ORIGINAL REPORT

Transrectal biopsy scheme can predict incorrect histological grading in prostate cancer

M.L. Nieto-Moralesa,*, J. Fernández-Ramosb, L. Pérez-Méndec,d, E. Alventosa-Fernándeza, M.S. Pastor-Santoveña, A. Aguirre-Jaime d

a Servicio de Radiodiagnóstico, Hospital Universitario Nuestra Señora de Candelaria, Santa Cruz de Tenerife, Islas Canarias, Spain
b Servicio de Radiodiagnóstico, Hospital Universitario de Canarias, La Laguna, Santa Cruz de Tenerife, Islas Canarias, Spain
c Enfermedades Respiratorias CIBER, Instituto Carlos III, Madrid, Spain
d Unidad de Investigación, Hospital Universitario Nuestra Señora de Candelaria, Santa Cruz de Tenerife, Islas Canarias, Spain

Received 27 June 2011; accepted 28 May 2012
Available online 30 July 2014

KEYWORDS
Prostate tumor; Transrectal ultrasound-guided prostate biopsy; Ultrasonography

Abstract
Objective: To identify factors that might explain why a prostate with a Gleason score (GS) < 7 in the biopsy specimen can turn out to have a GS ≥ 7 in the surgical specimen.

Material and methods: We compared the GS of biopsy specimens with the GS of surgical specimens in 185 patients who underwent surgery for prostate cancer. We calculated the sensitivity, specificity, and predictive values for the GS of the biopsy specimens. We used Cohen’s kappa to determine the degree of concordance between a GS of < 7 and ≥ 7 for the biopsy specimen and the surgical specimen. Age, a family history of prostate cancer, total prostate-specific antigen (tPSA), digital rectal examination, prostate structure and volume, and the number of biopsy cores (biopsy scheme) were analyzed using multivariable logistic regression.

Results: Histological study of biopsy specimens yielded high sensitivity (98%) but low specificity (49%) for GS ≤ 6 and low sensitivity (35%, 26%) and high specificity (93%, 99%) for GS = 7 and GS ≥ 7, respectively. Cohen’s kappa for the GS from the biopsy and surgical specimens was 0.43 (95% CI = 0.30–0.56). The biopsy scheme was the only predictor of discordance in the GS between the two techniques. Among the other variables included in the model, only tPSA showed a slightly significant association. Taking a scheme with less than 7 cores as a reference, we found no difference with 8 to 9 cores but we did find a difference with 10 to 11 cores and with 12 or more cores, with a prevalence ratio of 0.138 (95% CI = 0.030–0.513) and 0.277 (95% CI = 0.091–0.806), respectively.

Conclusion: The GS of the biopsy depends on the scheme. This factor must be taken into account when choosing a treatment option in patients with low tumor grade in biopsy specimens.

© 2011 SERAM. Published by Elsevier España, S.L.U. All rights reserved.

* Corresponding author.
E-mail address: luisaradazul@yahoo.es (M.L. Nieto-Morales).

Introduction

Patients with localized prostate cancer have several therapeutic options like radical prostatectomy, brachytherapy, external radiotherapy, hormonal therapy or active monitoring.\(^1\) However all these therapies can be insufficient, unnecessary or even be contraindicated in patients with more advanced type of tumors. Transrectal biopsy is the method most widely used to obtain the tumor samples whose analysis allows us to obtain the histological degree of prostate carcinoma. The Gleason system of histological gradation widely accepted for prostate carcinoma categorizes 5 different tumor degrees: Because several prostate carcinomas have more than one degree we granted one degree (from 1 to 5) to the most frequent primary tumor pattern and another degree (from 1 to 5) to the secondary pattern. The addition of these two Gleason degrees makes up the final score (from 2 to 10) known as the Gleason score (GS) or Vacurg. This is how tumors can be very well categorized with a 2-to-6 GS; moderately categorized with a 7-score GS and poorly categorized with an 8-to-10 GS. GS plays an important role in the therapeutic decision-making since it is a prognostic factor both of survival and recurrence. The differences between a 6 and a 7-score GS translates into different therapeutic options particularly in those patients in which a non-surgical therapy will not allow verifying the GS of the piece. A patient not considered a suitable candidate to surgery can only receive brachytherapy with a 6-score GS while in a 7-score GS both brachytherapy and external radiotherapy would be indicated. However mismatches have been reported between the GS value of biopsy and the surgical piece that can be different in 32–73% of cases.\(^2,3\) In an attempt to improve concordance between the biopsy GS and the surgical piece several studies have focused on how to obtain direct samples of tumors through modalities such as Doppler-color, magnetic resonance, contrast-enhanced ultrasounds, elastographies or 3D ultrasounds.\(^4,5\) However these modalities need to be evaluated to obtain the enough degree of validity and reliability so they can be used in daily routine clinical practice.

