Dosiomics for intensity-modulated radiotherapy in patients with prostate cancer: Survival analysis stratified by baseline PSA and Gleason grade group in a two-institutional retrospective study

Yu Murakami, MSc*,†; Daisuke Kawahara, PhD*; Takashi Soyano, MD‡; Takuyo Kozuka, MD, PhD§; Yuka Takahashi, BSc‖; Konatsu Miyake, MSc‖; Kenichi Kashihara, MD, PhD‖; Tairo Kashihara, MD, PhD¶; Tatsuya Kamima, BSc#; Masahiko Oguchi, MD, PhD#; Yuji Murakami, MD, PhD*; Yasuo Yoshioka, MD, PhD#; and Yasushi Nagata, MD, PhD*

**Department of Radiation Oncology, Graduate School of Biomedical Health Sciences, Hiroshima University, 1-3-2, Kagamiyama, Higashihiroshima, Hiroshima, 734-8551, Japan.*

†*Department of Physics, Cancer Institute, Japanese Foundation for Cancer Research, 3-8-31 Ariake, Koto-ku, Tokyo, 135-8550, Japan*

‡*Department of Radiology, Japan Self-Defense Forces Central Hospital, 1-2-24 Ikejiri, Setagaya-ku, Tokyo, 154-8532, Japan*

§*Department of Radiology, University of Tokyo Hospital, 7-3-1, Hongo, Bunkyo-ku, Tokyo, 113-8655 Japan*

‖*Tokyo Radiation Oncology Clinic, 3-5-7, Ariake, Koto-ku, Tokyo, 135-0063, Japan*

¶ *Department of Radiation Oncology, National Cancer Center Hospital, 5-1-1, Tsukiji, Chuo-ku, Tokyo, 104-0045, Japan*

Radiation Oncology Department, Cancer Institute Hospital, Japanese Foundation for Cancer Research, 3-8-31 Ariake, Koto-ku, Tokyo, 135-8550, Japan

Corresponding author:
Daisuke Kawahara, PhD
Department of Radiation Oncology, Graduate School of Biomedical Health Sciences, Hiroshima University, 1-3-2, Kagamiyama, Higashihiroshima, Hiroshima, 734-8551, Japan.
Tel: (+81) 82-257-1545; *E-mail:* daika99@hiroshima-u.ac.jp

Statistician:
Y Murakami, MSc
Department of Radiation Oncology, Graduate School of Biomedical Health Sciences, Hiroshima University, 1-3-2, Kagamiyama, Higashihiroshima, Hiroshima, 734-8551, Japan.
Tel: (+81) 82-257-1545; *E-mail:* mrkm.1125@gmail.com

Conflict of Interest: None; Funding statement: None; Acknowledgements: None
Data sharing statement: All data generated and analyzed during this study are included in this published article (and its supplementary information files)
Short title: Survival analysis for dosiomics in prostate cancer
Type of Manuscript: Full paper
Keywords: Dosiomics, Prostate, IMRT, Survival analysis, Biochemical recurrence (BCR)

Abstract

Objective: This study evaluated the prognostic impact of the quality of dose distribution using dosiomics in patients with prostate cancer, stratified by pretreatment prostate-specific antigen (PSA) levels and Gleason grade group (GG).

Methods: A total of 721 patients (Cohort A [anonymized]: N = 489 and Cohort B [anonymized]: N = 232) with localized prostate cancer treated by intensity-modulated radiation therapy were enrolled. Two predictive dosiomic features for biochemical recurrence (BCR) were selected, and patients were divided into certain groups stratified by pretreatment PSA levels and GG. Freedom from biochemical failure (FFBF) was estimated using the Kaplan–Meier method based on each dosiomic feature, and univariate discrimination was evaluated using the log-rank test. As an exploratory analysis, a dosiomics hazard (DH) score was developed, and its prognostic power for BCR was examined.

Results: The dosiomic feature extracted from planning target volume (PTV) significantly distinguished the high- and low-risk groups in patients with PSA levels >10 ng/ml (7-year FFBF: 86.7% vs. 76.1%, $p < 0.01$), GG 4 (92.2% vs. 76.9%, $p < 0.01$), and GG 5 (83.1% vs. 77.8%, $p = 0.04$). The DH score showed significant association with BCR (hazard score: 2.04; 95% confidence interval: 1.38–3.01; $p < 0.001$).

Conclusion: The quality of planned dose distribution on PTV may affect the prognosis of patients with poor prognostic factors, such as PSA levels >10 ng/ml and higher GGs.

Advances in knowledge: The effects of planned dose distribution on prognosis differ depending on the patient's clinical background.

INTRODUCTION

External beam radiation therapy (EBRT), along with surgery and brachytherapy, is an effective treatment option for patients with localized prostate cancer. However, approximately 15% of the patients develop a biochemical recurrence (BCR) after EBRT.[1] Baseline prostate-specific antigen (PSA) level and Gleason score are considered significant predictors for BCR after radiotherapy.[2,3]

Dosiomics is a method inspired by radiomics wherein numerous spatial features are extracted from dose-distribution images. Dosiomic features are expected to function as new potential metrics for evaluating the treatment plan, instead of conventional dose indices; this is because of the ability of dosiomics to detect the small differences in the dose distributions with and without recurrence or complications.[4–11] Our previous study demonstrated that the dosiomic features extracted from clinical target volume (CTV) and planning target volume (PTV) significantly correlated with BCR after

radiotherapy.[11] However, which patients with specific clinical backgrounds are significantly affected by planned dose distribution is still unclear. The sensitivity of prognostic prediction from the dose distribution may differ according to the patient's background. Previous clinical trials have revealed that dose-escalation improves freedom from biochemical failure (FFBF) with the largest benefit observed in patients with PSA ≥10 ng/ml.[1,12] Zelefsky et al. reported that dose escalation was associated with improved PSA relapse-free survival in unfavorable risk cases.[13] Thus, this study hypothesized that the quality of dose distribution may significantly affect the prognoses in patients with such clinical backgrounds, similar to those in dose escalation trials.

This study evaluated the prognostic impact of the quality of dose distribution using dosiomics in patients with prostate cancer, stratified by pretreatment PSA levels and Gleason grade group (GG). Moreover, as an exploratory analysis, a new evaluation metric for treatment planning was developed, and its prognostic power for BCR was examined.

