa biology (add some references). It is believed that multi-omic approach can increase sensitivity of biomarkers. Therefore, we evaluated the multi-omic approach in a matched cohort of pre-treatment blood plasma and MR imaging to identify biomarkers in localized PCa patients. Our data showed that several miRNA and image-
based features are differentially expressed in clinically significant PCa. By combining circulating extracellular transcriptomes and MRI-based radiomes, our multi-omic model improved performance in distinguishing aggressive disease. The integration of these features significantly enhances the accuracy of predicting clinically significant PCa, demonstrating the value of a mixed-modality approach in assessing disease aggressiveness.

**Limitations**

Our study has relatively small sample size, which may affect the generalizability of the findings. Additionally, the reliance on a single institutional cohort may introduce bias and limits the racial cross-sectional nature of the data that could add to biases. We used several mitigation strategies that include cross validation approach. The study also did not account for potential confounding factors such as providers, patient racial/treatment history and genetic variability. Further research with larger, diverse cohorts and longitudinal data is necessary to validate and expand upon these results.

**CONCLUSION**

Our study findings highlight the significant roles of circulating transcriptomics and radiomics in identifying aggressive PCa. By outlining specific miRNAs and MR imaging features, the study enhances our understanding of PCa pathogenesis through improved assessment of morphological characteristics. The combination of prostate miRNAs with imaging metrics offers a non-invasive method for assessing aggressive disease. However, validation of these findings in secondary, independent studies is essential.
# Table Legends

Table 1. Feature that are non-dependent across fluid and imaging modalities.

<table>
<thead>
<tr>
<th>Modalities</th>
<th>Patients Samples (Gleason score)</th>
<th>Number of Features</th>
<th>Uncorrelated Features ($R^2 \geq 0.99$)</th>
<th>Number of Significant Features (Wilcoxon, $p \leq 0.05$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3+3 Vs 3+4 vs ≥4+3</td>
</tr>
<tr>
<td><strong>Individual Modalities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1  miRNA</td>
<td>48</td>
<td>322</td>
<td>300</td>
<td>322</td>
</tr>
<tr>
<td>2  Radiomics on MRI:T2W</td>
<td>18</td>
<td>306</td>
<td>163</td>
<td>306</td>
</tr>
<tr>
<td>3  Radiomics on MRI:ADC</td>
<td>306</td>
<td>181</td>
<td>181</td>
<td>306</td>
</tr>
<tr>
<td><strong>Matched Samples (across modalities)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1  miRNA</td>
<td>13</td>
<td>322</td>
<td>285</td>
<td>322</td>
</tr>
<tr>
<td>2  Radiomics on MRI T2W</td>
<td>(3+3: 7, 3+4: 4)</td>
<td>306</td>
<td>143</td>
<td>306</td>
</tr>
<tr>
<td>3  Radiomics: on MRI:ADC</td>
<td>(3+4: 2)</td>
<td>306</td>
<td>166</td>
<td>306</td>
</tr>
</tbody>
</table>
Table 2. Features that discriminate clinically significant prostate cancer from indolent (3+3) using independent features in these modalities; a) miRNA, in 3+3 Vs ≥3+4 (n=48), b) miRNA, in 3+3 Vs ≥4+3 (n=34), c) MRI T2 W radiomics (n=18) d) MRI ADC radiomics (n=18).

### (a) miRNA – Univariate (n=48) (GS 3+3 Vs ≥3+4)

<table>
<thead>
<tr>
<th>miRNA (1-marker)</th>
<th>E[Sensitivity]/E[Specificity]</th>
<th>E[PPV]/E[NPV]</th>
<th>E[AUC], CI , Std</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 miR-151a-5p</td>
<td>0.798/0.68</td>
<td>0.756/0.726</td>
<td>0.76[0.388,0.956],0.162</td>
</tr>
<tr>
<td>2 miR-338-3p</td>
<td>0.807/0.573</td>
<td>0.696/0.709</td>
<td>0.699[0.213,0.956],0.201</td>
</tr>
<tr>
<td>3 miR-93-5p</td>
<td>0.775/0.547</td>
<td>0.692/0.67</td>
<td>0.767[0.457,0.971],0.15</td>
</tr>
<tr>
<td>4 miR-208a-5p</td>
<td>0.833/0.482</td>
<td>0.671/0.686</td>
<td>0.658[0.2,0.95],0.194</td>
</tr>
<tr>
<td>5 miR-190a-5p</td>
<td>0.842/0.469</td>
<td>0.654/0.687</td>
<td>0.702[0.392,0.968],0.173</td>
</tr>
</tbody>
</table>

