Multi-scale tissue architecture analysis of favorable-risk prostate cancer: Correlation with biochemical recurrence

Miha Pukl1, Sarah Keyes2, Mira Keyes3, Martial Guillaud2, Metka Volavšek4

1Department of Urology, General Hospital Celje, Celje, Slovenia, 2Department of Integrative Oncology, BC Cancer, Vancouver, BC, 3Department of Radiation Oncology, BC Cancer, Vancouver, BC, Canada, 4Institute of Pathology, Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia

Purpose: Prostate cancer (PCa) with biopsy-based grade group (GG) 1 or 2 characteristics has a favorable outcome, yet some cases still progress after radical prostatectomy and present with biochemical recurrence (BCR). We hypothesized that the multi-scale tissue architecture (MSTA) analysis score would correlate with the aggressive PCa phenotype and could be used as a tool for risk assessment to improve the management of patients with favorable-risk PCa.

Materials and Methods: MSTA was evaluated in needle-biopsy samples from 115 patients with favorable-risk PCa, as defined by GG1 and GG2, a prostate-specific antigen (PSA) level of <10 ng/mL, a clinical stage of cT1c to cT2b, and general Gleason GG (GGG) and expert pathologist-assessed GG (EGG). Algorithms based on Voronoi diagrams were applied to all Feulgen-thionin-stained diagnostic areas. One hundred tissue architecture features were calculated and an MSTA score, a linear combination of the most discriminant features, was generated. Correlation of MSTA score with BCR and other clinical variables was investigated.

Results: In a univariate regression model, EGG, clinical stage, and MSTA were significant predictors of BCR (respective p-values: 0.0016, 0.016, and 0.028). Survival analysis showed that patients with a high MSTA score were more likely to experience BCR than were patients with a low MSTA score (odds ratio, 2.9). Combining MSTA with GG assessment resulted in a significant stratification of risk for BCR.

Conclusions: MSTA score could be used as an objective adjunct risk stratification tool to pathologist assessments and could improve the management of patients with favorable-risk PCa.

Keywords: Biochemical recurrence; Image biomarkers; Prostate cancer; Tissue architecture analysis

INTRODUCTION

Prostate cancer (PCa) is a leading cause of cancer death among elderly men in the western world [1]. Because more than half of PCa is low-risk at diagnosis [2], radical treatments, particularly in this low-risk group, may have a detrimental impact on quality of life while causing national health care costs to soar without evidence of improvement in overall survival [3,4]. According to the European Association of Urology (EAU) guidelines, low-risk patients and selected intermediate-risk patients should be offered active surveillance (AS) and be informed about the odds of requiring radi-
cal treatment in the future [5]. However, up to 40% to 50% of men experience biochemical recurrence (BCR) after radical prostatectomy (RP), and 5% to 10% of men with low-risk PCa will have a poor outcome after radical treatment [5–7]. These facts underlie a need for new prognostic models that incorporate objective biomarkers that allow for early and more accurate prediction of BCR after RP, thus optimizing the decision-making process. Confirmation of low-risk biology in a pretreatment setting would allow patients to pursue AS, whereas a high probability of BCR after RP would suggest an aggressive PCa phenotype and call for treatment intensification. To date, the Gleason score (GS) has been the most reliable diagnostic and prognostic tool. The last modification of the GS system redefined GS 6 to 10 into ISUP grade groups (GGs) 1 to 5 [8]. The GS relies on subjective visual assessment of tissue architecture patterns. The subjective assessment results in significant interobserver variability, in particular among general pathologists (GGGs) [9–11]. One approach to objective and reproducible analysis of tissue architecture has been the use of mathematical graph theory, in which the Voronoi diagram serves as a geometric-topological model of prostatic tissue [12–14]. Very little is known, however, about how graph-based analysis of PCa biopsy tissue is associated with outcomes after RP.

We hypothesized that a robust, quantitative evaluation of prostatic tissue architecture could improve the objectivity of the GS assessment and provide an additional measure of the aggressiveness of the lesions, thereby classifying them as having either low or high probability of recurrence. We report here the results for multi-scale tissue architecture (MSTA) analysis as an imaging biomarker of PCa recurrence defined as BCR for biopsy specimens with GG1 and GG2. Additionally, we compare the prognostic value of GGG and expert pathologist (EGG) assessments.