METHODS AND MATERIALS

Patients

This is a retrospective, observational study. Flowchart of patient selection is presented in Figure 1. Four hundred and eighty-nine of the 712 patients who received intensity-modulated radiation therapy (IMRT) between May 2007 and September 2018 at Institution A (Cohort A) and 232 of 405 patients who received IMRT between June 2008 and June 2018 at Institution B (Cohort B) were included. All patients in both cohorts had adenocarcinoma of prostate and were prescribed a dose of 78 Gy/ 39 fractions to the PTV using static field IMRT. Patients with a follow-up time of ≤5 years[14] and pretreatment PSA level of ≥200 ng/mL[15] were excluded. If the patient had not relapsed at the last follow-up, it was considered as No-BCR. Finally, an integrated cohort that combined the Cohort A and Cohort B (N = 721) was created, and the patients were divided into certain groups stratified by pretreatment PSA levels and GG. Phoenix definition was used to define the BCR.[16] The study was approved by the ethics committee of the authors' institution (XXX-XXX).

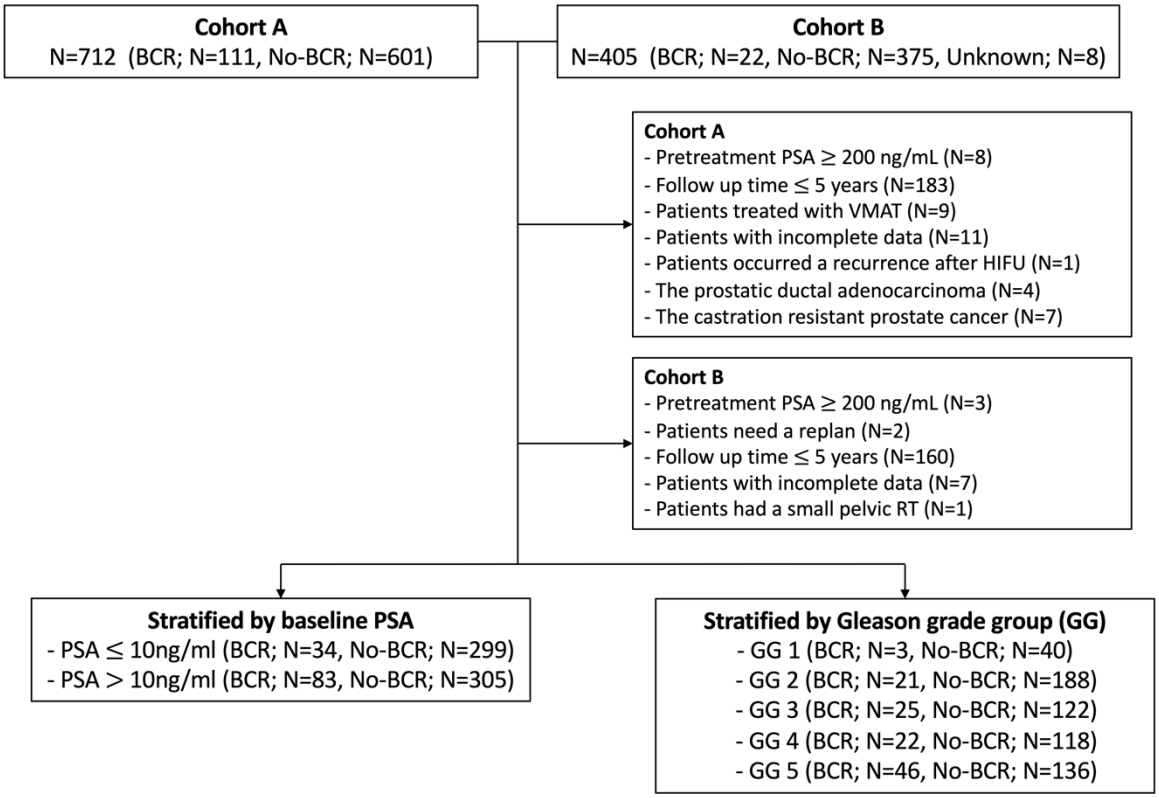

Figure 1. Flowchart of patient selection. BCR, biochemical recurrence; PSA, prostate-specific antigen; RT, radiation therapy; VMAT, volumetric-modulated arc therapy; HIFU, high-intensity focused ultrasound; GG, Gleason grade group.

Treatment planning for Cohort A

The details of delineation and treatment at Cohort A are described previously.[11,17] Five-field IMRT was used for all cases with the same beam angles (255, 315, 45, 105, and 180). All treatment plans were created by Eclipse treatment planning system (TPS) ver. 8.6 or 10.0 (Varian Medical Systems, Palo Alto, CA, USA). D95% = 100% to PTV was used for planning normalization. The treatment beams were 10-MV photon beams from Clinac 21EX accelerator (Varian Medical Systems). An analytical anisotropic algorithm (AAA) with a 2.5-mm dose grid was adopted for dose calculation. The preset dose constraints for IMRT are summarized in Table S1.

Treatment planning for Cohort B

Most patients (77.2%) were treated with seven-field IMRT (215, 260, 305, 0, 55, 100, and 145), whereas the remaining patients (22.8%) were treated with five-field IMRT (255, 315, 45, 105, and 180). D95% = 100% to PTV was used for nearly all patients (97.4%), while other normalization values were applied

for certain patients (2.6%). Delineation of the target and organs-at-risk (OARs) was the same as those in the Cohort A, while certain CTVs (33.2%) and PTVs (8.2%) were slightly adjusted according to the clinical judgement of the radiation oncologist. The TPS was the Eclipse software version 8.1 or 11.0, and the treatment beams were 10-MV photon beams from Clinac iX accelerator (Varian Medical Systems). Dose distribution was calculated using AAA with a 2.5-mm dose grid. The preset dose constraints are summarized in Table S2.

