Prostate Cancer

Adverse Disease Features in Gleason Score 3 + 4 “Favorable Intermediate-Risk” Prostate Cancer: Implications for Active Surveillance

Alessandro Morlacco, John C. Cheville, Laureano J. Rangel, Derek J. Gearman, R. Jeffrey Karnes

Abstract

Background: According to a recent National Comprehensive Cancer Network (NCCN) guidelines update, patients with Gleason score (GS) 3 + 4 prostate cancer (PCA) and “favorable intermediate-risk” (FIR) characteristics might be offered active surveillance (AS). However, the risk of unfavorable disease features and its prediction in this subset of patients is not completely understood.

Objective: To identify the risk of unfavorable disease and potential predictors of adverse outcomes among GS 3 + 4 FIR PCA patients.

Design, setting, and participants: The study included patients with biopsy GS 3 + 4 and otherwise fulfilling the NCCN low-risk definition (prostate-specific antigen [PSA] <10 ng/ml, cT2a or lower) undergoing radical prostatectomy (RP) from 2006 to 2014 at a single institution.

Outcome measurements and statistical analysis: Complete information on PSA, PSA density (PSAD), clinical stage, percentage of positive cores, percentage of maximum surface specimen involvement, and RP pathology were available. GS upgrade and downgrade, non–organ-confined and non–specimen-confined disease, unfavorable disease (pT3–T4 and/or pN1 and/or a pGS 4 + 3) were the outcomes. Statistical analysis included descriptive statistics and multivariable logistic regression.

Results and limitations: A total of 156 patients (13.1%) experienced GS upgrade; 201 (16.9%) were downgraded. Overall, 205 men (17.2%) harbored non–organ-confined disease, and 295 (24.8%) had unfavorable disease. Age (odds ratio [OR]: 1.06), percentage surface involvement (OR: 1.01), and PSAD (OR: 1.83) were the only significant predictors of upgrade. Age (OR: 1.05), clinical stage (OR: 1.74), percentage of positive cores >50% (OR 1.57), percentage of surface area (OR: 1.02), and perineural invasion (OR: 1.89) were significant predictors of unfavorable disease at RP. The retrospective design is a limitation.

Conclusions: AS is a possible option for a subset of men with FIR GS 3 + 4. However, clinical models alone have a limited role in GS upgrade prediction, and alternative tools warrant further investigation.

Patient summary: Patients with Gleason score 3 + 4 at biopsy, low prostate-specific antigen, and low stage might consider the option of active surveillance, but the use of clinical information alone might not be adequate for thorough risk-adapted counseling.

© 2016 European Association of Urology. Published by Elsevier B.V. All rights reserved.
1. **Introduction**

Whether or not active surveillance (AS) should be extended to men with intermediate-risk (IR) prostate cancer (PCa) is a highly debated issue. AS has emerged as a viable option for a considerable number of men to counterbalance the risk of overdiagnosis and overtreatment of clinically indolent disease. Some AS protocols have included a small proportion of men with Gleason score (GS) 7 disease [1,2] and, more recently, National Comprehensive Cancer Network (NCCN) guidelines [3] have listed AS an option for men with “favorable intermediate-risk” (FIR) PCa (GS 3 + 4, percentage of positive cores <50%, and only one additional IR factor). However, most evidence supporting this risk reclassification is indirect [4], and many studies relied on repeated biopsy results rather than on whole-specimen pathology [5], providing a partial perspective of the complete picture. Since its introduction, GS has been regarded as the single most important element in risk assessment [6].

Concerns about undergrading/staging at diagnosis and missing the opportunity for cure if radical intervention is delayed are the main drawbacks to the use of AS in the setting of FIR PCa. Understanding the risk of harboring unfavorable disease among this subgroup may foster more appropriate selection of patients for AS protocols. The aim of this study is to assess the risk of adverse characteristics in men with very favorable risk GS 3 + 4 PCa and to identify potential predictors of unfavorable features.