**MATERIALS AND METHODS**

1. **Clinical specimens**

The study was approved by the Slovenian National Medical Ethics Committee (approval number 147/02/14). Biopsy samples were used from 115 patients who underwent RP without lymphadenectomy between 2003 and 2009 at the General Hospital of Celje (GH Celje). Patients with an initial biopsy-based scoring of GG1 and GG2, initial prostate-specific antigen (iPSA) of 10 ng/mL or less, clinical stage cT1c to cT2b, and R0 resection (negative resection margin) were included in the study. None of the patients received neoadjuvant treatment. After surgery, follow-up was performed on patients every 4 months during the first year, every 6 months during the second and third years, and then yearly unless BCR was identified. BCR, the primary endpoint of the study, was defined according to EAU guidelines as PSA above 0.4 ng/mL and rising [5]. Needle biopsies were performed by a group of urologists with 5 cores obtained from each lobe (10 cores altogether). Pathologic evaluation and original PCa diagnosis (GGG) was done by a group of general pathologists from GH Celje. All available samples were reviewed by an experienced uropathologist (MV) from the Institute of Pathology in Ljubljana (EGG). The uropathologist identified samples with the worst Gleason pattern to be used in the study and delineated cancerous tissue in the original H&E-stained slides. Sections of 5 μm were cut from the corresponding paraffin block and stained with Feulgen-thionin, a stoichiometric, DNA-specific stain [15] for architectural analysis. Slides were scanned with a whole-slide Panoramic MIDI scanner (3DHISTECH) using a 20×0.8 NA objective coupled with a color CCD camera.

2. **Quantitative tissue analysis**

1) **Image processing**

Feulgen-thionin-stained images were analyzed using our in-house imaging platform Getafics [16] Regions of interest (ROIs) were delineated on the Feulgen-thionin scanned images based on the areas selected by the pathologist on the H&E-stained slide. When possible, up to two ROIs, all within the same Gleason grade pattern were selected. Within each ROI, individual gland contours were delineated by an experienced histotechnician. Automatic segmentation of nuclei was performed by using advanced algorithms [17] followed by manual editing to refine the results. The best-focused image of each identified nucleus was automatically selected and archived to an image library. The centers of gravity of each nucleus, together with the coordinates of the glands’ contour membranes and the coordinates of the ROIs were saved for MSTA analysis.

2) **MSTA analysis**

MSTA analysis is one of the three main components of our Tissue Phenotype Analysis platform (Fig. 1) and is based on the Voronoi diagram used as a geometric and topological model of prostatic tissue (Fig. 2). Nuclei-based tissue architecture (NbTA) corresponds to the global tissue architecture, which is based on the position of all nuclei, regardless of cell type or location. Gland-based tissue architecture (GbTA) corresponds to the tissue architecture as defined by the position of each gland. The glands’ phenotype (GlandPheno) includes the glands’ architecture, i.e., the spatial organization of nuclei.
within the gland and the gland morphology. The mean and standard deviation of each GlandPheno, NbTA, and GbTA features over the entire ROI were calculated. An exhaustive list and description of these features is given in Supplementary material.

3. Statistical analysis

All statistical analyses were performed by using STATISTICA software version 12 (TISCO). To discriminate between groups, we used a linear forward stepwise discriminant analysis (LDA) with default values of $p$ to enter and $p$ to remove both set to 0.05. Group comparison was performed by using the nonparametric Mann-Whitney U test. Univariate analysis and multivariate Cox proportional hazard modeling was done to assess the prognostic value of MSTA after adjustment for other covariates. Time to BCR was evaluated by using the Kaplan–Meier method.

