Needle biopsy size and pathological Gleason Score diagnosis: No evidence for a link

Antonio Cicione, MD;* Francesco Cantiello, MD;* Cosimo De Nunzio, MD;† Andrea Tubaro, PhD, MD;‡ Rocco Damiano, PhD, MD

*Urology Unit, Magna Graecia University of Catanzaro, Catanzaro, Italy; †La Sapienza University, Department of Urology, Sant’ Andrea Hospital, Rome, Italy

Abstract

Background: Biopsy Gleason score (GS), in combination with other clinical parameters, is important to take a therapeutic decision for patients with diagnosis of localized prostate cancer. However, preoperative GS is often upgraded after a radical prostatectomy. Increasing the amount of tissue in prostate biopsy may be a way to avoid this issue. We evaluate the influence of a larger biopsy needle size on the concordance between biopsy and pathological GS.

Methods: We analyzed paired biopsies and prostatectomy specimens from 104 cases of men with clinically localized prostate cancer. At the time of prostate biopsy, the patients were prospectively randomized into two needle groups (16-Gauge [G] and 18G) using a 1:1 ratio. GS concordance was estimated performing kappa statistic testing, overall concordance rate and risk to under grade biopsy GS=6. A logistic regression analysis was performed to evaluate the patients’ characteristics as possible risk factors.

Results: The overall concordance between prostate biopsy and pathological GS was 76.9% and 75.6% (p = 0.875) and the k values were 0.821 and 0.811 (p = 0.424), respectively, for 16G and 18G needle study groups. The risk to undergrade a biopsy GS=6 was 21.1% and 15.4% (p = 0.709) using a 16G and 18G needle, respectively. Age, prostate-specific antigen, prostate volume and needle calibre were not independently associated with a higher risk of GS discordance.

Conclusions: Needle size does not affect the concordance between biopsy and pathological GS. Although GS is not the only way to determine treatment, it is still an unresolved urological issue.

Introduction

Gleason score (GS) on needle biopsy is one of the most important parameters in planning treatment for clinically localized prostate cancer. It is also included in many nomograms to assess cancer recurrence risk after active treat-
GS was based on the core with the highest GS, while in radical prostatectomy specimens with multifocal disease, it was based on the nodule with the highest GS.

All patients received oral quinolone antibiotic (500 mg ciprofloxacin twice daily) the day before TRUS and this was continued for 4 days after biopsy. Moreover, each patient was instructed to perform a pre-biopsy rectal enema. A single operator (AC) performed the extended (12 cores) TRUS using a periprostatic block analgesia\(^6\) and a single pathologist was responsible for histological diagnosis and to allocate GS in biopsy and surgical specimens.

At the end of TRUS, each biopsy sampling was evaluated for the presence of fragmentation and short length core (less than 10 mm); a modified Clavien classification system was adopted to assess biopsy complications.\(^7\)

Prostate volume was determined by transrectal prostate ultrasound through the ellipse formula and a 7.5-MHz endorectal end-fire probe.

Radical retropubic prostatectomy was performed within 8 weeks after prostate biopsy using the technique described by Walsh and colleagues\(^8\) and prostate cancer stage was assigned according to the 2010 TNM.\(^9\)

If patients had a previous prostate biopsy, active anorectal disease, allergy to local anesthetic, androgen deprivation therapy or radiotherapy before the radical prostatectomy, they were excluded from the study analysis. The protocol was approved by our Institutional Review Board and informed consent was obtained by each participant.

**Statistical analysis**

The data were normally distributed. The Student’s t-test and chi-square test were used for ordinal and categorical variables, respectively. A bivariate and multivariate logistic regression was developed to determine whether age, total PSA, PSA ratio, prostate volume and needle calibre were important predictors of GS concordance. A \(p\) value <0.05 was considered statistically significant. All data were analyzed using Statistical Package for Social Science (SPSS) 18.0 for Windows.

**Results**

There were 140 patients in each group (Fig. 1). Of these, 104 patients with localized prostate cancer and who had under-
gone a radical retropubic prostatectomy were analyzed for GS concordance. In total, there were 52 patients in Group 1 (37.1%) with a diagnosis of prostate cancer compared to 54 in Group 2 (38.5%) \((p = 0.762)\). We excluded the last 2 patients from Group 2 (the 18G group) so that both groups were equal in number for comparison.