### (b) miRNA – Univariate (n=34) (GS 3+3 Vs ≥4+3)

<table>
<thead>
<tr>
<th>miRNA (1-marker)</th>
<th>E[Sensitivity]/E[Specificity]</th>
<th>E[PPV]/E[NPV]</th>
<th>E[AUC], CI , Std</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 miR-93-5p</td>
<td>0.684/0.857</td>
<td>0.762/0.798</td>
<td>0.821[0.3,0.95],0.208</td>
</tr>
<tr>
<td>2 miR-8072</td>
<td>0.631/0.855</td>
<td>0.74/0.767</td>
<td>0.776[0.05,0.95],0.237</td>
</tr>
<tr>
<td>3 miR-328-3p</td>
<td>0.589/0.861</td>
<td>0.727/0.73</td>
<td>0.798[0.417,0.958],0.178</td>
</tr>
<tr>
<td>4 miR-1469</td>
<td>0.574/0.845</td>
<td>0.652/0.772</td>
<td>0.772[0.333,0.958],0.183</td>
</tr>
<tr>
<td>5 miR-5583-5p</td>
<td>0.539/0.821</td>
<td>0.645/0.728</td>
<td>0.786[0.25,0.958],0.202</td>
</tr>
</tbody>
</table>

### (c) Univariate Imaging T2w (n=18) (GS 3+3 Vs ≥3+4)

<table>
<thead>
<tr>
<th>T2 (Radiomics)</th>
<th>E[Sensitivity]/E[Specificity]</th>
<th>E[PPV]/E[NPV]</th>
<th>E[AUC], CI , Std</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 3D_Laws_features_L5_E5_S5</td>
<td>0.588/0.882</td>
<td>0.627/0.748</td>
<td>0.83[0.05,0.95],0.267</td>
</tr>
<tr>
<td>2 3D_Laws_features_L5_E5_W5</td>
<td>0.522/0.897</td>
<td>0.645/0.713</td>
<td>0.871[0.325,0.963],0.192</td>
</tr>
<tr>
<td>3 3D_Laws_features_L5_S5_S5</td>
<td>0.575/0.802</td>
<td>0.587/0.712</td>
<td>0.827[0.05,0.95],0.264</td>
</tr>
<tr>
<td></td>
<td>ADC (Radiomics)</td>
<td>E[Sensitivity]/E[Specificity]</td>
<td>E[PPV]/E[NPV]</td>
</tr>
<tr>
<td>---</td>
<td>-----------------------------------------------------</td>
<td>------------------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>1</td>
<td>Volume_density_minimum_volume_enclosing_ellipsoid</td>
<td>0.72/0.832</td>
<td>0.702/0.808</td>
</tr>
<tr>
<td>2</td>
<td>Maximum_histogram_gradient_grey_level</td>
<td>0.413/1</td>
<td>0.61/0.708</td>
</tr>
<tr>
<td>3</td>
<td>GLSZM_Small_zone_high_grey_level_emphasis</td>
<td>0.582/0.81</td>
<td>0.663/0.693</td>
</tr>
<tr>
<td>4</td>
<td>avgCoocurrence_Difference_entropy</td>
<td>0.605/0.758</td>
<td>0.546/0.759</td>
</tr>
<tr>
<td>5</td>
<td>3D_Wavelet_P1_L2_C12</td>
<td>0.648/0.71</td>
<td>0.576/0.714</td>
</tr>
</tbody>
</table>

d) Univariate Imaging ADC (n=18) (GS 3+3 Vs ≥3+4)
Table 3. Multimodal features based (miRNA and imaging) model to discriminate clinically significant prostate cancer (≥3+4) from indolent (3+3) in a matched cohort of patients (n=13). (A1&2) extracellular exosomal miRNA with MRI T2W. (B-1&2) extracellular exosomal miRNA with MRI ADC features.