Trying to explain the existing mismatches between GS and biopsy and the surgical piece some have suggested that these mismatches can be due to borderline cases, or the location of the tumor\(^6,11\) or errors in the biopsy technique in such a way that the sample is taken out of the most representative tumor area while taking into consideration the multifocal character of this type of neoplasm. The level of the prostate-specific antigen (PSA), the percentage of positive cylinders in the biopsy,\(^7\) obesity\(^8\) and the expertise of the pathologist\(^9\) are factors predisposing to such mismatches. The identification of the factors that might influence such mismatches would be of great help to reduce errors in the biopsy when establishing the degree of prostate cancer.

The goal of this study is to be able to identify biopsy-associated technical, clinical or pathological factors that make the biopsy-GS vary from the surgical-GS especially in cases where the <7-score GS turns into a >7-score GS in the surgical piece. This mismatch is called “unacceptable mismatch” due to the difference in the therapeutic approach derived from these different gradings.
Material and methods

Patients

Data from this descriptive observational study were collected retrospectively selecting patients with prostate cancer managed with radical prostatectomy in our institution between January 2004 and December 2008. All patients underwent ultrasound-guided transrectal biopsies and had a preoperative diagnosis of prostate-confined disease. Within the biopsy technical protocol informed consent from each patient was obtained at least 24h before the test. The retrospective nature of the study did not require permission from the ethical committee for the realization of this study. The exclusion criteria were hormone therapy or radiotherapy before surgery, an incomplete clinical history or the radiological tests or histological analysis not taking place in our institution. Out of the clinical histories of patients included in the study researchers could collect the following information: (a) age; (b) family history of prostate cancer; (c) total PSA (tPSA); (d) transrectal palpation data; (e) the visualization of areas and hyper and hypoechoic nodules of the periphery of the gland described in the report of the diagnostic transrectal ultrasound; (f) the ultrasound-measured prostate volume; (g) the number of cylinders obtained in the biopsy; and (h) the GS of the biopsy material and prostatic piece.

Modality

Biopsies were performed by the same team of 4 radiologists with over 5 years experience doing transrectal biopsies and managing the ecogapher using a Toshiba Ecocee® machine equipped with a 5–7 MHz B-Mode endocavitary probe and a 18 G x 200 mm ACECUT® Tru-Cut needle with advancement and 11 mm, 16 mm, and 20 mm-cuts. However the cylinders obtained were categorized in ≤10 mm, from 11 to 15 mm and in ≥16 mm. Both the processing of the biopsy and the surgical piece were done following a standard proceeding using the step-sectioning modality. Considering the number of cylinders obtained the biopsy schemes were categorized as: A (<7), B (8–9), C (10–11), and D (12–15). In our hospital the extended model was based on taking 10–12–14 samples with a randomized sample of 10–12–11 peripheral cylinders associated some times with taking one or several samples of suspicious areas – especially hypoechoic.13 However this model could not always be followed due to the patient’s own intolerance. This is why the A model was established to apply it to these group of patients with a number of samples ≤7–classic sextant biopsy including one sample of the suspicious area among the cylinders extracted. The B Model increases acquisition with respect to the classic sextant biopsy but is still not what we mean by extended model like models C and D. The patients of our sample were divided according to biopsy schemes.

Data analysis

GS concordances between the biopsy and the surgical piece and the proportion of histological grading due to excess or deficit were determined for each and every one of these schemes. On the other hand in order to be able to compare our results with those of other authors14 patients were also categorized according to the biopsy GS and the post-surgical study – in GS <7, GS 7 and GS >7 based on the degree of tumor differentiation. In this study we tagged as ‘unacceptable mismatch’ all those cases in which the biopsy had a GS <7 turned into a GS >7 in the surgical piece given GS ≥7 involved a change in the therapeutical approach.