RESULTS

A total of 119 patients fulfilled the inclusion criteria. From this cohort, 4 were removed owing to a lack of diagnostic material on the slides, thereby reducing the final number of patients to 115. Patients had a median follow-up of 9.4 years. The clinical characteristics of the patients are shown in Table 1. A total of 87 patients (75.7%) had PCa GG1 and 28 patients had GG2 (24.3%). The median preoperative PSA was 4.8 ng/mL (range, 1.04 to 10.0 ng/mL), and the per-
percentage of positive cores (PPC) was 30% (range, 10% to 100%). BCR was detected in 27 patients (23.5%; BCR+ patients). Two patients had local recurrence and two had metastatic recurrence (one EGG1 and one EGG2). One patient died of PCa and five died of other causes. Sixteen patients received salvage therapy for recurrence.

Patients with EGG2 experienced significantly more BCR than did patients with EGG1; only 14.9% of EGG1 patients experienced BCR compared with 50% of EGG2 patients (Yates-corrected Chi-2, p-value=0.0007). A total of 21.5% of GGG1 patients experienced BCR compared with 29.6% of GGG2 patients (Yates-corrected Chi-2, p-value=0.32). The sensitivity and specificity of EGG/GGG to predict BCR are shown in Table 2. EGG and GGG differed not only in terms of predicting BCR but also in time to BCR, as shown in time-to-progression curves plotted as a function of the two GGS (Fig. 3A, B). As expected, there were statistically significant differences between BCR+ and BCR− patients (log-rank test p=3.10^{-5}) in EGG and no significant differences in GGG (p=0.206). Overall agreement between GGG and EGG was present in 93 cases (80.9%).

### 1. Multi-scale tissue architecture analysis

A total of 292 ROIs were analyzed. A total of 27 specimens (23.5%) contained only one ROI and 88 specimens (76.5%) contained more than one ROI.

We compared the respective distribution of the three groups of MSTA features between the two groups of patients (BCR+ and BCR−). Fourteen of the 40 NbTA features, 20 of the 63 GlandPheno features, and 2 of the 33 GbTA features showed a statistically significant difference between BCR+ and BCR− specimens (data not shown). As 52 of the 292 ROIs contained less than 20 glands, GbTA features were excluded from subsequent analyses. To calculate the MSTA discriminant score, one training set and one test set were generated. The test set consisted of 115 ROIs (one per patient, using the largest ROI when more than one ROI was available). The remaining 177 ROIs made up the training set.
Three features were selected by the LDA on the training set: the mean of the roundness factor values of the Voronoi polygons generated from the nuclei positions (NbTA), the mean (over all the glands) of the standard deviation of the distance to the three nearest Delaunay neighbors within each gland (GlandPheno feature), and the standard deviation of the roundness factor values calculated for each gland within the ROI (GlandPheno feature). The resulting MSTA score—the linear combination of these three features—was then calculated for all ROIs of the test set. LDA automatically set the threshold at 0. Patients whose ROI had a negative MSTA score were called MSTA-negative or MSTA- patients. Patients whose ROI had a positive MSTA score were called MSTA-positive or MSTA+ patients.

2. Correlation of MSTA with biochemical recurrence

The proportion of patients with BCR was significantly higher in the MSTA+ group than in the MSTA- group: 32.8% (9 cases of 28 MSTA+) compared with 14.0% (8 cases of 57 MSTA-), respectively, which corresponds to an odds ratio of 29 (relative risk, 103; confidence interval, 29–600). The sensitivity, specificity, negative predictive value, and positive predictive value of MSTA to predict BCR are given in Table 2. The sensitivity of MSTA was significantly better (70%) than the sensitivity of EGG (52%) as well as GGG (30%) but with a much lower specificity, 56% compared with 84% and 78%, respectively.

3. Correlation of MSTA with time to biochemical recurrence

We assessed the difference in time to recurrence between the MSTA- and MSTA+ patients by comparing the Kaplan–Meier curves (Fig. 3C). There was a statistically significant difference (log-rank test p-value of 0.20, 3.10^-5 and 0.003 respectively. Gleason score (GS)6 corresponds to ISUP grade group (GG)1 and GS7 corresponds to GG2.)