2. **Patients and methods**

2.1. **Patient selection and evaluation**

All patients with biopsy GS 3 + 4, prostate-specific antigen (PSA) <10 ng/ml, and clinical stage cT2a or lower who underwent radical prostatectomy (RP) at our institution between 2006 and 2014 were included in this retrospective cohort analysis. All had a histologically confirmed diagnosis of PCa (transrectal prostate biopsy) within 3 mo before surgery. Due to the referral nature of our practice, a pathologist at our institution reviewed all the biopsy slides. Thus the biopsy GS score reported represents the result of our internal review [7], and detailed biopsy information was available for all of the patients. RP was performed using an open retropubic, laparoscopic, or robot-assisted approach by experienced urologists. Pelvic lymph node dissection was carried out according to the operating surgeons’ preferences. Surgical specimens were processed and analyzed using a standardized technique, as previously described [8]. The institutional review board approved the study.

2.2. **Data collected**

The clinical and biopsy variables included age at surgery, presurgery PSA, PSA density (PSAD), clinical stage, primary and secondary Gleason grading on biopsy, number of positive cores, number of total cores and percentage of positive/total cores, maximum percentage of surface specimen tumor involvement, presence of perineural invasion, and/or high-grade prostatic intraepithelial neoplasia (HGPIN) in biopsy specimens.

Pathologic variables were primary and secondary GS, pT and pN stage, and surgical margin status. American Joint Committee on Cancer TNM 6th edition (2002) was used for pathologic staging, and GS was assigned according to the 2005 International Society of Urological Pathology (ISUP) modified Gleason scoring system [9]. Upgrade was defined as GS ≥4 + 3 at definitive pathology. Patients with advanced pathologic stage (pT3–T4 and/or pN1) and/or a pGS ≥4 + 3 were considered as “unfavorable disease” at RP [10]. Pathologic stage pT2 or lower and N0 was classified as “organ-confined disease,” whereas “specimen-confined disease” was pT2–pT3a PCa with negative margins (R0) and negative lymph nodes (pN0) [11].

2.3. **Statistical analysis**

In addition to descriptive statistics, we used the chi-square test for comparing categorical variables and the Student t test or Wilcoxon rank sum test for comparison of continuous variables. Clinical data (age, PSAD, with logarithmic transformation, and clinical stage) and detailed biopsy information were analyzed in multivariable prediction models using logistic regression. The percentage of positive cores/total and percentage of specimen surface tumor involvement were considered both as continuous values and categorized as <50% or ≥50%, in accordance with the cut-off suggested by the recent NCCN guidelines update [3]. All tests were two sided with α value of 0.05.

3. **Results**

A total of 1190 patients were included. Table 1 shows the baseline demographic and clinical features of the cohort. Median age at surgery was 62 yr (interquartile range [IQR]: 57–67); PSA was 5.2 ng/ml (IQR: 4.3–6.6). The cT stage was T1 in 812 patients (68%) and T2a in 378 (32%). Median percentage of positive cores was 33% (IQR: 19–50), and median surface area percentage of involvement was 25% (IQR: 10–50); perineural invasion was present in 273 patients (23%).

Table 2 lists the pathologic outcomes of the cohort. The pT stage was 2a–2b in 274 (23%), 2c in 723 (61%), 3a in 144 (12%), and 3b in 48 (4%). Pathologic GS was 6 in 17%, 3 + 4 in

| Table 1 – Demographic and preoperative features of the Gleason score 3 + 4 very favorable intermediate-risk cohort (n = 1190) |
|-------------|-----------------|
| Age at surgery, yr Median (IQR) | 62 (57–67) |
| Preoperative PSA, ng/ml Median (IQR) | 5.2 (4.3–6.6) |
| Clinical T stage (%) | |
| T1c | 812 (68) |
| T2a | 378 (32) |
| Total biopsy cores Median (IQR) | 12 (12–14) |
| Positive biopsy cores | |
| Median (IQR) | 4 (3–6) |
| Positive cores, % Mediann (IQR) | 33 (19–50) |
| Surface area, % Median (IQR) | 25 (10–50) |
| HGPIN (%) | |
| No | 1064 (89) |
| Yes | 126 (11) |
| Perineural invasion (%) | |
| No | 916 (77) |
| Yes | 274 (23) |