Data processing started as a description of the characteristics of the sample of patients using summary statistics in line with the type of variable and its distribution. GS contingency tables were obtained from both the biopsy material and the surgical piece for both GS categorizations. For the 1st GS categorization the sensibility, specificity, and predictive value of the biopsy GS positive and negative results were estimated using the surgical piece as the standard of reference. For the 2nd GS categorization and taking the concept of ‘unacceptable mismatch’ into consideration the concordance coefficient free from Cohen’s quadratic-weighted kappa randomized coincidences was estimated to assess the degree of concordance between the biopsy and the surgical proceeding to establish a GS <7 and ≥7 while offering details in line with accurate values. Lastly a logistic regression analysis was done using the multivariate binary model to explore the association between the biopsy scheme and the ‘unacceptable mismatch’ using as other possible predictors factors with differences according to the criterion of ‘unacceptable mismatch’. The result was expressed in terms of the prevalence ratio with its respective intervals of confidence at 95% (95% CI). The level of significance was set in p < 0.05 for bilateral tests of hypotheses. All analyses were done using the SPSS® 17.0 software package (SPSS Inc, Chicago, Illinois, U.S.A.).

Results

The sample consisted of 185 patients whose characteristics of interest are presented in Table 1. The biopsy A Model was used in 38 patients (21%), B Model in 37 patients (20%), C Model in 54 (29%) and D Model in 56 patients (30%). According to the clinical stratification 181 patients (98%) were T2-T3-stage patients. In 155 patients (84%) the biopsy GS was ≤6. Crude concordance between the biopsy GS and the surgical piece was 17 (45%), 20 (54%), 46 (85%), and 45 (80%) for A, B, C, and D biopsy schemes, respectively. The

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Characteristics of the studied sample (n = 185).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature</td>
<td>Valuea</td>
</tr>
<tr>
<td>Age (years)</td>
<td>65 (42–79)</td>
</tr>
<tr>
<td>tPSA (ng/ml)</td>
<td>8.0 (1.8–58.8)</td>
</tr>
<tr>
<td>Presence of symptoms while biopsing (%)</td>
<td>15</td>
</tr>
<tr>
<td>Volume of prostate (cc)</td>
<td>40.4 ± 20.4</td>
</tr>
<tr>
<td>Alterations in prostatic echogenicity (%)</td>
<td>53</td>
</tr>
<tr>
<td>Cylinders obtained (number)</td>
<td>10 ± 3</td>
</tr>
<tr>
<td>Most common GS in biopsy</td>
<td>6</td>
</tr>
<tr>
<td>Most common GS in surgical piece</td>
<td>6</td>
</tr>
<tr>
<td>tPSA: total prostate-specific antigen; GS Gleason summatory.</td>
<td></td>
</tr>
</tbody>
</table>
| a Median (range), average ± SD or %.

Transrectal biopsy scheme

| Table 2 | Gleason summatories of both the biopsy and surgical piece and validity parameters of the biopsy taking the surgical piece as the framework of reference. |
|-----------------|------------------------|-----------------|---------------------|-----------------|
| $\kappa = 43\%$ (95% CI: 30–56%; $p < 0.001$) | GS of the surgical piece | Validity parameters of biopsy (%) |
| | n (%) | 2–6 | 7 | 8–10 | Total | Sensitivity | Specificity | PVPR | PVNR |
| GS of the biopsy | | | | | | | | | |
| 2–6 | 129 (98) | 21 (62) | 6 (32) | 156 (84) | 98 | 49 | 83 | 90 |
| 7 | 3 (2) | 12 (35) | 8 (42) | 23 (13) | 35 | 93 | 52 | 86 |
| 8–10 | 0 (0) | 1 (3) | 5 (26) | 6 (3) | 26 | 99 | 83 | 92 |
| Total | 132 (100) | 34 (100) | 19 (100) | 185 (100) | | | | |

GS: Gleason summatory; PVNR: predictive value of a negative result; PVPR: predictive value of a positive result; $\kappa$: Cohen’s quadratic-weighted kappa for the assessment of concordance/matching free from pure chance between the GS of the biopsy and the GS of the surgical piece.

The histological degree of biopsy was significantly lower than that of the surgical piece with 21 (55%), 16 (43%), 8 (15%), and 10 (18%) for these schemes and in that order ($p < 0.001$). Table 2 shows the matching between the biopsy GS and the pathological anatomy for the categorization 2–6, 7 and >7. The degree of unadjusted matching was 79% and the concordance/matching free of coincidences due to pure chance was 43% (95% CI: 30–56%; $p < 0.001$). In Table 2 we can also see the sensitivity, specificity and predictive values with positive and negative result of the biopsy GS for this categorization.

