

Article

# Deep Learning with Quantitative Features of Magnetic Resonance Images to Predict Biochemical Recurrence of Radical Prostatectomy: A Multi-Center Study

Ye Yan <sup>1,†</sup>, Lizhi Shao <sup>2,3,†</sup>, Zhenyu Liu <sup>2,4,5</sup>, Wei He <sup>6</sup>, Guanyu Yang <sup>3</sup>, Jiangang Liu <sup>7,8</sup>, Haizhui Xia <sup>1</sup>, Yuting Zhang <sup>1</sup>, Huiying Chen <sup>6</sup>, Cheng Liu <sup>1</sup>, Min Lu <sup>9</sup>, Lulin Ma <sup>1</sup>, Kai Sun <sup>2</sup>, Xuezhi Zhou <sup>2</sup>, Xiongjun Ye <sup>10</sup>, Lei Wang <sup>11</sup>, Jie Tian <sup>2,7,8,\*</sup> and Jian Lu <sup>1,\*</sup>

<sup>1</sup> Department of Urology, Peking University Third Hospital, Peking University, Beijing 100191, China; yanye@bjmu.edu.cn (Y.Y.); haizhuixia@bjmu.edu.cn (H.X.); zyt100582@163.com (Y.Z.); chengliu@bjmu.edu.cn (C.L.); malulin@medmail.com.cn (L.M)

<sup>2</sup> CAS Key Laboratory of Molecular Imaging, Beijing Key Laboratory of Molecular Imaging, the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China; lz\_shao@seu.edu.cn (L.S.); zhenyu.liu@ia.ac.cn (Z.L.); 15120478@bjtu.edu.cn (K.S.); zxxidian@163.com (X.Z.)

<sup>3</sup> School of Computer Science and Engineering, Southeast University, Nanjing 210096, China; yang.list@seu.edu.cn

<sup>4</sup> CAS Center for Excellence in Brain Science and Intelligence Technology, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

<sup>5</sup> School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100080, China

<sup>6</sup> Department of Radiology, Peking University Third Hospital, Peking University, Beijing 100191, China; heweiwh@126.com (W.H.); hychen@bjmu.edu.cn (H.C.)

<sup>7</sup> Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, School of Engineering Medicine, Beihang University, Beijing 100191, China; jgliu@buaa.edu.cn

<sup>8</sup> Ministry of Industry and Information Technology of the People's Republic of China, Key Laboratory of Big Data-Based Precision Medicine (Beihang University), Beijing 100191, China

<sup>9</sup> Department of Pathology, Peking University Third Hospital, Peking University, Beijing 100191, China; lumin@bjmu.edu.cn

<sup>10</sup> Urology and Lithotripsy Center, Peking University People's Hospital, Peking University, Beijing 100044, China; yexiongjun@bjmu.edu.cn

<sup>11</sup> Department of Urology, Beijing Friendship Hospital, Capital Medical University, Beijing 100050, China; sclare@163.com

\* Correspondence: jie.tian@ia.ac.cn (J.T.); lujian@bjmu.edu.cn (J.L.)

† These authors contributed equally to this work and should be considered as co-first authors.

**Citation:** Yan, Y.; Shao, L.; Liu, Z.; He, W.; Yang, G.; Liu, J.; Xia, H.; Zhang, Y.; Chen, H.; Liu, C.; et al. Deep Learning with Quantitative Features of Magnetic Resonance Images to Predict Biochemical Recurrence of Radical Prostatectomy: A Multi-Center Study. *Cancers* **2021**, *13*, 3098. <https://doi.org/10.3390/cancers13123098>

Academic Editor: Andreas Stadlbauer

Received: 19 March 2021

Accepted: 8 June 2021

Published: 21 June 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

**Simple Summary:** Biochemical recurrence after radical prostatectomy is vitally important for long-term oncological control and subsequent treatment of these patients. We applied radiomic technique to extract features from MR images of prostate cancer patients, and used deep learning algorithm to establish a predictive model for biochemical recurrence with high accuracy. The model was validated in 2 independent cohorts with superior predictive value than traditional stratification systems. With the aid of this model, we are able to distinguish patients with higher risk of developing biochemical recurrence at early stage, thus providing a window to initiate neoadjuvant or adjuvant therapies for prostate cancer patients.

**Abstract:** Biochemical recurrence (BCR) occurs in up to 27% of patients after radical prostatectomy (RP) and often compromises oncologic survival. To determine whether imaging signatures on clinical prostate magnetic resonance imaging (MRI) could noninvasively characterize biochemical recurrence and optimize treatment. We retrospectively enrolled 485 patients underwent RP from 2010 to 2017 in three institutions. Quantitative and interpretable features were extracted from T2 delineated tumors. Deep learning-based survival analysis was then applied to develop the deep-radiomic signature (DRS-BCR). The model's performance was further evaluated, in comparison with conventional clinical models. The model achieved C-index of 0.802 in both primary and validating cohorts,

outweighed the CAPRA-S score (0.677), NCCN model (0.586) and Gleason grade group systems (0.583). With application analysis, DRS-BCR model can significantly reduce false-positive predictions, so that nearly one-third of patients could benefit from the model by avoiding overtreatments. The deep learning-based survival analysis assisted quantitative image features from MRI performed well in prediction for BCR and has significant potential in optimizing systemic neoadjuvant or adjuvant therapies for prostate cancer patients.

**Keywords:** prostate cancer; biochemical recurrence; survival prediction; deep learning; MRI

## 1. Introduction

Prostate cancer (PCa) is the most common cancer among men and the second leading cause of death for men worldwide [1]. Radical prostatectomy (RP) is one option of the multimodal approaches for organ-confined or locally advanced PCa. Biochemical recurrence (BCR) after RP is known to harbor more advanced or aggressive disease. The 10-year estimated BCR rate after RP has been reported to be up to 27% [2]. More than two-thirds of BCR cases develop in the first two years after surgery [3]. BCR is known to be a surrogate of local recurrence, distant metastasis, and cancer-specific death [4]. Therefore, early identification of patients who are predisposed to developing BCR could provide a window for early intervention with systemic or local therapies.

In fact, the natural history of biochemical relapse is heterogeneous. It is difficult to stratify patients according to conventional TNM system, several clinical predictive models were gradually developed for BCR prediction. Among which, the Cancer of the Prostate Risk Assessment (CAPRA/CAPRA-S) score, National Comprehensive Cancer Network (NCCN) model and Gleason grade group (GG-RP) system were three kinds of most widely adopted models [5]. According to these systems, patients were assigned into low-, intermediate- and high-risk groups, while patients in the middle group would face controversial prognosis and unclear therapeutic guidance. Besides, these models only incorporate pure clinical factors, neglecting more comprehensive information obtained from imaging and genetic data.