Feature extraction for Dosiomics

Dose distributions were resampled to have isotropic voxels (1, 1, 1 mm) using B-spline interpolation.[10,11] Thereafter, dosiomic features for CTV and PTV were calculated from the discretized 3D dose voxel dimensions with fixed bin widths of 1 Gy.[11] In total, 1,650 features including 210 (105 × 2) original features and 1,440 (720 × 2) wavelet features were extracted using PyRadiomics version 3.0.[18] The wavelet filter computed eight decompositions for each level. Summary of dosiomic features used in this study is shown in Table S3. Spearman's correlation coefficient (SCC) was calculated between the features of all possible two combinations, and then the features with SCC of ≥0.80 were eliminated.[11,19]

Dosiomics hazard score

Z-score normalization was used to standardize each feature. Then, certain prognostic dosiomic features for BCR were selected via five-fold cross-validation using the univariate Cox proportional hazard (CPH) regression. The cross-validation was performed 20 times (100 loops), with randomization of the inner dataset in each loop. Subsequently, C-index was computed for each random validation dataset, and the top three features with a higher mean C-index were chosen for both CTV and PTV. The C-index is an indicator to evaluate the goodness of fit measure for created model; C-index = 0.5 and 1 imply random and perfect predictions, respectively. Six prognostic features were selected as candidate features for constructing the multivariate CPH regression model. Prior to building the model, a variance inflation factor (VIF)[20] was computed between the features of all possible two combinations. The features with VIF >10 were excluded to avoid collinearity between the features.[7] Finally, the dosiomics hazard (DH) score was calculated as:

$$Dosiomics\ hazard\ score = \omega_1 x_1 + \omega_2 x_2 + \cdots \omega_n x_n \qquad (1)$$

Where ω is the estimated relative hazard risk of BCR in the multivariate model, and x corresponds to the value of the dosiomic feature. The DH score was inspired by the radiomics score,[21] which estimates

the individual risk of BCR from the whole dose distribution including the CTV and PTV.

Evaluation

The two predictive dosiomic features (*CTV_wavelet-HHH_glrlm_HighGrayLevelRunEmphasis* [HGLRE] and *PTV_wavelet-HHH_firstorder_Entropy*) and DH score were selected as candidate features for the Kaplan–Meier analyses. The two dosiomic features were selected based on the previous results[11] to examine the generalizability of the dosiomic features with a larger cohort. BCR time was calculated from the date of IMRT completion to the detection of BCR, while No-BCR time was calculated from the date of IMRT completion to the date of last visit (censored). Kaplan–Meier analysis was applied to all patient groups stratified by pretreatment PSA levels of 10 ng/ml[1,12] and GGs 1–5. The GG has a five-grade group based on the original Gleason score.[22] The high- and low-risk groups were split using the median feature value of the dataset in each stratified group. To examine the feasibility of DH score, a univariate CPH regression model was built using seven clinical factors, one dose-volume histogram (DVH) parameter, and three dosiomic features including the DH score. Following the calculations for the VIF, the variable with p-value < 0.05 in the univariate model was further applied to the multivariate model.

Statistics

Differences in patient characteristics between with and without BCR were assessed using the Mann–Whitney U- and Fisher's tests for continuous and categorical variables. 7-year FFBF rates were estimated using the Kaplan–Meier method, and the log-rank test was used to assess differences in FFBF between the high- and low-risk groups. Kaplan–Meier curves were estimated using the Python package *lifelines* version 0.25.4 (https://doi.org/10.5281/zenodo.4002777), and the p-values were computed using the log-rank test. Statistical analyses were conducted using the R software version 3.6.3 (https://cran.r-project.org/bin/macosx/). All p-values were two-sided and all tests were performed using a 5% significance level.

RESULTS

Patient characteristics

The patient characteristics of the integrated cohort are presented in Table 1. The median follow-up times were 100.2 (range: 60.3–153.7) and 82.6 (range: 60.3–126.4) months in Cohorts A and B, respectively, except for two patients who had developed BCR. These two patients with a median follow-up time of

47.9 and 55.8 months, respectively, were included in this analysis owing to the limited number of patients with BCR in the Cohort B (N = 21). The median times to BCR were 55.2 (range: 9.7–149.3) and 46.8 (range: 4.8–100.7) months in Cohorts A and B, respectively. The FFBF rates at 3 and 5 years were 94.7 and 89.0% in Cohort A (Figure S1) and 96.1 and 93.5% in Cohort B (Figure S2), respectively. Patient characteristics of the Cohorts A and B are presented in Tables S4 and S5, respectively.

Dosiomics hazard score

After assessing the SCC, the 1,650 dosiomic features were reduced to 150 and 143 robust features for the CTV (Table S6) and PTV (Table S7), respectively. Consequently, a multivariate model was created using the top three features for the CTV and PTV (Table S8), and the DH score for BCR was calculated. Forest plots of the hazard ratios for FFBF in the multivariate CPH regression model using these six prognostic features are shown in Figure S3.

Kaplan–Meier estimates for patient groups

FFBF curves separated by *CTV_wavelet-HHH_glrlm_HGLRE*, *PTV_wavelet-HHH_firstorder_Entropy* and DH scores were obtained from the patient groups according to the pretreatment PSA levels and GG (Figures 2–5). Significant differences in the survival curves were observed between the high- and low-risk groups in patients with pretreatment PSA levels >10 ng/ml for *PTV_wavelet-HHH_firstorder_Entropy* (7-year FFBF: 86.7% vs. 76.1%, $p < 0.01$) and DH score (7-year FFBF: 86.0% vs. 76.8%, $p < 0.01$), respectively. Regarding GG, the *CTV_wavelet-HHH_glrlm_HGLRE* significantly distinguished the high- and low-risk groups in patients with GG 4 (7-year FFBF: 89.6% vs. 79.3%, $p = 0.04$). Further, *PTV_wavelet-HHH_firstorder_Entropy* significantly discriminated between high- and low-risk groups in patients with GG 4 (7-year FFBF: 92.2% vs. 76.9%, $p < 0.01$) and GG 5 (7-year FFBF: 83.1% vs. 77.8%, $p = 0.04$), respectively. Furthermore, the DH score significantly distinguished the high- and low-risk groups in patients with GG 5 (7-year FFBF: 88.5% vs. 72.9%, $p < 0.001$).