<table>
<thead>
<tr>
<th>(A1) Combined: miRNA + Img (T2W): (GS 3+3 Vs ≥3+4): 1-dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>miRNA</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(A2) Combined: miRNA + Img (T2W): (GS 3+3 Vs ≥3+4): 2-dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>miRNA</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(B1) Combined: miRNA + Img (ADC): (GS 3+3 Vs ≥3+4): 1-dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>miRNA</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td><strong>1</strong></td>
</tr>
<tr>
<td><strong>2</strong></td>
</tr>
</tbody>
</table>

*(B2) Combined: miRNA + Img (ADC): (GS 3+3 Vs ≥3+4): 2-dimension*

<table>
<thead>
<tr>
<th></th>
<th>miRNA (2-dim)</th>
<th>Imaging (ADC) – (2-dim)</th>
<th>E[AUC], CI , Std</th>
<th>miRNA</th>
<th>Img: ADC</th>
<th>Combined (miRNA, ADC)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1</strong></td>
<td>miR-190b-5p; miR-106b-3p</td>
<td>avgCoocurrence_Difference_entropy; avg_3D_LGRE_(Low grey level run emphasis)</td>
<td>0.843[0.05, 0.95], 0.27 2</td>
<td>0.88[0.05, 0.95], 0.276</td>
<td>0.88[0.4, 0.95], 0.237</td>
<td></td>
</tr>
<tr>
<td><strong>2</strong></td>
<td>miR-151a-5p; miR-338-3p</td>
<td>3D_Laws_features_E5_E5_W5; 3D_Laws_features_E5_R5_E5</td>
<td>0.6[0.05, 0.95], 0.381</td>
<td>0.898[0.05, 0.95], 0.256</td>
<td>0.76[0.2, 0.95], 0.261</td>
<td></td>
</tr>
</tbody>
</table>
### Table 4. Predictive biomarkers relationship to biological pathways

<table>
<thead>
<tr>
<th>Pathways</th>
<th>Biomarkers</th>
<th>Target Gene</th>
<th>Biological Function</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>miRNA's</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>miR-151a-5p, miR-338-3p</td>
<td>Neuregulin 1 (c-Myc, p53)</td>
<td>Promotes proliferation and metastasis</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Inhibits Proliferation and Promotes Apoptosis</td>
</tr>
<tr>
<td>2</td>
<td>miR-93-5p, miR-208a-5p, miR-190a-5p</td>
<td>TGF-β1/Smad3, N/A, PHLPP1</td>
<td>Suppressing proliferation and invasion of PCa cells&lt;sup&gt;67&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Promoting migration and invasion&lt;sup&gt;68&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Radiomics (MRI)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>T2W: Volume_at_intensity_fraction; Weighted_CoM; avgCoocurrence_Autocorrelation; 3D_Wavelet_P1_L2_C12; GLSZM_High_grey_level_zone_emphasis; AvgCoocurrence-Difference_entropy; Volume_density_minimum; Volume_enclosing_ellipsoid</td>
<td>Texture, Morphology related to heterogeneity</td>
<td>Disease heterogeneity, proliferation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>67</sup> Weng et al. 2012, J Clin Oncol
<sup>68</sup> Weng et al. 2011, J Clin Oncol
Figure Legends:

**Figure 1.** Process flow to identify discriminate features from multimodalities, plasma-based miRNA and MR image based radiomics (T2W, ADC) to discriminate aggressive grade prostate disease.

**Figure 2.** Feature distribution that differentiates clinically significant (≥3+4) from indolent (3+3) for a) individual cohorts (miRNA, MR T2w, MR ADC), b) combined cohort (miRNA, MR T2w, MR ADC).

**Figure 3.** Multi-feature scatter plot to show spread of aggressive from indolent grade prostate cancer along with discrimination boundary A) using miRNA-based features (miR-93-5p, miR-151a-5p) b) MR-T2w features (3-features).

**Figure 4.** Receiver operating characteristic curve (ROC) in using top predictors to discriminate clinically significant (Gleason 3+4) from indolent grade prostate cancers A) miRNA (hsa-miR-7704) with T2W radiomics (univariate), C) miRNA (has-miR-338-3p) with ADC radiomics (univariate).

**Figure 5.** Regulators pathways enriched for miRNAs (miR-151a-5p, miR-338-3p, miR-7704, miR-93-5p, miR-190b-5p), that was identified in our predictive analysis. Pathway tools using; a) GO functional pathway (Pather.org), b) p52 pathway enriched, seeded with two miRNAs (miR-151a-5p, miR-338-3p) using (GeneGo®).
REFERENCES