MRI has been routinely performed as one component of the multi-modal diagnostic procedure of prostate cancer. MRI can overcome the biopsy bias in a non-invasive fashion and better highlight tumoral heterogeneity. Adoption of the PI-RADs 2.1 protocol led to a sensitivity of 79% in distinguishing benign lesions of malignancies [6]. However, PI-RADs system demonstrated limited potentials for oncologic prognoses such as BCR or overall survival estimation. Radiomics has emerged as an approach converting conventional images into high volume quantitative features [7]. The radiomics may not only analyze anatomical information of micro structural but also information representing the underlying pathophysiologic process, which may be correlated with oncological prognosis. Thus, radiomics has successfully introduced to discriminate benign and malignant lesions [8], predict aggressiveness [9] and estimate Gleason scores [10]. However, there were still few studies focusing on BCR prediction for patients after radical prostatectomy with relatively large amount datasets [11–13].

In this study, based on pathological ground truth registration annotations, we developed a novel model for BCR prediction, combing quantitative features of magnetic resonance imaging (MRI) and deep learning-based survival analysis. We then validated the model and calculated its incremental predictive value in comparison with several acknowledgeable nomograms in a multicenter dataset.

## 2. Materials and Methods

This study was approved by the institutional review board of the Peking University Third Hospital Medical Science Research ethics committee with a waiver of informed consent and compliant with the principles in the Declaration of Helsinki (S2019326).

### 2.1. Study Design

A total of 584 consecutive patients diagnosed with prostate cancer were included from three institutions in Beijing. All patients were enrolled with strict inclusion and exclusion criteria. The inclusion criteria were shown as follows: (a) primary prostate adenocarcinoma confirmed by radical prostatectomy; (b) locally or locally advanced disease according to the 8th edition of the AJCC Staging Criteria; (c) No neoadjuvant androgen deprivation treatment (ADT) before surgery; (d) had documented BCR or, (e) did not have BCR but followed over 3 years. BCR was defined as two consecutive total PSA readings  $>0.2$  ng/mL after RP [14]; (f) Radical prostatectomy was performed. The exclusion criteria were as follows: (a) patients with neoadjuvant ADT before surgery; (b) patients with  $<3$  years of follow-up data without BCR status; (c) patients with incomplete clinical information; (d) patients with distant metastasis; (e) patients with other pathology types or mixed pathology types; (f) patients with postoperative PSA nadir  $>0.1$  ng/mL within 3 months after RP. According to the criteria, 485 patients were enrolled in the final analysis (Table 1). Patients who were free from BCR were censored at the last follow-up.

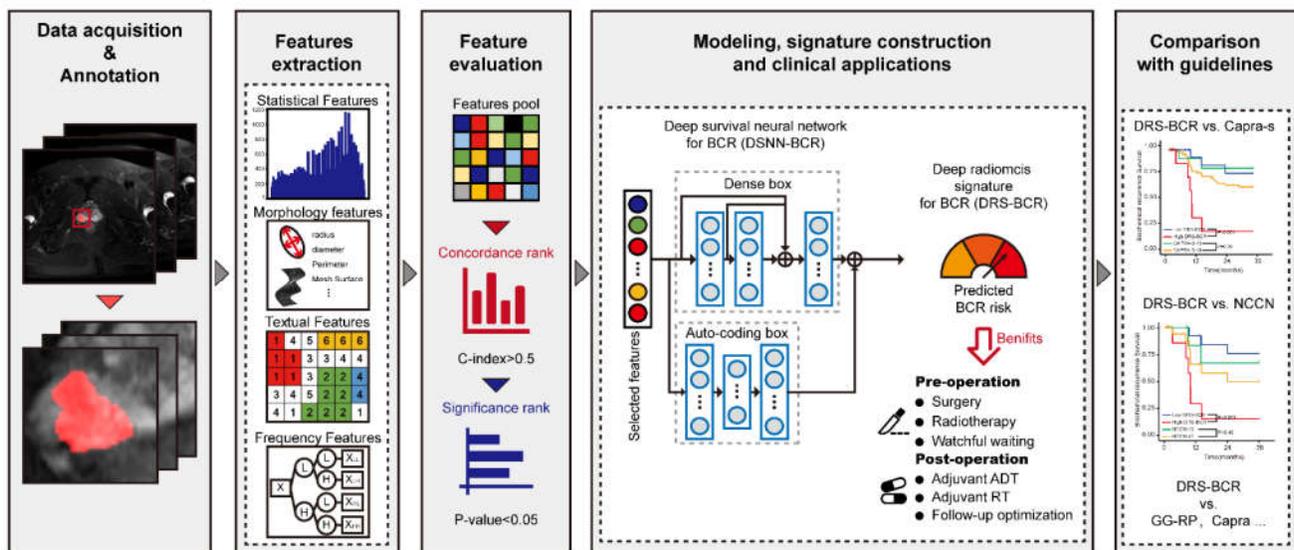
**Table 1.** Performance of clinical models and deep radiomic models of BCR prediction by univariate cox proportional hazards regression.

	PC (PUH)	VC (BJFH)	VC2 (PUPH)	<i>p</i>
	<i>n</i> = 368	<i>n</i> = 34	<i>n</i> = 83	
No. of BCR event (%)	99 (27.1)	16 (47.1)	31 (37.3)	0.016
Age (Median)	70	68.5	68.0	0.275
PSA (Median)	10.7	9.89	13.4	0.883
GG-NB				0.125
1	138	13	29	
2	114	10	23	
3	63	6	14	
4	89	3	11	
5	81	2	6	
cT				0.436
2	221	15	44	
3	264	19	39	
4	9			
pT				0.687
1	2		1	
2	303	22	57	
3	165	12	22	
4	15		3	
GG-RP				0.173
1	82	4	13	
2	127	10	28	
3	93	8	18	
4	65	9	8	
5	118	3	16	
SM				0.172
Positive	170	12	38	
Negative	315	21	45	
EPE				0.639
Positive	171	11	25	
Negative	314	23	58	
SVI				0.940
Positive	67	4	11	
Negative	418	30	72	
CAPRA				0.248

Low	53	9	10
Intermediate	168	12	34
High	148	13	39

Note: BCR: biochemical recurrence; PC, primary cohort; VC, validation cohort; PUTH, Peking University Third Hospital; BJFH, Beijing Friendship Hospital; PUPH, Peking University People’s Hospital; PSA, prostate specific antigen; GG-NB, Gleason grade group of needle biopsy; cT, clinical T stage; pT, pathological T stage; GG-RP, Gleason grade group of radical prostatectomy; SM, surgical margin; EPE, extracapsular extension; SVI, seminal vesicle invasion; CAPRA, Cancer of Prostate Risk Assessment Score.