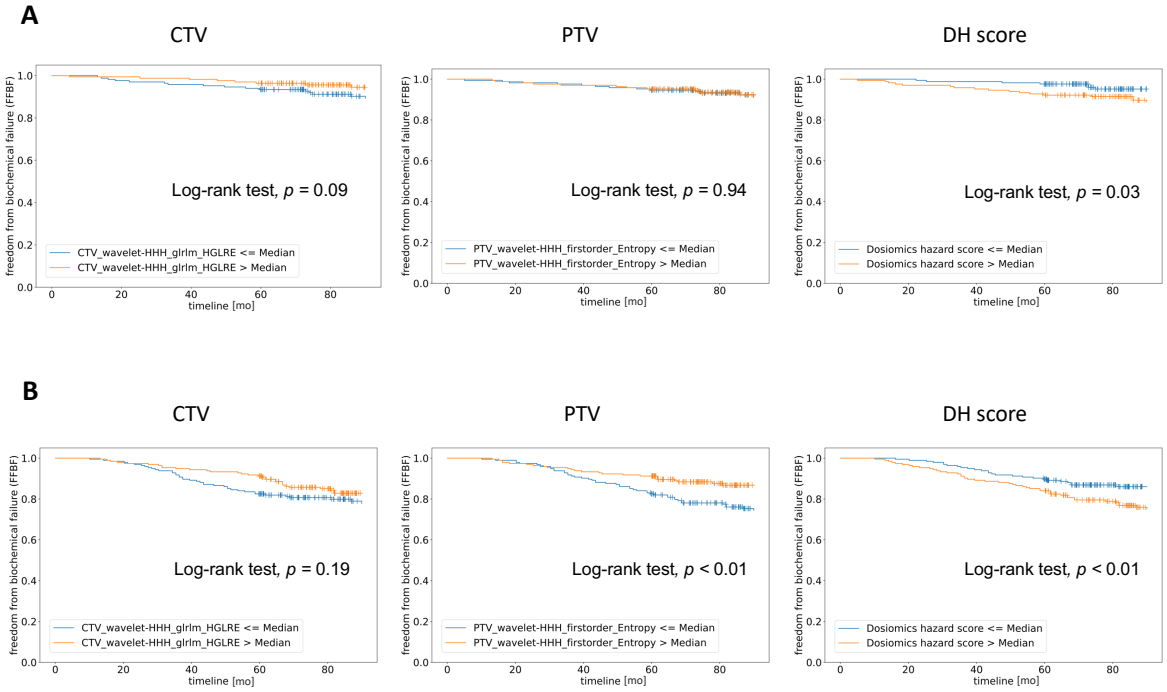

Figure 2. Kaplan–Meier estimates of FFBF for high- and low-risk BCR groups separated by two predictive dosiomic features and DH score. (A) Patient group with pretreatment PSA level ≤ 10 ng/ml, (B) Patient group with pretreatment PSA level >10 ng/ml.

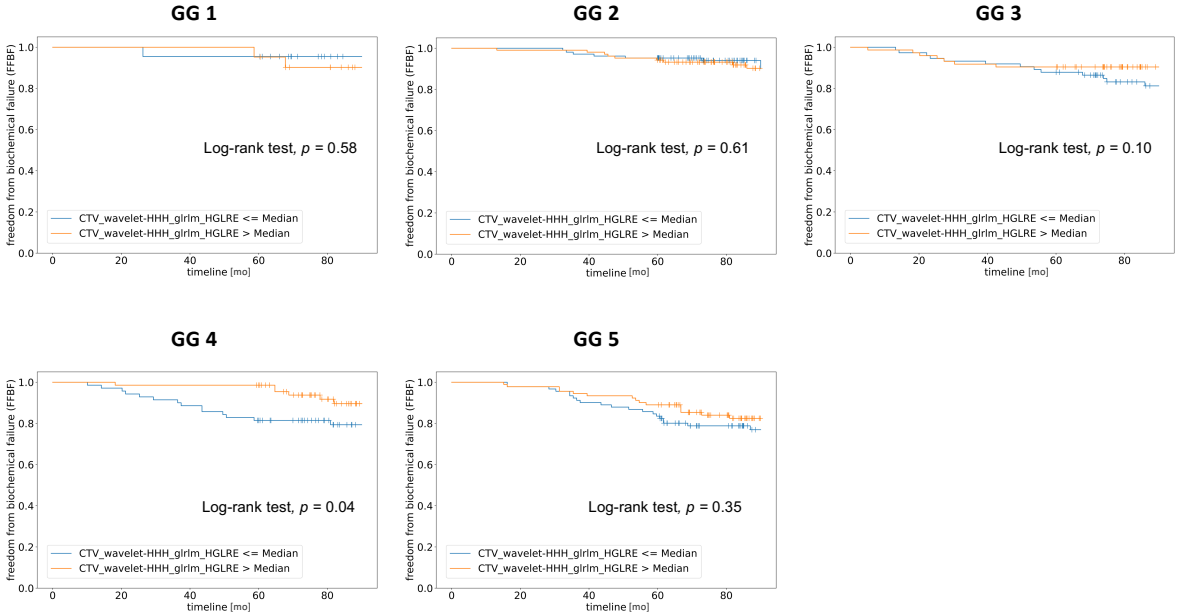

Figure 3. Kaplan–Meier estimates of FFBF for high- and low-risk BCR groups separated by CTV-derived dosiomic feature, according to GG.

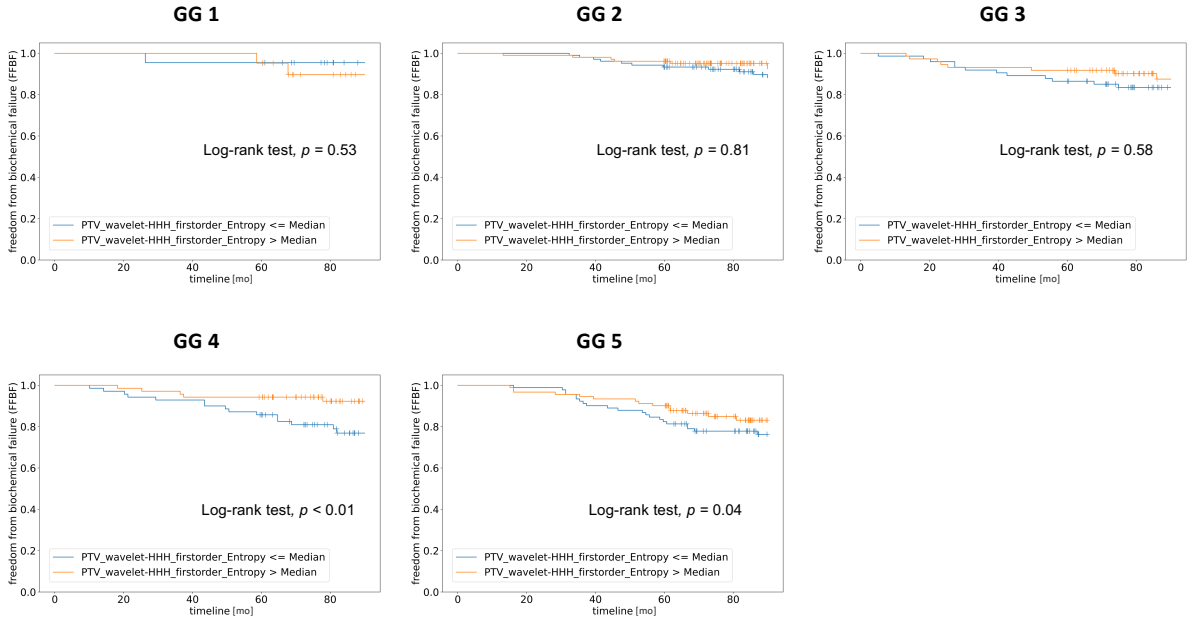

Figure 4. Kaplan–Meier estimates of FFBF for high- and low-risk BCR groups separated by PTV-derived dosiomic feature, according to GG.