Table 2 shows the matching between the biopsy GS and the surgical piece was shown in Table 3 together with an unadjusted concordance/matching of 83% free from coincidences due to pure chance of 54% (95% CI: 40–68%; $p < 0.001$). Twenty seven (15%) patients showed GS ‘‘unacceptable mismatch’’. Among them, 24 (89%) showed a GS of 6.

After analyzing all the variables collected the only factor other than the biopsy scheme ($p = 0.001$) that showed some difference for the ‘‘unacceptable mismatch’’ was the tPSA ($p = 0.003$).

In Table 4 we can see the outcomes of the logistic regression analysis with the biopsy scheme as the factor explaining the ‘‘unacceptable mismatch’’ adjusted by the tPSA level reaching some significant trend only ($p = 0.07$). When it comes to the A model both C and D Models reduced the ‘‘unacceptable mismatch’’ with a prevalence ratio close to 0.138 (95% CI: 0.030–0.513) and 0.277 (95% CI: 0.091–0.806), respectively whereas the B Model did not vary the frequency of these mismatches showing a 0.433 prevalence ratio (95% CI: 0.143–1.317) that is to say that the scheme with ≤7 cylinders increased the ‘‘unacceptable mismatch’’ 7 times more than the ≥12 cylinders scheme did.

Table 4 Assessment of the prevalence ratio of ‘‘unacceptable mismatch’’ between the Gleason summatory and the surgical piece as the result of the last iteration in the adjustment of the logistic regression analysis.

<table>
<thead>
<tr>
<th>Biopsy scheme</th>
<th>PR (95% CI)</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheme/model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>Ref</td>
<td>–</td>
</tr>
<tr>
<td>B</td>
<td>0.433 (0.143–1317)</td>
<td>0.140</td>
</tr>
<tr>
<td>C</td>
<td>0.138 (0.030–0.513)</td>
<td>0.003</td>
</tr>
<tr>
<td>D</td>
<td>0.277 (0.091–0.806)</td>
<td>0.019</td>
</tr>
<tr>
<td>tPSA</td>
<td>1038 (0.997–1080)</td>
<td>0.070</td>
</tr>
</tbody>
</table>

Table 3 Concordance/matching of Gleason summatories of biopsy and surgical piece using the cut-off value of 7 as the criterion used when changing the therapeutic attitude.

<table>
<thead>
<tr>
<th>$\kappa = 54%$ (95% CI: 40–68%; $p &lt; 0.001$)</th>
<th>Number (%)</th>
<th>Prostatectomy</th>
</tr>
</thead>
<tbody>
<tr>
<td>GS &lt; 7</td>
<td>GS ≥ 7</td>
<td></td>
</tr>
<tr>
<td>Biopsy</td>
<td>129 (69)</td>
<td>27 (15)</td>
</tr>
<tr>
<td>SG &lt; 7</td>
<td>3 (2)</td>
<td>26 (14)</td>
</tr>
</tbody>
</table>

GS: Gleason summatory; $\kappa$: Cohen’s quadratic-weighted kappa for the assessment of concordance/matching free from pure chance between the biopsy GS and the surgical piece.

Discussion

Our results show that in prostate cancer the matching of tumor grading through ultrasound guided transrectal-biopsy and surgical piece is low. This means that it is necessary to identify the factors that drive this matching. Of all the variables analyzed the biopsy scheme was the only factor capable of predicting the ‘‘unacceptable mismatch’’ or the number of cylinders – to be a little more precise. The more cylinders obtained the fewer mismatches. There are no significant differences between the schemes of 10–11 and ≥12 cylinders.

The sensibility and specificity of the biopsy and the anatomo-pathological study of the surgical piece in our study are similar to those published previously. However we have also seen that the predictive values have been different which seems to rely on the fact that the prevalence ratios of the different degrees of prostate cancer are distinct. In our series 71% of low-degree prostate cancer-patients underwent surgery, 40% in another series and a 19% of moderate differentiation vs 48%. Similarities could be seen in the prevalence ratio of high degree-tumors (10% vs 12%). These interesting results for comparison purposes of both series give us a static view on the degree of matching between the biopsy and the surgical piece since unlike in our series in the meta-analysis the modality used by the radiologist is
not taken into account to obtain the cylinders when doing the biopsy—which is why it is impossible to assess if the proceeding actually modifies the degree of mismatch or not.

In a great number of cancer patients when doing the biopsy for the categorization of tumor the degree obtained was 6 or 7 and it is precisely in this group of patients in which any biopsy grading errors due to excess or deficit play an more important role. We estimate that in one third of patients diagnosed with prostate cancer there is some kind of "unacceptable mismatch" yet not all authors categorize patients exactly the same. For instance Chun et al. called "unacceptable mismatch" all GS changes from ≤ 6 to ≥ 7 or from 7 to ≥ 8.