The patients were divided into three cohorts, the primary cohort (PC:  $n = 368$ , from Peking University Third Hospital, Beijing, China) and other two external validation cohorts (VC1:  $n = 34$ , from Beijing Friendship Hospital; VC2:  $n = 83$ , from Peking University People’s Hospital). Conceptual workflow was demonstrated in Figure 1. The first step included data acquisition of images and tumor segmentation. Then, we extracted radiomic features from tumor area for tumor heterogeneity. Thirdly, feature evaluation based on rank is used to roughly remove features that are of little value in modeling based on statistical approaches. Finally, prognosis prediction of BCR was constructed and validated in our multi-center dataset. Comparison with guidelines was also supplemented. Details of patient recruitment and study design were shown in Supplementary Figure S1.



**Figure 1.** The workflow of radiomics-based model for data preparation, modeling, signature construction, predicting BCR-free survival, and clinical applications. DSNN: deep survival neural network; BCR: biochemical recurrence; DRS: deep radiomic signature; CAPRA: Cancer of Prostate Risk Assessment; NCCN: National Comprehensive Cancer Network; GG: Gleason grade group system.

### 2.2. MRI Protocol

All patients received 3T MRI examination before systematic transrectal needle biopsy. All MR images were obtained with T1W, T2W, DWI and Apparent diffusion coefficient (ADC) maps. Only DICOM data of T2WI were used for analysis in this study. Details of scanning parameters were shown in Table S1.

### 2.3. Annotations

Pathological hematoxylin-eosin sections of each patient from radical prostatectomy specimens were scanned at 40× magnification, converting to computational pathological sections (NanoZoomer S360, HAMAMATSU, Hamamatsu City, Japan). First, one pathologist of urology expertise delineated the lesions that were responsible for diagnosis

on each section. Second, this pathologist and one urological radiologist together recognized and delineated lesions on MRI, which were correlated to pathological whole-mount slides. The tumor regions of interest (ROIs) were delineated using ITK-SNAP software ([www.itksnap.org](http://www.itksnap.org), April 27, 2018). Detailed examples of annotations were demonstrated in Figure S2.

#### 2.4. Radiomic Features of Mp-MRI Extraction

Standard normal distribution of image intensities was obtained through Z-score normalization. Subsequently, 702 quantitative features extracted from the T2WIs of individual patients were calculated to characterize intratumoral heterogeneity and complexity. Five common groups of features (first-order, shape, texture, wavelet and Laplacian of Gaussian filter) were extracted by using “Pyradiomics” (Version 2.1.1, <https://github.com/Radiomics/pyradiomics>, April 14, 2021) [15]. Details of the features were summarized in Table S5.

#### 2.5. Deep Learning-Based Survival Analysis for Signature Construction to BCR Survival

In this study, we developed a deep survival neural network for BCR prediction and constructed a signature named deep radiomic signature for BCR (DRS-BCR) to predict BCR probability. After radiomic feature extraction, we firstly evaluated all the radiomic features in the PC by univariate analysis and recorded the concordance (C-index) and  $p$ -value ( $p$ ) between each feature and BCR. Features with predictive potential (C-index >0.5 and  $p < 0.05$ ) were selected from the features pool to eliminate redundant information. During the modeling phase, we used a new design of neural network structure to describe the correlation between image features and BCR-free survival, which contained a dense box to gather information of multi-level abstraction [16], and an auto-coding box to generate sparse features [17]. Three hidden layers with forty-eight neurons were used for the dense box. Two hidden layers of forty-eight neurons and a hidden layer with twenty-four neurons were used in the auto-coding box. The dropout and early stop strategy were employed in the training process to mitigate overfitting. The loss function of the model was deep survival loss [18], in which the optimizer was Adam [19], and the training batch was the total amount of data in the PC. Finally, the output of the neural network was the risk of BCR for a patient, namely DRS-BCR.

To evaluate the interaction between clinical indicators and our proposed DRS-BCR, important clinical factors with prognosis power and DRS-BCR were added in a Cox proportional-hazards regression model (CPH) for BCR survival. Comparisons of DRS-BCR were evaluated in all cohorts, including combination model and clinical factors (Supplementary Material, Supplementary Information I and II).

#### 2.6. Statistics

We compared the patients with and without BCR using the t-test for continuous variables and the chi-square test or Fisher’s test for categorical variables. The C-index and 95% confidence interval (CI) were calculated, in order to evaluate the performance of the BCR-free survival model. The Kaplan–Meier (K–M) method and log-rank test were used to estimate BCR-free survival. The time-dependent analysis (including ROC curves, AUC, sensitivity, and specificity) was performed to evaluate accuracy of predicting BCR. The cutoff in the time-dependent analysis was selected using criteria based on Youden Index. The Wilcoxon signed rank test was used to compare the C-index distributions of different models. All packages were based on R software (version 3.1.0). A two-sided  $p$  value < 0.05 was considered significant.

### 3. Results

#### 3.1. Patient Characteristics

A total of 485 patients were finally enrolled in the study, with a median age of 69.86 [95% CI, 62.79–76.93] years old. The median BCR-free survival was 57.7 months. There were 81 (22.0%), 9 (26.5%) and 23 (27.7%) patients with 3-year BCR in the PC, VC1 and VC2, respectively, which was balanced ( $p = 0.479$ ), detail information is shown in Table 1.

To assess the independent prognostic power of each clinical parameter, we grouped them as preoperative clinical parameters (PSA, cT stage, GG-NB, PPB) and postoperative parameters (GG-RP, surgical margin, extracapsular extension, seminal vesicle invasion) in a univariable cox regression model. The parameters were normalized by CAPRA (cT stage, GG-NB, PPB) or CAPRA-S (PSA, GG-RP, surgical margin, extracapsular extension, seminal vesicle) criteria. The results of the univariate analysis are shown in Table S2.

We integrated all significant clinical variables ( $p < 0.05$ ), and grouped them by the acquisition time (pre-operative, post-operative, a combination of both pre- and post-operative), and built three clinical prognostic models for BCR prediction, respectively (Tables S3–S5). We termed the clinical variables as clinical signature (CS). The combination of pre- and post-operative clinical variables (CS-combine) yielded C-index [95% CI] of 0.693 [0.634–0.752] in PC, 0.651 [0.471–0.831] in VC1 and 0.641 [0.529–0.753] in VC2, which is the highest among the three.