4. Multiple regression analysis

We first performed a univariate regression analysis looking at the value of MSTA, GGG, EGG, clinical stage, age, PSA, and PPC with BCR as the primary endpoint. Binary transformation of variables was made for regression analysis. The results of the univariate regression analyses are shown in Table 3. Only EGG, clinical stage, and MSTA were significant predictors of BCR (with respective p-values of 0.0016, 0.016, and 0.028).
We tested different models combining these three variables in a multivariate regression model to predict BCR (Table 4). EGG was the best predictive variable of recurrence in any combination followed by clinical stage and then MSTA. When EGG, clinical stage, and MSTA were combined in the multivariate regression analysis, only EGG remained significant; however, in a model with GGG instead of EGG, MSTA remained significant.

5. Combining Gleason grade with MSTA

We investigated whether the combination of Gleason grades and MSTA could improve the risk stratification of BCR. Using general or EGG’s Gleason grade, we obtained the four groups with increased risk of recurrence (Table 5). Using the combination with EGG, there was a continuous increase in BCR risk from risk group 1 to 4 when not using GGG (Table 5). Using Kaplan–Meier plots, we compared the respective time to recurrence of these four groups of patients (Fig. 4). There was a statistically significant difference between groups A and D (p = 0.004), whereas the difference between groups A and C was close to significant (p = 0.052). On the other hand, there was no statistically significant difference between groups C and D (p = 0.30) or between groups B and C (p = 0.38).

**DISCUSSION**

In this study, we showed that MSTA analysis of GG1 and GG2 specimens provides prognostic information regarding the likelihood of BCR. Our results indicate that the MSTA score can identify aggressive PCa phenotypes and significantly correlates with relapse-free recurrence, whereas assessment by a GGG does not. Furthermore, MSTA showed significant risk stratification of BCR when combined with pathologic assessments. We found that MSTA showed high sensitivity and GG assessments showed high specificity for BCR (Table 2).

**Table 3. Results of the univariate regression analysis to predict biochemical recurrence**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>Std Err</th>
<th>Wald</th>
<th>p-value</th>
<th>Odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.081</td>
<td>0.227</td>
<td>0.128</td>
<td>0.719</td>
<td>1.177</td>
</tr>
<tr>
<td>PPC</td>
<td>-0.24</td>
<td>0.22</td>
<td>1.26</td>
<td>0.26</td>
<td>0.608</td>
</tr>
<tr>
<td>PSA</td>
<td>-0.14</td>
<td>0.236</td>
<td>0.369</td>
<td>0.543</td>
<td>0.750</td>
</tr>
<tr>
<td>cT</td>
<td>0.569</td>
<td>0.23</td>
<td>5.79</td>
<td>0.016</td>
<td>3.120</td>
</tr>
<tr>
<td>GGG</td>
<td>0.298</td>
<td>0.24</td>
<td>1.52</td>
<td>0.21</td>
<td>1.81</td>
</tr>
<tr>
<td>EGG</td>
<td>0.754</td>
<td>0.239</td>
<td>3.19</td>
<td>0.0016</td>
<td>4.519</td>
</tr>
<tr>
<td>MSTA</td>
<td>0.546</td>
<td>0.236</td>
<td>2.407</td>
<td>0.028</td>
<td>2.983</td>
</tr>
</tbody>
</table>

PPC, proportion of positive cores; PSA, prostate-specific antigen; cT, clinical stage; GGG, general Gleason grade group; EGG, expert Gleason grade group; MSTA, multi-scale tissue architecture score.

**Table 4. Results of the multivariate regression analysis to predict biochemical recurrence.** Two models were tested. Model 1 combined GGG with MSTA and clinical stage. Model 2 combined EGG with clinical stage and MSTA