Even though several studies have examined the clinical and pathological factors capable of making the histological grading of biopsy material lower than the conclusive one few studies have examined what factors might predict the "unacceptable mismatch". Its results have also been variables. In Kulinkar et al.'s sample of low risk-patients (tPSA ≤ 10 ng/ml and a 6 GS value in the biopsy) both the tPSA value and the inexperience of the anatopathologist when interpreting prostatic biopsies played a part in the occurrence of "unacceptable mismatches". Dong et al. said that both the tPSA and the prostatic volume and the cancer volume in the biopsy predict "unacceptable mismatches" in patients scoring 6 in the GS. However for King et al. the prostatic tumor volume of biopsy independently predicting both the organ-confined disease and the biochemical relapse could not predict "unacceptable mismatches", yet despite our univariate study shows that both the tPSA value and the biopsy scheme independently predict "unacceptable mismatches" the multivariate analysis concludes that only the biopsy scheme is the one actually predicting stuff. This latter analysis showed prevalence ratios indicating that the risk of "unacceptable mismatches" between the biopsy GS and the prostatectomy GS for 10–11 and ≥12 cylinder-schemes is lower than with ≥7 cylinder-schemes.

Our study is different from others in that for the multivariate analysis we used a greater number of variables and in that on top of using the number of cylinders obtained in each biopsy we categorized the population into 4 different groups based on the biopsy scheme used which allowed us to evaluate how each and every one of the biopsy schemes used predicts the occurrence of "unacceptable mismatches". Also our study was innovative in that we tried to identify what factors can make the therapy of prostate cancer-patients be incorrect leading to a biopsy-based lower tumor grading in turn leading to the occurrence of "unacceptable mismatches". Based on our results we found out that among these results there are no clinical factors like age, family history or prostatic volume and that the tPSA values borderline statistical significance so it can be a false negative which means we have to take it into account during the process of decision making. In a recent study coinciding with our findings the number of cylinders was the most determinant factor in the prediction of "unacceptable mismatches". However in our study the biopsy scheme that fewer "unacceptable mismatches" was the C Model. This is probably due to the fact that we distinguished 4 different biopsy schemes with a small difference of cylinders among them—also to improve the level of concordance/matching we would have to enlarge the number of samples to reach the point of saturation biopsies.

The main limitations of our study are those of its retrospective nature, the heterogeneity of the samples with respect to the variables collected and the small number of patients showing "unacceptable mismatches". We did not assess the time elapsed between biopsy and the surgery which might have been a confounding variable. Similarly we did not analyze the influence of the operator or anatopathologists yet these were 2 homogeneous groups when it comes to years of experience managing genitourinary conditions so it unlikely that they were predicting factors of "unacceptable mismatches".

In sum it is important that a specialist involved in the therapeutic decision-making process knows the clinical and pathological parameters leading to the occurrence of "unacceptable mismatches". Our results suggest that the number of cylinders obtained is the most determinant predictor in the occurrence of "unacceptable mismatch".

Lacking other methods validated for the preoperative grading of prostate cancer we recommend doing an early biopsy scheme of at least 10 cylinders and if this is not possible we believe that rebiopsing the patient should be taken into consideration in cases of patients with a low grade-GS in the early biopsy who initially are candidates to conservative therapy. It is in these patients in which we believe that other diagnostic imaging modalities to improve the accuracy of needle biopsies may be indicated.

Authorship

1. Manager of the integrity of the study: MLNM, JFR and SPS.
2. Original Idea of the Study: MLNM, JFR and SPS.
3. Study Design: MLNM, JFR, LPM and AAJ.
4. Data Mining: MLNM.
5. Data Analysis and Interpretation: LPM and AAJ.
7. Reference Search: MLNM and JFR
8. Writing: MLNM and EAF.
9. Manuscript critical review with intellectually relevant reviews: EAF, MLNM, JFR, SPS, LPM and AAJ.
10. Final Version Approval: MLNM, JFR, EAF, SPS, LPM and AAJ.

Conflict of interests

Authors reported no conflicts of interests.

Acknowledgement

Special thanks to the Ultrasound and Anatopathology Units of the University Hospital of Canarias, Spain.

References

2. Cookson MS, Flesher NE, Soloway SM, Fair WR. Correlation between Gleason score of needle biopsy and radical