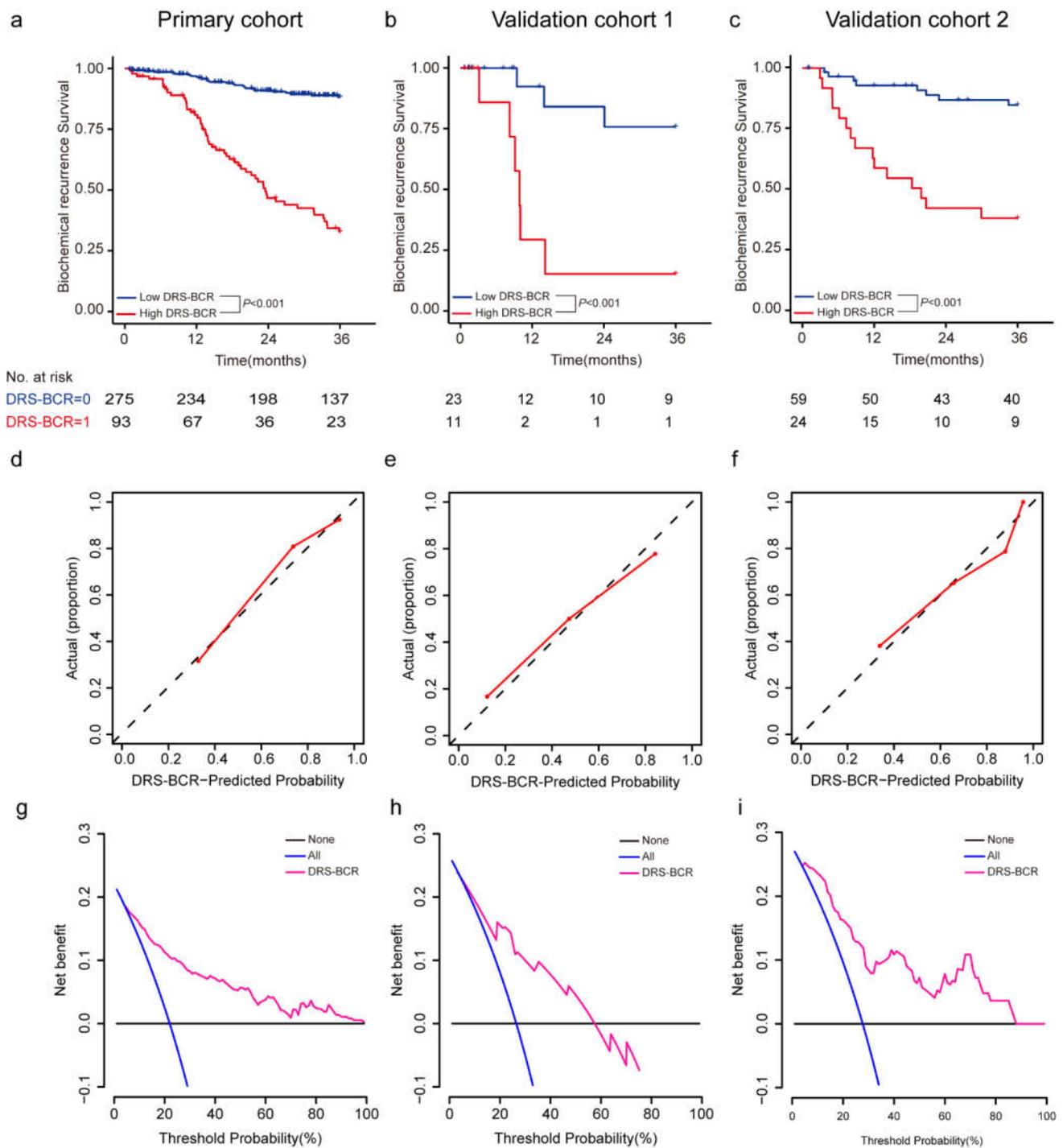
#### 3.2. Evaluation of Quantitative Features of Mp-MRI for BCR-Free Survival

The radiomic pool of 702 features was evaluated by univariate analysis of BCR-free survival. From which, 155 features with both concordance and significance were finally selected for modeling. Among these, 125 (80.6%) of them were in texture category, 22 (14.2%) were in first-order category, and 8 (5.16%) were in shapes related category (Table S6). Features with the highest concordance and significance were the intensity of gray value (first-order) from the original imaging.

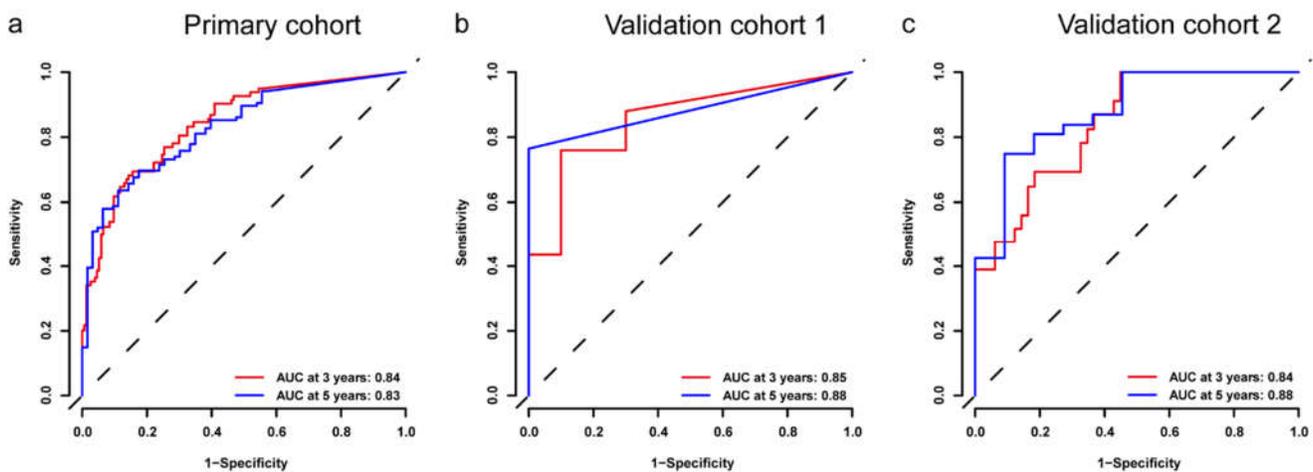
#### 3.3. Assessment of DRS-BCR for Predicting BCR-Free Survival

The DRS-BCR was generated from a deep survival radiomic neural network (DSNN) to predict BCR-free survival, which yielded C-index of 0.802 [95% CI, 0.758–0.846] in PC, 0.811 [95% CI, 0.722–0.9] in VC1 and 0.794 [95% CI, 0.718–0.87] in VC2. Time-dependent ROC curves were then applied to evaluate 3-year and 5-year BCR detection rates (Figure 2). The DRS-BCR of 3-year BCR yielded AUCs of 0.84 in PC, 0.85 in the VC1, and 0.84 in the VC2, respectively (Figure 3). The DRS-BCR of 5-year BCR yielded AUCs of 0.83 in PC, 0.88 in the VC1, and 0.88 in the VC2, respectively (Figure S3). The HR of DRS-BCR was 1.705 [95% CI, 1.531–1.893] ( $p < 0.001$ ) in univariate analysis.