HGPIN = high-grade prostatic intraepithelial neoplasia; IQR = interquartile range; PSA = prostate-specific antigen.
70%, 4 + 3 in 10%, 8 in 1.8%, and 9 in 1.1%. GS upgrade was therefore present in 156 patients (13%) and GS downgrade in 201 (17%). A total of 295 men (25%) had unfavorable disease at RP; 205 (17%) had non–organ-confined disease.

Using multivariable analysis, we looked at outcome predictive features. Table 3 shows the results of this analysis. For GS upgrade at definitive pathology, age (odds ratio [OR]: 1.06; 95% confidence interval [CI], 1.03–1.09; \( p = 0.0007 \)), Log2 PSA density (OR: 1.83; 95% CI, 1.17–2.85; \( p = 0.007 \)), and percentage of surface involvement (OR: 1.01; 95% CI, 1.00–1.02; \( p = 0.03 \)) were the only elements reaching statistical significance in the multivariable analysis. Perineural invasion showed an OR of 1.37 (95% CI, 0.91–2.07) without statistical significance (\( p = 0.13 \)). PSAD (OR: 0.67), higher cT (OR: 0.52), percentage of positive cores (OR: 0.98), percentage of surface area (OR: 0.98), and perineural invasion (OR: 0.59) were inversely associated with GS downgrade. Age (OR: 1.05), clinical stage (OR: 1.74 for T2a), percentage of positive cores >50% (OR: 1.57), percentage of surface area (OR: 1.02), and perineural invasion (OR: 1.89) were significant predictors of unfavorable disease at RP. Increasing age, clinical stage, percentage of positive cores >50% (OR: 1.87), percentage of surface area, and the presence of perineural invasion were significantly associated with the risk of non–organ-confined disease. HGPIN did not show significant associations with pathologic outcomes (data not shown).

4. Discussion

Our data suggest that very favorable GS 3 + 4 IR men have a relatively small but unpredictable risk of harboring higher grade disease, and one in four patients may have unfavorable disease features. GS downgrading is also fairly common in this setting, and few clinical features can predict this outcome.

The use of definitive RP pathology as a gold standard is a distinctive feature of the present study, whereas many previous works on risk reclassification were based on repeated biopsy alone. All the pathologic material underwent single-center review, and GS was assigned following ISUP 2005 criteria. The high variability in GS and upgrading criteria between different centers, especially during the first years after the introduction of the ISUP 2005 criteria [9], we believe this is a significant limitation.