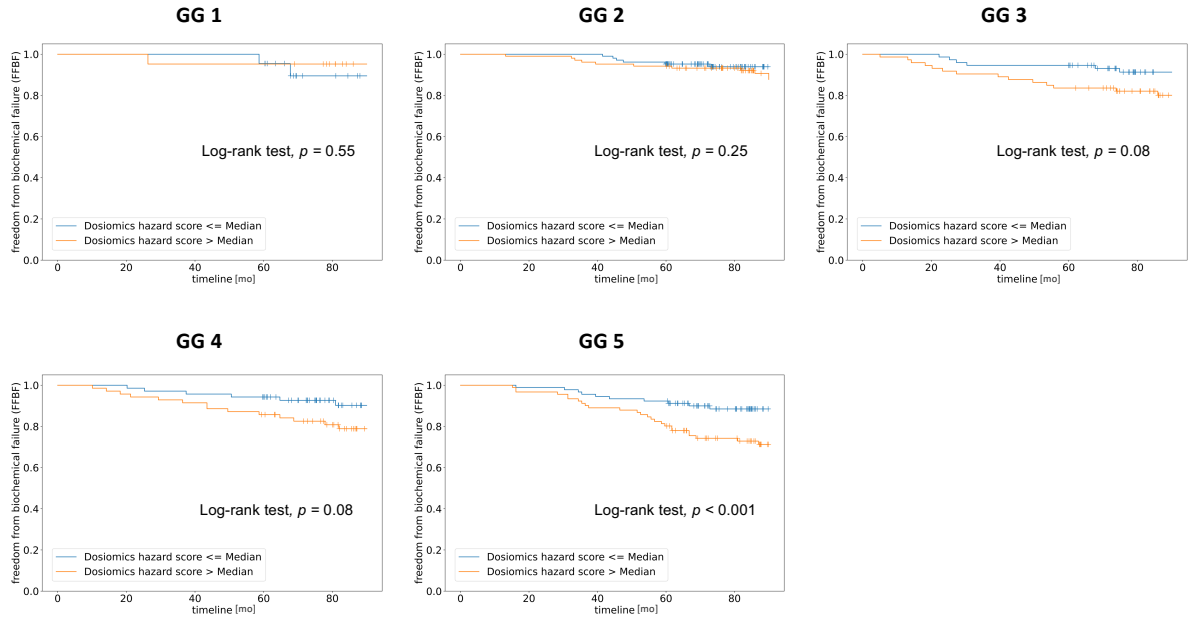

Figure 5. Kaplan–Meier estimates of FFBF for high- and low-risk BCR groups separated by DH score, according to GG.

Efficacy of the dosiomics hazard score

Univariate and multivariate CPH regression analyses results are summarised in Table 2, and Figure S4 presents the forest plots of the hazard ratios for FFBF in the multivariate model. In the multivariate model, T-stage (hazard ratio [HR]: 1.53, 95% confidence interval [CI]:1.01–2.34, $p = 0.04$), pretreatment

PSA (HR: 1.61, 95% CI: 1.06–2.46, $p = 0.02$), GG (HR: 1.25, 95% CI: 1.07–1.47, $p < 0.01$), positive biopsy core rate (PBCR) (HR: 1.55, 95% CI: 1.04–2.32, $p = 0.03$), and DH score (HR: 2.04, 95% CI: 1.38–3.01, $p < 0.001$) were significantly correlated with BCR. Notably, the NCCN risk group was excluded from the multivariate model because the calculated VIF between the NCCN risk group and GG was >10.

DISCUSSION

This study evaluated the prognostic impact of the planned dose distribution quality using dosiomics in patients with prostate cancer stratified by pretreatment PSA levels and GG. Consequently, it was demonstrated that the dosiomic feature extracted from PTV can significantly discriminate between the high- and low-risk BCR groups with PSA levels >10 ng/ml and GG of 4 and 5. However, no statistically significant differences were observed in any patient group, except for patients with GG 4, when the CTV-derived dosiomic feature was used. Thus, the sensitivity of prognostic prediction from dose distribution differed according to the patient's background and the type of dosiomic features used. Although dosiomics can capture differences in dose distributions in the CTV,[11] it can be assumed that the quality of dose distributions in the PTV had a greater influence on the value of the dosiomic features. This was because the dose of 5% of the PTV volume was not considered in treatment planning (i.e., D95% = 100% was used in both institutions). The observed larger differences in the FFBF curves for the PTV compared with those for the CTV are consistent with a previous report.[11]

PTV_wavelet-HHH_firstorder_Entropy specifies the uncertainty and randomness of the dose distribution after wavelet transformation with high-pass filtering in the x-, y-, and z-dimensions. Feature maps of high- and low-risk patients using *PTV_wavelet-HHH_firstorder_Entropy* are shown in Figure S5. The feature maps of low-risk patients tended to be more homogeneous than those of high-risk patients. Considering that smaller feature values in the voxels indicate a higher risk of BCR, the feature map could aid a treatment planner in assessing whether the dose distribution quality is appropriate, in addition to information from conventional DVH parameters. Interestingly, despite exhibiting the lowest feature value in the high-risk group, the patient in the upper-left panel with a PSA level of 5.4 ng/ml did not experience BCR. This indicates that patients with favourable clinical backgrounds may be more tolerant to inferior dose distributions in treatment planning. Our future research aims to elucidate the physiological and biological mechanisms underlying the relationship between spatial dose distribution, feature maps, and the occurrence of BCR.