Next, we performed K–M analysis of 3-year BCR-free survival in the PC, VC1, and VC2 (Figure 2a–c). The K–M curves of cohorts all revealed significant differences by log-rank test between groups of high and low risk ( $p < 0.001$ ). The calibration curve of the radiomics model estimated the probability of 3-year BCR, which demonstrated good agreement in the primary cohorts (Figure 2d). The Hosmer–Lemeshow test yielded a non-significant statistic ( $p = 0.657$ ), suggesting no departure from the perfect fit. Good performance was also observed for the probability of 3-year BCR in all validation cohorts ( $p = 0.419$ ,  $p = 0.583$ ) (Figure 2e,f). The decision curve showed relatively good performance for the radiomic signature as well (Figure 2g–i). The results of incorporating DRS-BCR with clinical variables were evaluated in Supplementary Information III (Table S7).



**Figure 2.** Performance of DRS-BCR for predicting 3-year BCR-free survival. (a–c) The Kaplan–Meier (K–M) curves of DRS-BCR-free survival for BCR-free survival within three years in the primary cohort, validation cohort 1, and validation cohort 2, respectively. The *p* values were calculated by log-rank test between subgroup with high-risk and low-risk, and significant discrimination was revealed by *p* values less than 0.05. (d–f) The calibration curves for 3-year BCR. (g–i) Clinical benefits by decision curve analysis for 3-year BCR. DRS: deep radiomic signature; BCR: biochemical recurrence.



**Figure 3.** The time-dependence receiver operating characteristic curve (ROC) of DRS-BCR in the primary cohort (a), validation cohort 1 (b), and validation cohort 2 (c). DRS: deep radiomic signature; AUC: area under the curve.

### 3.4. Comparison between DRS-BCR and Clinical Nomogram

We performed Cox proportional hazards regression analysis for the GG-RP, CAPRA-S, NCCN, and CAPRA models in our multicenter dataset. Performances of these four conventional popular models for BCR prediction were listed in Table 2. Statistically, the GG-RP, CAPRA-S, NCCN, and CAPRA all demonstrated discriminative power for BCR prediction ( $HR > 1$  and  $p < 0.001$ ). Significant differences of C-index between DRS-BCR/DRC-BCR and the other four clinical models were tested by the U-statistics-based C estimator, and the C-index of DRS-BCR/DRC-BCR significantly outperformed others in multicenter validation ( $p < 0.05$ ). For the performance of discriminating BCR at 3-year, IDI and continuous NRI tests were applied to evaluate the significance of the incremental performance. The test results showed that DRS-BCR maintained better concordance of 3-year BCR than other clinical models in all datasets ( $p < 0.05$ ) (Tables S8–12). The K–M curves were further utilized to compare the models for assigning patients into high-risk and low-risk of BCR-free survival, the log-rank test was used to evaluate the significant discrimination ( $p < 0.05$ ) between high-risk and low-risk. Comparisons of K–M analysis and log-rank tests between DRS-BCR and other conventional models with multi cut-off groups were carried out (Figures S4 and S6). DRS-BCR achieved stronger discriminative power than any other clinical models at almost any cut-off settings ( $p < 0.001$ ). More comparisons between DRS-BCR and acknowledgeable indicators were listed in Supplementary information IV (Table S13).

**Table 2.** Patient characteristics.

	HR	<i>p</i>	PC ( <i>n</i> = 369)	VC1 ( <i>n</i> = 34)	VC2 ( <i>n</i> = 83)
GG-RP	1.645	0.001	0.583 [0.53–0.636]	0.564 [0.419–0.709]	0.689 [0.601–0.777]
CAPRA-S	1.339	<0.001	0.677 [0.62–0.734]	0.654 [0.49–0.818]	0.654 [0.544–0.764]
NCCN	1.9022	<0.001	0.586 [0.548–0.624]	0.535 [0.408–0.662]	0.583 [0.498–0.668]
CAPRA	1.306	<0.001	0.677 [0.618–0.736]	0.552 [0.385–0.719]	0.614 [0.509–0.719]
DRS-BCR	1.705	<0.001	0.802 [0.758–0.846]	0.811 [0.722–0.9]	0.794 [0.718–0.87]
DRC-BCR	1.654	<0.001	0.807 [0.76–0.854]	0.794 [0.685–0.903]	0.8 [0.723–0.877]

Note: GG-RP: Gleason grade group of radical prostatectomy; BCR: biochemical recurrence; PC, primary cohort; VC, validation cohort; HR: hazard ratio; CAPRA: Cancer of Prostate Risk Assessment; NCCN: National Comprehensive Cancer Network; DRS-BCR: deep radiomic signature of biochemical recurrence; DRC-BCR: deep radiomic combing signature of biochemical recurrence. *p* values were two-sided.

#### 4. Discussion

Prostate cancer is characterized by its notable heterogeneity followed by a wide variation of oncologic prognosis. The majority of prostate cancers are indolent, while the rest could be very aggressive and even life threatening. Thus, developing high discriminative prognostic models to distinguish low risk of BCR patients from high-risk ones and to provide directly instructive therapeutic assistant is of great importance. In this study, we developed a survival model to predict biochemical recurrence after radical prostatectomy. This model combined conventional radiomic technique for feature extraction and deep learning algorithm for survival analysis. In multicenter validation, the deep radiomic model (DRS-BCR) outperformed three popular clinical conventional models in prediction of BCR-free survival.

Clinicians have made great efforts to identify clinical variable combinations in pursuit of better oncological prediction and controls. Some models have been widely adopted and validated in western countries such as the NCCN stratification, CAPRA/-S score and Gleason grade group system [5]. However, the performance of these models in Asian population have been reported to be suboptimal than those in the European and American population [20]. American-Asian men were reported to be more likely to have unfavorable risk profile with worse prognosis [21]. In the current study, nearly two third patients were classified into intermediate- or high-risk groups, which is in line with Korean and Japanese reports [22,23]. Conversely, most patients in European were low risk [24]. These facts have increasingly drawn our attention for a hypothesis that the entity of prostate cancer in east and west are taking distinct evolution paths. Recently, evidence supporting such differences has been discovered in genetic level. In a large national prostatic genomic analysis, Chinese population were reported to have distinct driver gene profile comparing to American population [25]. Thus, conventional predictive models might not be suitable for Asian people. These facts motivated us to develop a new stratification system according to our own clinical data. In the current study, we developed and validated a deep learning bootstrapped radiomic signature (DRS-BCR) to predict BCR-free survival. The model achieved very good performance in all centers, outweighing other four clinical models (NCCN, CAPRA/-S and Gleason grade group system) in the same setting. This was in line with several reports that radiomic models have better performance than conventional ones [13].