<table>
<thead>
<tr>
<th>Model</th>
<th>Variable</th>
<th>Estimate</th>
<th>Std Error</th>
<th>Wald stat</th>
<th>p-value</th>
<th>Odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Intercept</td>
<td>1.22</td>
<td>0.27</td>
<td>20.04</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GGG</td>
<td>0.235</td>
<td>0.254</td>
<td>0.858</td>
<td>0.354</td>
<td>1.601</td>
</tr>
<tr>
<td></td>
<td>cT</td>
<td>0.507</td>
<td>0.242</td>
<td>3.97</td>
<td>0.036</td>
<td>2.758</td>
</tr>
<tr>
<td></td>
<td>MSTA</td>
<td>0.483</td>
<td>0.242</td>
<td>4.38</td>
<td>0.046</td>
<td>2.631</td>
</tr>
<tr>
<td>2</td>
<td>Intercept</td>
<td>1.066</td>
<td>0.267</td>
<td>15.88</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EGG</td>
<td>0.624</td>
<td>0.251</td>
<td>6.167</td>
<td>0.013</td>
<td>3.484</td>
</tr>
<tr>
<td></td>
<td>cT</td>
<td>0.419</td>
<td>0.250</td>
<td>2.795</td>
<td>0.094</td>
<td>2.312</td>
</tr>
<tr>
<td></td>
<td>MSTA</td>
<td>0.429</td>
<td>0.249</td>
<td>2.967</td>
<td>0.0849</td>
<td>2.362</td>
</tr>
</tbody>
</table>

GGG, general Gleason grade group; MSTA, multi-scale tissue architecture score; EGG, expert Gleason grade group; cT, clinical stage.

**Table 5. Proportion of BCR+ patients in different risk groups according to the combination of Gleason grade and MSTA**

<table>
<thead>
<tr>
<th>Risk groups</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>GGG/MSTA-</td>
<td>15.4% (6/39)</td>
<td>28.6% (12/42)</td>
<td>16.7% (2/12)</td>
<td>43.8% (7/16)</td>
</tr>
<tr>
<td>GGG/MSTA+</td>
<td>8.7% (4/46)</td>
<td>22.5% (9/40)</td>
<td>36.4% (4/11)</td>
<td>56.6% (10/18)</td>
</tr>
</tbody>
</table>

BCR, biochemical recurrence; GGG, grade group. 

- Risk group 1: GG1/MSTA-, 
- Risk group 2: GG1/MSTA+, 
- Risk group 3: GG2/MSTA-, 
- Risk group 4: GG2/MSTA+.

- First row: combination of general Gleason grade group/multi-scale tissue architecture score, 
- Second row: combination of expert Gleason grade group/multi-scale tissue architecture score.

**Fig. 4.** Kaplan–Meier plot of relapse-free recurrence for patients grouped according to expert pathologist-assessed GG (EGG) diagnosis and multi-scale tissue architecture (MSTA) score. (A) EGG1/MSTA-, (B) EGG1/MSTA+, (C) EGG2/MSTA- and (D) EGG2/MSTA+. Using log-rank test, there was a significant difference between groups A and C (p = 0.05), between groups A and D (p = 0.00004) and between groups B and D (p = 0.004), BCR, biochemical recurrence.
ferentiation. Similarly, GS grading incorporates the degree of abnormality of the glandular architecture [18]. Unlike GG, which consists of a combination of primary and secondary pattern assessments (whole-slide examination), our score measures the architectural organization of the worst area on the slide only. It is interesting to note that two of the three features defining the MSTA score measure architectural changes within each individual gland. This could indicate that architectural changes within glands are the first detectable signs of a potentially aggressive phenotype.

In a multivariate regression model that included clinical stage, EGG, and MSTA, EGG and clinical stage remained the only significant variables. Even though the contribution of MSTA was not statistically significant, the p-value was borderline (p=0.08), indicating that MSTA still holds important prognostic information. The MSTA score combined with a pathologist’s assessment could serve as an additional objective tool, particularly in a general hospital setting where EGGs are in high demand. A continuous increase in EGG and MSTA in combination suggests that the MSTA score could further facilitate clinical decision-making in combination with EGG assessment.

For example, if a GGG defines a case as GG2 and if the MSTA score is negative, an EGG could be consulted for additional assessment, which could affect treatment strategy (AS vs. treatment intensification). Patients having a GG1 and an MSTA score could be advised against AS with any assessment. In contrast, patients with a GG1 biopsy and a negative MSTA score may have a very low probability of BCR. These patients could confidently be advised to pursue an AS strategy. Patients with a GG2 biopsy and a positive MSTA score may have a much higher probability of PCa progression in the next 9 years after RP and may benefit from a more intense treatment up front or closer monitoring for recurrence to facilitate an earlier salvage treatment strategy after RP.