No significant differences were observed in the survival curves of patients with PSA levels

≤10 ng/ml and GGs of 1–3. This indicates that the dose distribution quality does not influence the prognosis of patients with good prognostic factors, provided that the predetermined dose constraints satisfy the clinical criteria. Therefore, a treatment planner could prioritise sparing the OARs to prevent gastrointestinal or genitourinary complications in these patients. In addition, these patients are good candidates for rapid planning,[23,24] which is essential for efficient treatment planning and allows clinicians to focus on more challenging cases.

However, the current TPS does not support dosiomics-based treatment planning. Thus, a retrospective evaluation using in-house software after calculating the final dose distribution is essential to assess the quality of the dose distribution using dosiomics. As this process is very time-consuming, identifying the patient groups to be used for dosiomics-based evaluation is important. Our findings may aid decision-making in clinical practice. The optimal solution is to incorporate dosiomics into the iterative dose optimisation process for IMRT in treatment planning.

To date, the relationships between dosimetric factors related to treatment planning and BCR remain unclear. A previous study demonstrated that certain dosiomic features are associated with BCR, and their predictive performance for BCR outperformed that of DVH parameters.[11] The present study observed that PTV-derived feature can significantly distinguish the high- and low-risk groups in patients with poor prognostic factors in a two-institution integrated cohort. However, the best dosiomic feature for PTV in this study was not identical to that previously reported (i.e., *PTV_wavelet-LHH_glszm_SizeZoneNonUniformity* [SZNU] vs. *PTV_wavelet-HHH_firstorder_Entropy*). This implies that several dosiomic features may be related to BCR. Interestingly, *PTV_wavelet-LHH_glszm_SZNU* significantly discriminated between the high- and low-risk groups across all PSA levels and in patients with GGs of 4 and 5 (Figures S6 and S7). If several dosimetric factors for BCR are present, several dosiomic features can be considered for prognostic prediction similar to the case of the DH score. The prognostic power using DVH parameters was limited in prostate cancer (Figures S8 and S9), highlighting the importance of dosiomic analysis for dose distribution.

Previous studies reported that the prognosis differed between Gleason scores of 3 + 4 (i.e., GG 2) and 4 + 3 (i.e., GG 3).[22,25] Therefore, GG was used as a stratification factor rather than the Gleason score for the Kaplan–Meier analysis. Unexpectedly, no statistically significant differences were observed in the survival curves of these groups, although there were moderate differences in patients with GG 3 based on the DH score.

The DH score, inspired by the concept of the Rad score,[21] was developed to evaluate whole dose distributions, including the CTV and PTV. Interestingly, certain differences in the survival curves were observed early following the completion of treatment, and an elevated DH score increased the risk

of BCR. The DH score may provide valuable information to a treatment planner regarding whether whole dose distributions, including the CTV and PTV, are appropriate in terms of individual risk of BCR after radiotherapy. However, this indicator may be further improved because there were no statistically significant differences in survival curves of certain groups, such as GGs of 3 and 4. For example, determining the optimal number of dosiomic features for calculating the DH score could potentially enhance the prognostic power.

Our findings should be interpreted with caution. Although we merged cohorts from the two sites, the small number of patients with BCR in each group may have reduced the statistical power when comparing the FFBF curves between the high- and low-risk groups. In particular, uncertainty exists in the analysis stratified by the GGs. Therefore, a future multi-institutional study is warranted to identify the patient groups that are truly affected by the dose distribution quality.

In general, a prognostic feature is aggregated from a training cohort, and generalisability is tested using an independent validation cohort. This study combined the two cohorts because the number of patients with BCR in Cohort B was too small for Kaplan–Meier analysis (N=21). If the two dosiomic features used in this study did not work in Cohort B, the separation of the FFBF curves would be underestimated in the integrated cohort. The generalisability of these features should be tested in an independent cohort in future studies.

Previous efforts related to dosiomics have included patients with highly diverse clinical backgrounds, including different clinical stages, histology, treatment modalities, and prescribed doses, to examine the prognostic power of certain dosiomic features.[4–11, 26–28] Consequently, the patient groups that were significantly affected by the dose distribution quality remained unclear. To the best of our knowledge, this is the first study to demonstrate that the effects of dose distribution on prognosis differ depending on patient's background. This highlights the importance of stratified analysis in dosiomics research, even for specific cancer types. In other words, dosiomics may not be useful for all patients; however, it can be considered a novel metric for treatment planning in specific populations.

This study has several limitations. First, the limited number of patients with BCR in each patient group may have affected the results of the Kaplan–Meier analysis. For example, there were no statistically significant differences in patients with GGs of 3 and 4 when using the DH score, although the FFBF curves appeared to differ in those groups. Second, the usefulness of the DH score was not examined in an independent validation cohort. Therefore, it remains unclear whether this is a strong prognostic marker in other institutional cohorts. Stratified analysis renders performing external validation challenging. Third, androgen deprivation therapy status was not considered in the Kaplan–Meier analysis. Finally, the robustness of dosiomic features was not considered in this study. Recent

studies suggested that dosiomic features can be sensitive to changes in dose calculation algorithms, dose grid sizes, and dose cube pixel spacing.[29–32] However, the robustness of filtered dosiomic features such as wavelets for treatment plans in patients with prostate cancer remains unclear and should be clarified in future studies.

CONCLUSION

The dosiomic feature extracted from PTV significantly distinguished high- and low-risk groups in patients with PSA levels >10 ng/ml and GGs of 4 and 5. This indicates that the quality of the planned dose distribution on the PTV may affect the prognosis of patients with poor prognostic factors.