Prostate cancer is a multifocal malignancy with highly pronounced heterogeneity, treated by multimodal approaches. Patients are generally stratified into different risk groups followed by distinct therapies. Thus, accurate stratification is critical to certain group of patients, especially for high-risk ones. Hormonal therapy is one of the key to prostate cancer. Available evidence indicated that the use of neoadjuvant ADT in RP candidates could reduce positive surgical margin rates, EPE rates and lymph node involvements [26,27]. Of note, these benefits only occurred in high-risk patients [28]. With the widespread application of PSMA-PET-CT [29–31] and single-cell sequencing [32,33], early micrometastasis is considered to be one critical cause for biochemical recurrence or progressive metastasis after radical prostatectomy. Therefore, accurate prediction of the risk of biochemical recurrence will help the early use of systemic therapy for the control of systemic micrometastasis in order to obtain longer-term survival benefits. Based on results of predictive models, many ongoing clinical studies were designed to provide adjuvant new-generation antiandrogens for high-risk patients, such as abiraterone (NCT04513717), darolutamide (NCT04484818) and apalutamide (NCT03767244). Compared with the traditional predictive tools, our model can use prebiopsy MR make recurrence risk predictions, so it can be a good basis for stratification to guide the early application of postoperative systemic adjuvant therapy, and even guide the development of neoadjuvant systemic treatment.

Considering the clinical implementation of the DRS-BCR model, we made a comparison with CAPRA-S score by fourfold table (Table S13). According to the CAPRA-S system, 355 (73.20%) patients were classified into high-risk group, while 260 of them had not

developed BCR. Within the 260 patients, only those who harbored latent metastatic cancer cells but had not developed BCR yet could benefit from adjuvant therapies, while the rest “majority” of them were more likely to only suffer from side effects of over treatments. As comparison, DRS-BCR model significantly reduced the proportion of the false positive patients and maintained equivalent high level of negative predictive value (89.4% vs. 86.20%). This indicated that DRS-BCR model can help prevent more than one third patients from overtreatments and keep the undertreatment rate at a similarly low level as CAPRA system.

Comparing with machine learning-based Cox proportional-hazards regression model (CPH), deep survival neural network is able to describe a non-linear relationship between features and survival events. Usually, the relationship between features and survival events is often non-linear [18], which makes conventional linear models (e.g., CPH) difficult to obtain optimal performance. In the current study, imaging features were extracted and primarily selected by conventional radiomic approach. And then, a deep survival neural network algorithm was applied to develop a more complex non-linear model, naming DRS-BCR. Besides, to further mitigating overfitting, strategies such as dropout [34] Early stopping [16], transfer learning [35] were also used.

Additionally, we attempted to incorporate perioperative parameters with DRS-BCR to make a new model (DRC-BCR) in pursuit of better performance. Unluckily, DRC-BCR only demonstrated equivalent performance when comparing with DRS-BCR. This was not in line with those of most previous studies, wherein the final performance of the model can always be improved by incorporating clinical factors [11,36]. It could be assumed that isolated clinical variables can only reflect limited tumor features from one single perspective, as with one color from a prism. With a noninvasive way to analysis all information from a region, it is quite possible for radiomics to provide more complex and deeper level of information to better characterize tumor nature.

Finally, relatively short follow-up is a major limitation of our study. Selecting 3-year and 5-year BCR-free survival in a small cohort decreased the diagnostic power of the model. Further, patients in clinical high-risk group might have received long-term adjuvant ADT or adjuvant RT, which may lead to delayed BCR occurrence. Thus, 10-year long-term follow-up might better reflect the true biological process. Additionally, incorporating of DWI or ADC maps may further improve the overall accuracy and consistency of the model in the future.

## 5. Conclusions

In conclusion, quantitative features of MRI of prostate cancer showed the potential to improve the description of tumor heterogeneity. These features were empowered by deep learning-based method to build a powerful prognostic prediction model for BCR after RP. With the use of this model, it is promising to optimize neoadjuvant or adjuvant systemic therapies for suitable patients.

**Supplementary Materials:** The following are available online at [www.mdpi.com/article/10.3390/cancers13123098/s1](http://www.mdpi.com/article/10.3390/cancers13123098/s1), Figure S1: Study design, Figure S2: Detailed examples of annotations, Figure S3: Performance of DRS-BCR for predicting 5-year BCR free survival, Figure S4: Comparisons between DRS-BCR and NCCN model by Kaplan-Meier (K-M) analysis, Figure S5: Comparisons between DRS-BCR and CAPRA score by Kaplan-Meier (K-M) analysis, Figure S6: Comparisons between DRS-BCR and GG-RP by Kaplan-Meier (K-M) analysis, Table S1: MRI Parameters, Table S2: Univariate analysis of clinical factors, Table S3: Multivariate analysis of pre-operative clinical factors, Table S4: Multivariate analysis of post-operative clinical factors, Table S5: Multivariate analysis of combination of perioperative clinical factors (CS-combine), Table S6: Statistic of selected features from radiomic features, Table S7: Evaluation of DRC-BCR for BCR survival, Table S8: Comparison between DRC-BCR and DRS-BCR, Table S9: Comparison between DRS-BCR and GG-RP, Table S10: Comparison between DRS-BCR and CAPRA-S, Table S11: Comparison between DRS-

BCR and NCCN, Table S12: Comparison between DRS-BCR and CAPRA, Table S13: Fourfold comparison between DRS-BCR/CAPRA-S with true BCR events. Supplementary Information I, II, III, and IV.

**Author Contributions:** Conceptualization, Y.Y. and L.S.; methodology, Y.Y., L.S., Z.L., and M.L.; software, L.S., G.Y., K.S., X.Z.; validation, Y.Y., L.S., X.Z., K.S., L.W.; investigation, Y.Y., J.L. (Jiangang Liu), L.M. and C.L.; resources, W.H., Y.Z., C.L., M.L., X.Y., L.W.; data curation, W.H., Y.Z., C.L., X.Y., L.W.; formal analysis, L.S., and J.L. (Jiangang Liu); visualization, W.H., L.S., M.L., and X.Y.; writing—original draft preparation, Y.Y. and L.S., H.C.; writing—review and editing, H.C., Z.L., L.M., and J.L. (Jian Lu); supervision, G.Y., J.L. (Jiangang Liu), L.M., J.T. and J.L.; funding acquisition, Z.L., J.T. and J.L., project administration, J.T. and J.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by grants from the Beijing Natural Science Foundation (Z200027), the National Natural Science Foundation of China (No. 61871004, 81922040), the National Key R&D Program of China (No. 2018YFC0115900) and the Youth Innovation Promotion Association CAS (No. 2019136).