Table 3 – Multivariable analysis of outcome predictive features

<table>
<thead>
<tr>
<th></th>
<th>Gleason upgrade</th>
<th></th>
<th></th>
<th>Gleason downgrade</th>
<th></th>
<th></th>
<th>Non–organ confined</th>
<th></th>
<th>Unfavorable disease</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR</td>
<td>95% CI</td>
<td>( p )</td>
<td>OR</td>
<td>95% CI</td>
<td>( p )</td>
<td>OR</td>
<td>95% CI</td>
<td>( p )</td>
<td>OR</td>
</tr>
<tr>
<td>Age</td>
<td>1.06</td>
<td>1.02</td>
<td>1.09 &lt;0.001</td>
<td>0.97</td>
<td>0.94</td>
<td>1.00 0.05</td>
<td>1.03</td>
<td>1.00</td>
<td>1.06 0.04</td>
<td>1.05</td>
</tr>
<tr>
<td>Log2 PSA density</td>
<td>1.83</td>
<td>1.17</td>
<td>2.85 0.007</td>
<td>0.67</td>
<td>0.51</td>
<td>0.89 0.006</td>
<td>1.04</td>
<td>0.79</td>
<td>1.35 0.79</td>
<td>1.14</td>
</tr>
<tr>
<td>cT2a</td>
<td>1.13</td>
<td>0.84</td>
<td>1.52 0.43</td>
<td>0.52</td>
<td>0.32</td>
<td>0.84 0.007</td>
<td>1.95</td>
<td>1.32</td>
<td>2.88 &lt;0.001</td>
<td>1.74</td>
</tr>
<tr>
<td>Positive cores, % (cont)</td>
<td>0.99</td>
<td>0.99</td>
<td>1.01 0.55</td>
<td>0.98</td>
<td>0.97</td>
<td>0.99 0.007</td>
<td>1.02</td>
<td>1.01</td>
<td>1.03 &lt;0.001</td>
<td>1.02</td>
</tr>
<tr>
<td>&gt;50</td>
<td>0.96</td>
<td>0.57</td>
<td>1.62 0.88</td>
<td>0.76</td>
<td>0.45</td>
<td>1.28 0.3</td>
<td>1.87</td>
<td>1.22</td>
<td>2.85 0.004</td>
<td>1.57</td>
</tr>
<tr>
<td>Surface area, % (cont)</td>
<td>1.01</td>
<td>1.00</td>
<td>1.02 0.03</td>
<td>0.98</td>
<td>0.97</td>
<td>0.99 &lt;0.001</td>
<td>1.02</td>
<td>1.01</td>
<td>1.03 &lt;0.001</td>
<td>1.02</td>
</tr>
<tr>
<td>&gt;50</td>
<td>0.87</td>
<td>0.29</td>
<td>2.63 0.8</td>
<td>2.62</td>
<td>0.69</td>
<td>9.94 0.16</td>
<td>1.60</td>
<td>0.58</td>
<td>4.44 0.36</td>
<td>1.09</td>
</tr>
<tr>
<td>Yes</td>
<td>0.70</td>
<td>0.43</td>
<td>1.13 0.15</td>
<td>0.86</td>
<td>0.55</td>
<td>1.36 0.52</td>
<td>1.32</td>
<td>0.87</td>
<td>1.99 0.19</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>1.37</td>
<td>0.91</td>
<td>2.07 0.13</td>
<td>0.59</td>
<td>0.36</td>
<td>0.96 0.03</td>
<td>2.50</td>
<td>1.75</td>
<td>3.56 &lt;0.001</td>
<td>1.89</td>
</tr>
</tbody>
</table>

CI = confidence interval; OR = odds ratio; PSA = prostate-specific antigen; Ref = reference. Boldface shows \( p < 0.05 \).
advantage of the present work. Finally, the presence of detailed core biopsy information enabled us to examine the potential predictive value of these elements. Biopsy information was unable to predict GS upgrading accurately, neither using a continuous nor a dichotomized model.

In a comparable study, Ploussard et al [10] analyzed 2323 patients with localized GS 3 + 4 PCa who underwent RP and applied various AS criteria to test their ability to stratify the risk of unfavorable disease in this group. The overall rate of unfavorable disease was 46%, and for patients without any additional risk factor (PSA level <10 ng/ml, PSAD 0.15 ng/ml per gram, T1c, two or fewer positive cores), the same rate was 19%. Not surprisingly, the study showed large heterogeneity in the results, strictly dependent on the AS criteria used (Prostate Cancer Research International Active Surveillance [PRIAS], Toronto Criteria, Royal Marsden Hospital). However, the authors were not able to use detailed biopsy information, such as percentage of core/specimen surface involvement, which certainly can be relevant when counseling a man about AS eligibility.

In the present work, we used a FIR definition and consequently found a risk of unfavorable disease (24.8%) slightly lower than the 30.5% reported by the authors using the most restrictive criteria. In our study, clinical stage higher than T1 and information on positive cores were predictors of unfavorable disease, confirming the findings of Ploussard et al. We were able to assess the role of age and perineural invasion that also played a significant role in the prediction of unfavorable disease.