REFERENCES

1. Kuban DA, Tucker SL, Dong L, Starkschall G, Huang EH, Cheung MR, et al. Long-term results of the M.D. Anderson randomized dose-escalation trial for prostate cancer. *Int J Radiat Oncol Biol Phys* 2008;70:67-74. doi: 10.1016/j.ijrobp.2007.06.054.
2. Kattan MW, Zelefsky MJ, Kupelian PA, Scardino PT, Fuks Z, Leibel SA. Pretreatment nomogram for predicting the outcome of three-dimensional conformal radiotherapy in prostate cancer. *J Clin Oncol* 2000;18:3352-9. doi: 10.1200/JCO.2000.18.19.3352.
3. Zelefsky MJ, Kattan MW, Fearn P, Fearon BL, Stasi JP, Shippy AM, et al. Pretreatment nomogram predicting ten-year biochemical outcome of three-dimensional conformal radiotherapy and intensity-modulated radiotherapy for prostate cancer. *Urology* 2007;70:283-7. doi: 10.1016/j.urology.2007.03.060.
4. Liang B, Yan H, Tian Y, Chen X, Yan L, Zhang T, et al. Dosiomics: Extracting 3D spatial features from dose distribution to predict incidence of radiation pneumonitis. *Front Oncol* 2019;9:269. doi: 10.3389/fonc.2019.00269.
5. Wu A, Li Y, Qi M, Lu X, Jia Q, Guo F, et al. Dosiomics improves prediction of locoregional recurrence for intensity modulated radiotherapy treated head and neck cancer cases. *Oral Oncol* 2020;104:104625. doi: 10.1016/j.oraloncology.2020.104625.
6. Adachi T, Nakamura M, Shintani T, Mitsuyoshi T, Kakino R, Ogata T, et al. Multi-institutional dose-segmented dosiomic analysis for predicting radiation pneumonitis after lung stereotactic body radiation therapy. *Med Phys* 2021;48:1781-91. doi: 10.1002/mp.14769.
7. Lee SH, Han P, Hales RK, Voong KR, Noro K, Sugiyama S, et al. Multi-view radiomics and

dosiomics analysis with machine learning for predicting acute-phase weight loss in lung cancer patients treated with radiotherapy. *Phys Med Biol* 2020;65:195015. doi: 10.1088/1361-6560/ab8531.

8. Ren W, Liang B, Sun C, Wu R, Men K, Xu Y, et al. Dosiomics-based prediction of radiation-induced hypothyroidism in nasopharyngeal carcinoma patients. *Phys Med* 2021;89:219-25. doi: 10.1016/j.ejmp.2021.08.009.
9. Gabryś HS, Buettner F, Sterzing F, Hauswald H, Bangert M. Design and selection of machine learning methods using radiomics and dosiomics for normal tissue complication probability modeling of xerostomia. *Front Oncol* 2018;8:35. doi: 10.3389/fonc.2018.00035.
10. Rossi L, Bijman R, Schillemans W, Aluwini S, Cavedon C, Witte M, et al. Texture analysis of 3D dose distributions for predictive modelling of toxicity rates in radiotherapy. *Radiother Oncol* 2018;129:548-53. doi: 10.1016/j.radonc.2018.07.027.
11. XXX (anonymized).
12. Hanks GE, Lee WR, Hanlon AL, Hunt M, Kaplan E, Epstein BE, et al. Conformal technique dose escalation for prostate cancer: Biochemical evidence of improved cancer control with higher doses in patients with pretreatment prostate-specific antigen > or = 10 NG/ML. *Int J Radiat Oncol Biol Phys* 1996;35:861-8. doi: 10.1016/0360-3016(96)00207-6.
13. Zelefsky MJ, Fuks Z, Hunt M, Lee HJ, Lombardi D, Ling CC, et al. High dose radiation delivered by intensity modulated conformal radiotherapy improves the outcome of localized prostate cancer. *J Urol* 2001;166:876-81. Doi: 10.1016/S0022-5347(05)65855-7.
14. Pound CR, Partin AW, Eisenberger MA, Chan DW, Pearson JD, Walsh PC. Natural history of progression after PSA elevation following radical prostatectomy. *JAMA* 1999;281:1591-7. doi: 10.1001/jama.281.17.1591.
15. Spratt DE, Zhang J, Santiago-Jiménez M, Dess RT, Davis JW, Den RB, et al. Development and validation of a novel integrated clinical-genomic risk group classification for localized prostate cancer. *J Clin Oncol* 2018;36:581-90. doi: 10.1200/JCO.2017.74.2940.
16. Roach M 3rd, Hanks G, Thames H Jr, Schellhammer P, Shipley WU, Sokol GH, et al. Defining biochemical failure following radiotherapy with or without hormonal therapy in men with clinically localized prostate cancer: Recommendations of the RTOG-ASTRO phoenix consensus conference. *Int J Radiat Oncol Biol Phys* 2006;65:965-74. doi: 10.1016/j.ijrobp.2006.04.029.
17. XXX (anonymized).
18. Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res* 2017;77:e104-7. doi:

10.1158/0008-5472.CAN-17-0339.

19. Kakino R, Nakamura M, Mitsuyoshi T, Shintani T, Kokubo M, Negoro Y, et al. Application and limitation of radiomics approach to prognostic prediction for lung stereotactic body radiotherapy using breath-hold CT images with random survival forest: A multi-institutional study. *Med Phys* 2020;47:4634-43. doi: 10.1002/mp.14380.
20. Graham MH. Confronting multicollinearity in ecological multiple regression. *Ecology* 2003;84:2809-15. doi: 10.1890/02-3114.
21. Zheng BH, Liu LZ, Zhang ZZ, Shi JY, Dong LQ, Tian LY, et al. Radiomics score: a potential prognostic imaging feature for postoperative survival of solitary HCC patients. *BMC Cancer* 2018;18:1148. doi: 10.1186/s12885-018-5024-z.
22. Epstein JI, Zelefsky MJ, Sjoberg DD, Nelson JB, Egevad L, Magi-Galluzzi C, et al. A contemporary prostate cancer grading system: A validated alternative to the Gleason score. *Eur Urol* 2016;69:428-35. doi: 10.1016/j.eururo.2015.06.046.
23. Ueda Y, Fukunaga JI, Kamima T, Adachi Y, Nakamatsu K, Monzen H. Evaluation of multiple institutions' models for knowledge-based planning of volumetric modulated arc therapy (VMAT) for prostate cancer. *Radiat Oncol* 2018;13:46. doi: 10.1186/s13014-018-0994-1.
24. Kubo K, Monzen H, Ishii K, Tamura M, Kawamorita R, Sumida I, et al. Dosimetric comparison of RapidPlan and manually optimized plans in volumetric modulated arc therapy for prostate cancer. *Phys Med* 2017;44:199-204. Doi: 10.1016/j.ejmp.2017.06.026.
25. Epstein JI, Egevad L, Amin MB, Delahunt B, Srigley JR, Humphrey PA, et al. The 2014 international society of urological pathology (ISUP) consensus conference on gleason grading of prostatic carcinoma: Definition of grading patterns and proposal for a new grading system. *Am J Surg Pathol* 2016;40:244-52. doi: 10.1097/PAS.0000000000000530.
26. Monti S, Xu T, Liao Z, Mohan R, Cella L, Palma G. On the interplay between dosiomics and genomics in radiation-induced lymphopenia of lung cancer patients. *Radiother Oncol* 2022;167:219-25. doi: 10.1016/j.radonc.2021.12.038.
27. Lam SK, Zhang Y, Zhang J, Li B, Sun JC, Liu CY, et al. Multi-organ omics-based prediction for adaptive radiation therapy eligibility in nasopharyngeal carcinoma patients undergoing concurrent chemoradiotherapy. *Front Oncol* 2022;11:792024. doi: 10.3389/fonc.2021.792024.
28. Buizza G, Paganelli C, D'Ippolito E, Fontana G, Molinelli S, Preda L, et al. Radiomics and dosiomics for predicting local control after carbon-ion radiotherapy in skull-base chordoma. *Cancers (Basel)* 2021;13:339. doi: 10.3390/cancers13020339.
29. Placidi L, Cusumano D, Lenkowicz J, Boldrini L, Valentini V. On dose cube pixel spacing pre-