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board (or Ethics Committee) of Peking University Third Hospital (S2019326).

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The related data and materials are available for sharing upon request to Jian Lu and Jie Tian.

**Acknowledgments:** The authors thank the Department of Pathology and Department of Radiology, Peking University Third Hospital, the Department of Urology, Peking University People's Hospital and the Department of Radiology of Beijing Friendship Hospital, for their support with data collection.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer statistics, 2019. *CA Cancer J. Clin.* **2019**, *69*, 7–34.
2. Diaz, M.; Peabody, J.O.; Kapoor, V.; Sammon, J.; Rogers, C.G.; Stricker, H.; Lane, Z.; Gupta, N.; Bhandari, M.; Menon, M. Oncologic outcomes at 10 years following robotic radical prostatectomy. *Eur. Urol.* **2015**, *67*, 1168–1176.
3. Walz, J.; Chun, F.K.; Klein, E.A.; Reuther, A.; Saad, F.; Graefen, M.; Huland, H.; Karakiewicz, P.I. Nomogram predicting the probability of early recurrence after radical prostatectomy for prostate cancer. *J. Urol.* **2009**, *181*, 601–607.
4. Freedland, S.J.; Humphreys, E.B.; Mangold, L.A.; Eisenberger, M.; Partin, A.W. Time to prostate specific antigen recurrence after radical prostatectomy and risk of prostate cancer specific mortality. *J. Urol.* **2006**, *176*, 1404–1408.
5. Zelic, R.; Garmo, H.; Zugna, D.; Stattin, P.; Richiardi, L.; Akre, O.; Pettersson, A. Predicting Prostate Cancer Death with Different Pretreatment Risk Stratification Tools: A Head-to-head Comparison in a Nationwide Cohort Study. *Eur. Urol.* **2020**, *77*, 180–188.
6. Turkbey, B.; Rosenkrantz, A.B.; Haider, M.A.; Padhani, A.R.; Villeirs, G.; Macura, K.J.; Tempany, C.M.; Choyke, P.L.; Cornud, F.; Margolis, D.J.; et al. Prostate Imaging Reporting and Data System Version 2.1: 2019 Update of Prostate Imaging Reporting and Data System Version 2. *Eur. Urol.* **2019**, *76*, 340–351.
7. Liu, Z.; Wang, S.; Dong, D.; Wei, J.; Fang, C.; Zhou, X.; Sun, K.; Li, L.; Li, B.; Wang, M.; et al. The Applications of Radiomics in Precision Diagnosis and Treatment of Oncology: Opportunities and Challenges. *Theranostics* **2019**, *9*, 1303–1322.
8. Zamboglou, C.; Carles, M.; Fechter, T.; Kiefer, S.; Reichel, K.; Fassbender, T.F.; Bronsert, P.; Koeber, G.; Schilling, O.; Ruf, J.; et al. Radiomic features from PSMA PET for non-invasive intraprostatic tumor discrimination and characterization in patients with intermediate- and high-risk prostate cancer—A comparison study with histology reference. *Theranostics* **2019**, *9*, 2595–2605.
9. Chen, T.; Li, M.; Gu, Y.; Zhang, Y.; Yang, S.; Wei, C.; Wu, J.; Li, X.; Zhao, W.; Shen, J. Prostate Cancer Differentiation and Aggressiveness: Assessment With a Radiomic-Based Model vs. PI-RADS v2. *J. Magn. Reson. Imaging* **2019**, *49*, 875–884.
10. Shao, L.; Yan, Y.; Liu, Z.; Ye, X.; Xia, H.; Zhu, X.; Zhang, Y.; Zhang, Z.; Chen, H.; He, W.; et al. Radiologist-like artificial intelligence for grade group prediction of radical prostatectomy for reducing upgrading and downgrading from biopsy. *Theranostics* **2020**, *10*, 10200–10212.
11. Shiradkar, R.; Ghose, S.; Jambor, I.; Taimen, P.; Ettala, O.; Purysko, A.S.; Madabhushi, A. Radiomic features from pretreatment biparametric MRI predict prostate cancer biochemical recurrence: Preliminary findings. *J. Magn. Reson. Imaging* **2018**, *48*, 1626–1636.