Prediction of GS upgrade plays a major role when considering AS, given the known risk of underestimation in biopsy samples. In our series, only age, PSAD, and percentage of surface involvement showed association with Gleason upgrade, whereas in the Ploussard et al clinical stage, PSA, PSAD, and total number of cores more than two reached significance. These discrepancies are probably reflective of a more heterogeneous population in their study, with an overall 30% of GS upgrade (vs 13.1% in our cohort, which was restricted by PSA and cT at the inclusion).

Interestingly, percentage of positive biopsy cores categorized as <50% or ≥50%, and 33% did not show a significant value in predicting GS upgrade either in Ploussard et al or in the present study, suggesting a limited role for the parameter in this setting.

Motamedinia and colleagues [12] also analyzed a low-risk group and reported a risk classification rate of 33% at rebiopsy, with microfoci of GS 4 and the presence of prostatic intraepithelial neoplasia as important risk factors for progression, whereas PSA, PSAD, and the total number of positive cores did not play a significant role. Jain and colleagues [5] found a 30% upgrade at 1-yr follow-up biopsy in a cohort of 592 men with low-risk or FPR disease undergoing AS. T2, higher PSA at biopsy, and higher percentage of involved cores on initial biopsy predicted upgrading. However, only 9.8% of patients were GS 3 + 4, and most of them were all nonsurgical candidates >75 yr of age, reducing the applicability to contemporary AS candidates. These studies were not able to differentiate GS underestimation at initial biopsy and GS progression over time.

The role of age as an independent predictor of upgrading confirms previous findings in the setting of low- and lower IR PCa [13,14]. The precise reason for this phenomenon is incompletely understood, but it might be related to the influence of aging on tumor biology, leading to more aggressive differentiation.

As far as downgrading is concerned, PSAD, clinical stage, and the presence of perineural invasion were inversely associated with the outcome in our cohort, and biopsy core information, although statistically significant, did not show a clinically meaningful association. These results partially compare with Ploussard et al [10], where the overall prevalence was 8.4%, rising to 14.6% when they applied the more restrictive PRIAS criteria. Age, prostate-specific antigen doubling time (PSA DT), and biopsy core information were significant predictors in their series.

Our results fit well in the current debate about extension of AS eligibility. IR PCa has been historically considered an indication for immediate treatment, and the Prostate Cancer Intervention Versus Observation Trial (PIVOT) [15] and the SPCCG-4 [16] trials make quite a strong case for definitive treatment of IR men, with a mortality risk reduction of 31% in the PIVOT and 15.5% in the SPCCG-4. These results are based on older risk-defining criteria and lack the additional information used today. It is therefore debatable whether they apply to contemporary newly diagnosed GS 3 + 4 patients.

Some major prospective AS experiences have enrolled a subset of GS 3 + 4 patients. The University of Toronto protocol, initially including patients with GS 3 + 4 (17%), was eventually restricted to GS 6 after 4 yr, or better IR patients (PSA 10–20 ng/ml and/or GS 3 + 4) were admitted only if they had significant comorbidities and a life expectancy <10 yr. GS >6 and Gleason pattern 4 were independent predictors of deferred treatment in that cohort, but the receipt of ISUP 2005 criteria is not certain in all patients. A limitation of this approach is the inclusion of patients who would not be good candidates for curative treatment, indeed reducing the applicability of the findings to the modern AS concept.

Yamamoto et al [17] analyzed the risk of metastatic progression in 980 patients in the Toronto cohort, 133 of whom (13.6%) were GS 7 (3 + 4 or 4 + 3). Men with Gleason pattern 4 on biopsy have an increased risk for metastatic progression when treated with AS (hazard ratio: 2.9). PSA DT <3 yr and more than three positive cores are additional risk factors. However, we do not know how many of these patients would fit the GS 3 + 4 FIR definition and what outcome would have been observed in this specific subset.