processing for features extraction stability in dosiomic studies. *Phys Med* 2021;90:108-14. doi: 10.1016/j.ejmp.2021.09.010.

30. Adachi T, Nakamura M, Kakino R, Hirashima H, Iramina H, Tsuruta Y, et al. Dosiomic feature comparison between dose-calculation algorithms used for lung stereotactic body radiation therapy. *Radiol Phys Technol* 2022;15:63-71. doi: 10.1007/s12194-022-00651-9.
31. Placidi L, Gioscio E, Garibaldi C, Rancati T, Fanizzi A, Maestri D, et al. A multicentre evaluation of dosiomics features reproducibility, stability and sensitivity. *Cancers (Basel)* 2021;13:3835. doi: 10.3390/cancers13153835.
32. Sun L, Smith W, Kirkby C. Stability of dosiomic features against variations in dose calculation: An analysis based on a cohort of prostate external beam radiotherapy patients. *J Appl Clin Med Phys*. 2023;24:e13904. doi: 10.1002/acm2.13904.

Figure legends

Figure 1. Flowchart of patient selection. BCR, biochemical recurrence; PSA, prostate-specific antigen; RT, radiation therapy; VMAT, volumetric-modulated arc therapy; HIFU, high-intensity focused ultrasound; GG, Gleason grade group.

Figure 2. Kaplan–Meier estimates of FFBF for high- and low-risk BCR groups separated by two predictive dosiomic features and DH score. (A) Patient group with pretreatment PSA level ≤ 10 ng/ml, (B) Patient group with pretreatment PSA level >10 ng/ml.

Figure 3. Kaplan–Meier estimates of FFBF for high- and low-risk BCR groups separated by CTV-derived dosiomic feature, according to GG.

Figure 4. Kaplan–Meier estimates of FFBF for high- and low-risk BCR groups separated by PTV-derived dosiomic feature, according to GG.

Figure 5. Kaplan–Meier estimates of FFBF for high- and low-risk BCR groups separated by DH score, according to GG.

Table 1 Patient characteristics

Characteristics	Integrated cohort (N=721)		
	BCR (n=117)	No-BCR (n=604)	*p* value
Age [years]			0.033
≥ 70	55 (47.0)	349 (57.8)	
< 70	62 (53.0)	255 (42.2)	
T-stage [No. (%)]			< 0.000001
T1a-T2a	40 (34.2)	326 (54.0)	
T2b-T2c	17 (14.5)	121 (20.0)	
≥ T3a	60 (51.3)	157 (26.0)	
Pretreatment PSA [No. (%)]			< 0.0001
≤ 10 ng/ml	34 (29.1)	299 (49.5)	
> 10 ng/ml	83 (70.9)	305 (50.5)	
Gleason grade group [No. (%)]			0.007
1	3 (2.6)	40 (6.6)	
2	21 (17.9)	188 (31.1)	
3	25 (21.4)	122 (20.2)	
4	22 (18.8)	118 (19.6)	
5	46 (39.3)	136 (22.5)	
NCCN risk group [No. (%)]			0.001
Low	0 (0.0)	2 (0.3)	
Intermediate	32 (27.4)	268 (44.4)	
High	85 (72.6)	334 (55.3)	
Status of HTx [No/Yes (%)]	37 (31.6) / 80 (68.4)	181 (30.0) / 423 (70.0)	0.742
PBCR [%, Median (Range)]	41.7 (0.06–100.0)	25.0 (0.00–100.0)	< 0.00001

BCR, biochemical recurrence; PSA, prostate specific antigen; HTx, hormone therapy; PBCR, positive biopsy core rate

Table 2. Univariate and multivariate Cox proportional hazard regression analysis

	Univariate		Multivariate	
Feature or variable	HR (95% CI)	*p* value	HR (95% CI)	*p* value
Age (≥70 vs. <70)	0.70 (0.49-1.01)	0.058		
T-stage (≥T3a vs. <T3a)	2.53 (1.76-3.64)	< 0.000001	1.53 (1.01-2.34)	0.045
Pretreatment PSA (>10 ng/ml vs. ≤10 ng/ml)	2.24 (1.50-3.34)	< 0.0001	1.61 (1.06-2.46)	0.027
Gleason grade group (vs. among Tier1–5)	1.37 (1.18-1.59)	< 0.0001	1.25 (1.07-1.47)	0.005
NCCN risk group (low vs. int vs. high)	2.04 (1.36-3.06)	< 0.001		
Status of HTx (yes vs. no)	1.00 (0.68-1.48)	0.990		
PBCR (≥50% vs. <50%)	2.24 (1.56-3.22)	< 0.0001	1.55 (1.04-2.32)	0.030
PTVandRECT_D95 [Gy] (>Median vs. ≤Median)	1.14 (0.79-1.65)	0.470		
CTV_wavelet-HHH_glrlm_HGLRE (>Median vs. ≤Median)	0.71 (0.49-1.03)	0.073		
PTV_wavelet-HHH_firstorder_Entropy (>Median vs. ≤Median)	0.57 (0.39-0.83)	0.004	0.68 (0.46-1.01)	0.053
DH score (>Median vs. ≤Median)	1.87 (1.27-2.75)	0.002	2.04 (1.38-3.01)	< 0.001

PSA, prostate specific antigen; HTx, hormone therapy; PBCR, positive biopsy core rate; PTVandRECT, PTV ∩ Rectum; glrlm, gray-level run length matrix; HGLRE, HighGrayLevelRunEmphasis; DH, Dosiomics hazard; HR, hazard ratio; CI, confidence interval.