12. Bourbonne, V.; Fournier, G.; Vallieres, M.; Lucia, F.; Doucet, L.; Tissot, V.; Cuvelier, G.; Hue, S.; Du, H.I.; Perdriel, L.; et al. U. External Validation of an MRI-Derived Radiomics Model to Predict Biochemical Recurrence after Surgery for High-Risk Prostate Cancer. *Cancers* **2020**, *12*, 814.
13. Bourbonne, V.; Vallieres, M.; Lucia, F.; Doucet, L.; Visvikis, D.; Tissot, V.; Pradier, O.; Hatt, M.; Schick, U. MRI-Derived Radiomics to Guide Post-operative Management for High-Risk Prostate Cancer. *Front. Oncol.* **2019**, *9*, 807.
14. Cookson, M.S.; Aus, G.; Burnett, A.L.; Canby-Hagino, E.D.; D'Amico, A.V.; Dmochowski, R.R.; Eton, D.T.; Forman, J.D.; Goldenberg, S.L.; Hernandez, J.; et al. Variation in the definition of biochemical recurrence in patients treated for localized prostate cancer: The American Urological Association Prostate Guidelines for Localized Prostate Cancer Update Panel report and recommendations for a standard in the reporting of surgical outcomes. *J. Urol.* **2007**, *177*, 540–545.
15. Van Griethuysen, J.J.M.; Fedorov, A.; Parmar, C.; Hosny, A.; Aucoin, N.; Narayan, V.; Beets-Tan, R.G.H.; Fillion-Robin, J.C.; Pieper, S.; Aerts, H. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res.* **2017**, *77*, e104–e107.
16. Huang, G.; Liu, Z.; Pleiss, G.; van der Maaten, L.; Weinberger, K. Convolutional Networks with Dense Connectivity. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, doi:10.1109/TPAMI.2019.2918284.
17. Coates, A.; Ng, A.Y. The Importance of Encoding Versus Training with Sparse Coding and Vector Quantization. In Proceedings of the 28th International Conference on Machine Learning, ICML 2011, Bellevue, Washington, DC, USA, 28 June–2 July 2011.
18. Katzman, J.L.; Shaham, U.; Cloninger, A.; Bates, J.; Kluger, Y. DeepSurv: Personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC Med. Res. Methodol.* **2018**, *18*, 24.
19. Kingma, D.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2014**, arXiv:1412.6980. Available online: <https://arxiv.org/abs/1412.6980> (accessed on January 30, 2017).
20. Hu, M.B.; Yang, T.; Hu, J.M.; Zhu, W.H.; Jiang, H.W.; Ding, Q. Prognostic factors in Chinese patients with prostate cancer receiving primary androgen deprivation therapy: Validation of Japan Cancer of the Prostate Risk Assessment (J-CAPRA) score and impacts of pre-existing obesity and diabetes mellitus. *Int. J. Clin. Oncol.* **2018**, *23*, 591–598.
21. Lichtensztajn, D.Y.; Gomez, S.L.; Sieh, W.; Chung, B.I.; Cheng, I.; Brooks, J.D. Prostate cancer risk profiles of Asian-American men: Disentangling the effects of immigration status and race/ethnicity. *J. Urol.* **2014**, *191*, 952–956.
22. Ahn, H.; Kim, H.J.; Jeon, S.S.; Kwak, C.; Sung, G.T.; Kwon, T.G.; Park, J.Y.; Paick, S.H. Establishment of Korean prostate cancer database by the Korean Urological Oncology Society. *Investig. Clin. Urol.* **2017**, *58*, 434–439.
23. Tanaka, N.; Nakai, Y.; Miyake, M.; Anai, S.; Inoue, T.; Fujii, T.; Konishi, N.; Fujimoto, K. Trends in risk classification and primary therapy of Japanese patients with prostate cancer in Nara urological research and treatment group (NURTG)—comparison between 2004–2006, 2007–2009, and 2010–2012. *BMC Cancer* **2017**, *17*, 616.
24. Punnen, S.; Freedland, S.J.; Presti, J.C., Jr.; Aronson, W.J.; Terris, M.K.; Kane, C.J.; Amling, C.L.; Carroll, P.R.; Cooperberg, M.R. Multi-institutional validation of the CAPRA-S score to predict disease recurrence and mortality after radical prostatectomy. *Eur. Urol.* **2014**, *65*, 1171–1177.
25. Li, J.; Xu, C.; Lee, H.J.; Ren, S.; Zi, X.; Zhang, Z.; Wang, H.; Yu, Y.; Yang, C.; Gao, X.; et al. A genomic and epigenomic atlas of prostate cancer in Asian populations. *Nature* **2020**, *580*, 93–99.
26. Schulman, C.C.; Debruyne, F.M.; Forster, G.; Selvaggi, F.P.; Zlotta, A.R.; Witjes, W.P. 4-Year follow-up results of a European prospective randomized study on neoadjuvant hormonal therapy prior to radical prostatectomy in T2-3N0M0 prostate cancer. European Study Group on Neoadjuvant Treatment of Prostate Cancer. *Eur. Urol.* **2000**, *38*, 706–713.
27. Labrie, F.; Cusan, L.; Gomez, J.L.; Diamond, P.; Suburu, R.; Lemay, M.; Tetu, B.; Fradet, Y.; Belanger, A.; Candas, B. Neoadjuvant hormonal therapy: The Canadian experience. *Urology* **1997**, *49*, 56–64.
28. McClintock, T.R.; von Landenberg, N.; Cole, A.P.; Lipsitz, S.R.; Gild, P.; Sun, M.; Fletcher, S.A.; Roghmann, F.; Menon, M.; Nguyen, P.L.; et al. Neoadjuvant Androgen Deprivation Therapy Prior to Radical Prostatectomy: Recent Trends in Utilization and Association with Postoperative Surgical Margin Status. *Ann. Surg. Oncol.* **2019**, *26*, 297–305.
29. Klingenberg, S.; Jochumsen, M.R.; Ulhøi, B.P.; Fredsøe, J.; Sørensen, K.D.; Borre, M.; Bouchelouche, K. 68Ga-PSMA PET/CT for Primary Lymph Node and Distant Metastasis NM Staging of High-Risk Prostate Cancer. *J. Nucl. Med.* **2021**, *62*, 214–220.
30. Giesel, F.L.; Kesch, C.; Yun, M.; Cardinale, J.; Haberkorn, U.; Kopka, K.; Kratochwil, C.; Hadaschik, B.A. 18F-PSMA-1007 PET/CT detects micrometastases in a patient with biochemically recurrent prostate cancer. *Clin. Genitourin Cancer* **2017**, *15*, e497–e499.
31. Sanseverino, R.; Intilla, O.; di Mauro, U.; Molisso, G.; Adesso, M.; Napodano, G. Predictive Factors of Nodal Micrometastasis in Patients with Organ Confined Prostate Cancer. *Anticancer Res.* **2017**, *37*, 2082–2083.
32. Chen, S.; Zhu, G.; Yang, Y.; Wang, F.; Xiao, Y.T.; Zhang, N.; Bian, X.; Zhu, Y.; Yu, Y.; Liu, F.; et al. Single-cell analysis reveals transcriptomic remodellings in distinct cell types that contribute to human prostate cancer progression. *Nat. Cell Biol.* **2021**, *23*, 87–98.
33. Murray, N.P.; Aedo, S.; Fuentealba, C.; Reyes, E.; Salazar, A.; Lopez, M.A.; Minzer, S.; Orrego, S.; Guzman, E. Minimal Residual Disease Defines the Risk and Time to Biochemical Failure in Patients with Pt2 and Pt3a Prostate Cancer Treated With Radical Prostatectomy: An Observational Prospective Study. *Urol. J.* **2020**, *17*, 262–270.
34. Jindal, I.; Nokleby, M.; Chen, X. Learning Deep Networks from Noisy Labels with Dropout Regularization. In Proceedings of the 2016 IEEE 16th International Conference on Data Mining (ICDM), Barcelona, Spain, 12–15 December 2016; pp. 967–972.
35. Pan, S.J.; Yang, Q. A Survey on Transfer Learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359.
36. Osman, S.O.S.; Leijenaar, R.T.H.; Cole, A.J.; Lyons, C.A.; Hounsell, A.R.; Prise, K.M.; O'Sullivan, J.M.; Lambin, P.; McGarry, C.K.; Jain, S. Computed Tomography-based Radiomics for Risk Stratification in Prostate Cancer. *Int. J. Radiat. Oncol. Biol. Phys.* **2019**, *105*, 448–456.