Because of its objectivity, an MSTA-based biomarker has the potential to significantly reduce subjectivity and interobserver variability from GG assessments. However, the correlation between MSTA scores from core biopsies and prostatectomy specimens to assess the initial sampling error has yet to be defined. The variability of GG assessment is well-known and is confirmed by the 20% interobserver variability as seen between GGG and EGG in our data. Furthermore, it is interesting to note that GGG is not a significant predictor of BCR. This is due to a significant number of specimens being graded as GG2 by GGGs and then being downgraded to GG1 by an EGG.

1. Comparisons with previous studies

Kayser et al. [14] found that the concept of entropy is closely related to structural entropy and the prognosis of malignancies when applied to PCa. Architectural features are highly correlated with underlying biological processes governing changes in growth pattern characteristics [19]. Graph-based textural features from PCa biopsy tissue have mainly been used for automated determination of GG rather than for investigating the correlation to PCa outcomes [20-22].

Sudbo assessed the prognostic value of quantitative tissue architecture analysis for 30 PCa cases, attempting to differentiate good versus poor prognosis. They extracted 10 features from the Voronoi diagrams and its subgraphs (H&E images). Five of these features showed some potential prognostic value, in particular, the average distance between cells from the Delaunay [13]. Fogarasi et al. [23] used glandular object-oriented image analysis on 1,027 PCa biopsy samples stained with H&E. He found out that glandular tissue architecture combined with preoperative clinical variables can improve the prediction of post-treatment outcome. MacAulay et al. [24]’s group used a combination of nuclear phenotypic characteristics (GOALS features) and cell sociology features, and compared this with GS on Feulgen-thionin-stained tissue microarray samples from a data set of 78 patients who underwent RP (16 of whom had biochemical failure). Prediction of progression was improved by this technique compared with GS.

2. Study limitations

Our study does not take into account issues related to sampling error and tumor heterogeneity. This issue has been extensively addressed [25-27]. Cyll et al. [25] have shown significant intra-prostatic and intra-focal heterogeneity of GS, DNA ploidy, and PTEN expression. They concluded that interpreting the prognostic power of biomarkers with a single sample limits the representativity and validity owing to extensive tumor heterogeneity. However, our study used a homogeneous population with favorable-risk PCa and subsequently lower tumor volume. Therefore, we speculate that the effect of heterogeneity on biomarker prognostic value was not as significant in our study as it was in Cyll’s study where the cohort consisted of mainly locally advanced, high-volume PCa [25]. The small number of BCR+ cases must also be kept in mind. However, the positive results from this study allow us to further explore the significance of these and other imaging biomarkers on a larger sample set, including samples with higher GG and more abundant cancer tissue.
CONCLUSIONS

We have shown that an MSTA analysis of PCa biopsies with GG1 and GG2 can measure changes related to the aggressiveness of the tumor and may eventually help in risk stratification of patients with favorable-risk PCa by identifying patients who are more likely to experience BCR after RP. The significance and eventual implementation of this imaging biomarker in clinical settings will require further investigation.

CONFLICTS OF INTEREST

The authors have nothing to disclose.

ACKNOWLEDGMENTS

There was no specific grant or funding. The study was supported by the Department of Integrative Oncology, BC Cancer Vancouver, Institute of Pathology, Ljubljana, and General Hospital Celje.

We would like to acknowledge Anita Carraro, Jagoda Korbelik, Alan Harisson, and Zhaoyang Chen, all of whom contributed to this study with dedicated work, especially with data acquisition.

AUTHORS’ CONTRIBUTIONS

Miha Pukl contributed to research conception and design, data acquisition, data analysis and interpretation, drafting of the manuscript. Sarah Keyes contributed to data acquisition and original draft preparation. Mira Keyes contributed to data interpretation and critical revision of the manuscript. Martial Guillaud contributed to statistical analyses, drafting of the manuscript, and critical revision of the manuscript. Metka Volavšek contributed to data acquisition, validation, and revision of the manuscript.

SUPPLEMENTARY MATERIAL

Supplementary material can be found via https://doi.org/10.4111/icu.20200018.

REFERENCES

14. Kayser K, Kayser G, Metze K. The concept of structural en-