Cooperberg et al [2] reported on 90 patients with IR PCa in their AS cohort, only 27 of whom had GS 3 + 4. The rates of active treatment-free survival did not vary between low-risk and IR patients; in spite of that, 30% of IR men were upgraded at repeated biopsy. Among patients who eventually underwent RP, the upstage rate was 50%. The Prostate Testing for Cancer and Treatment (ProtecT) randomized trial [18], whose results are awaited in 2016, included 20% of GS 7 men randomized to AS, RP, or radiation therapy. These data will shed some more light on long-term
outcomes, but it is uncertain if this trial will be able to add significant upgrade/upstage information using whole-specimen pathology as the gold standard.

The present study is not devoid of limitations. First, our findings come from a cohort of RP-treated men and may only indirectly be generalized to the population of AS men. Despite that, the use of stringent inclusion/exclusion criteria, and the inclusion limited men fit for surgery enhance the applicability of these results to a contemporary subset of men with newly diagnosed FIR GS 3 + 4 PCa. Our results refer mostly to the PCa population during the PSA screening era. Also in the small percentage (11.8%) of patients operated on in 2013–2014, after the formal US Preventive Services Task Force recommendation against PSA screening [19], we did not observe any significant difference in patient features. Perhaps it is too early to see any effect of this change in our cohort; however, our results might be interpreted differently in the context of nonscreened populations. The retrospective nature and the lack of a single pathologist reexamination of all the cases are other potential limitations. However, the single-institution analysis of all the biopsy and RP specimen was carried out by a very limited number of fully trained uropathologists. This enhances internal validity but could make our assumptions slightly less generalizable to centers when the complete review of the material is not a consuetude. Finally, the short follow-up for a considerable number of patients prevented us from presenting survival data.

Our results show that the risk of harboring unfavorable disease among FIR GS 3 + 4 patients may be up to 25%. However, downgrade is not an uncommon event. Overall, these findings underscore the need for appropriate counseling when a management decision is made and suggest that additional risk-assessing tools might be useful to improve the prediction models. Recently, 3-T multiparametric magnetic resonance imaging (mpMRI) using the Prostate Imaging Reporting and Data System has emerged as an independent predictor of downgrade in patients with biopsy GS 3 + 4 [20,21]. It is yet to be determined whether this technique may be helpful in the prediction of unfavorable features. Genetic biomarkers also look promising in this setting. The Genomic Prostate Score (GPS) [22] has proved more efficacious than clinical variables alone to predict adverse tumor features at RP when used on biopsy samples, yet this retrospective cohort included a small fraction (23%) of GS 3 + 4 patients, and the diagnostic accuracy is still not ideal. However, the test has now entered the 2016 NCCN guidelines for patient counseling after the initial diagnosis.

5. Conclusions

Based on the risk of upgrade and upstage at RP, AS is a viable option for a subset of FIR GS 3 + 4, but there is clearly room for improvement of risk-predictive strategies and patient selection. Further investigation with the use of biologically based tests, genomic classifiers, and mpMRI, as well as prospective AS results in this population, are expected to shed more light in this clinical setting.

Author contributions: R. Jeffrey Karnes had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Karnes, Morlacco, Gearman.

Acquisition of data: Morlacco, Rangel.

Analysis and interpretation of data: Morlacco, Rangel, Cheville, Karnes.

Drafting of the manuscript: Morlacco, Gearman, Karnes.

Critical revision of the manuscript for important intellectual content: Morlacco, Cheville, Rangel, Gearman, Karnes.

Statistical analysis: Rangel, Morlacco.

Obtaining funding: None.

Administrative, technical, or material support: None.

Supervision: Karnes, Cheville.

Other (specify): None.

Financial disclosures: R. Jeffrey Karnes certifies that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (eg, employment/affiliation, grants or funding, consultancies, honoraria, stock ownership or options, expert testimony, royalties, or patents filed, received, or pending), are the following: None.

Funding/Support and role of the sponsor: None.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.eururo.2016.08.043.

References