We tallied patient characteristics (Table 1). There were no significant differences for age and prostate volume between the two groups. The median PSA at diagnosis was 7.8 and 6.86 ng/mL, in Group 1 and 2, respectively. The median PSA ratio was 14.6% and 15.9%, in Group 1 and 2, respectively.

In each study group, most prostate cancers were GS=6 at prostate biopsy: 40 (76.9%) patients in Group 1 and 38 (73.1%) patients in Group 2. Subsequently, pathological GS=6 and GS ≥7 were revealed in 29 (55.7%) and 23 (44.3%) patients and in 30 (57.6%) and 22 (42.4%) patients, in Groups 1 and 2, respectively.

The main pathological prostate cancer stage was organ confined (pT2 in 32 and 35 patients) and all the biopsy samplings had cores longer than 10 mm without fragmentation.

Moreover, the use of a 16G needle appeared to be safe (Table 2): 10 patients (7.3%) and 6 (4.2%), respectively, had Grade 1 and Grade 2 complications compared to 9 (6.3%) and 4 (3%) patients in the 18G group.

We tallied the concordance between biopsy and pathological Gleason scores for each group (Table 3). The K coefficient was 0.821 and 0.811 \((p = 0.424)\), the overall GS concordance rate was 76.9% and 75.6% \((p = 0.875)\), the risk to undergrade a biopsy GS=6 was 21.1% and 15.4% \((p = 0.709)\) using 16G or 18G needle biopsy, respectively. Using the multivariate logistic regression analysis (Table 4), none of the variables evaluated in our study was independently associated with a higher risk of discordance between biopsy and pathological GS.

### Discussion

The Gleason grading system is one of the most important means to predict and choose treatment for men with prostate cancer. GS is indispensable to predict pathological stage, lymph node or distant metastasis.\(^1\)\(^2\) Moreover, in the era of non-invasive treatment options for prostate cancer, such as radiotherapy or active surveillance, where the only tissue sampled is on prostate biopsy, it is important that the grade obtained from the biopsy accurately reflects that of the tumour in the prostate after radical prostatectomy.

However, a risk of discrepancy between biopsy and surgical GS is decrypted and many reasons can cause it, such as sampling error, inter- and intra-observer variability and the pathologist’s experience.\(^1,1\)

Inal and colleagues\(^1,2\) reported a better histological quality sampling using a 16G needle to perform transrectal prostate biopsy and McCormack and colleagues\(^5\) recently showed that the 16G needle does not increase prostate cancer detection.

The goal of this study was to evaluate whether using a larger needle, 16G, to perform a prostate biopsy might increase the concordance between clinical and pathological GS. A 16G needle is about 1.5 times wider than a 18G needle, 16G can increase the quantity and quality of tissue specimen needed, therefore it is reasonable to speculate that using a larger needle, 16G, to perform a prostate biopsy might increase the concordance between clinical and pathological GS.\(^1,2\)

### Table 1. Characteristics of patients with prostate cancer diagnosis at prostate biopsy undergone to radical retropubic prostatectomy

<table>
<thead>
<tr>
<th>Age, years ((\text{Mean} \pm \text{SD}))</th>
<th>16-Gauge needle ((n=52))</th>
<th>18-Gauge needle ((n=52))</th>
<th>(p) value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median ((\text{range}))</td>
<td>67.53 (6.24)</td>
<td>67.11 (7.40)</td>
<td>0.759</td>
</tr>
<tr>
<td>Prostate volume, mL ((\text{Mean} \pm \text{SD}))</td>
<td>57.02 (12.57)</td>
<td>52.67 (12.59)</td>
<td>0.093</td>
</tr>
<tr>
<td>Median ((\text{range}))</td>
<td>55 (32–90)</td>
<td>49 (25–84)</td>
<td>\n</td>
</tr>
<tr>
<td>Median ((\text{range}))</td>
<td>7.07 (2.5–16.6)</td>
<td>7.30 (2.35–13.9)</td>
<td>\n</td>
</tr>
<tr>
<td>Median ((\text{range}))</td>
<td>16 (3–28)</td>
<td>15 (5–39)</td>
<td>\n</td>
</tr>
<tr>
<td>Pathological stage, no. (%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2a</td>
<td>5 (9.6)</td>
<td>4 (7.7)</td>
<td>\n</td>
</tr>
<tr>
<td>T2c</td>
<td>18 (34.6)</td>
<td>22 (42.3)</td>
<td>\n</td>
</tr>
<tr>
<td>6</td>
<td>40 (76.9)</td>
<td>38 (73.1)</td>
<td>0.920</td>
</tr>
<tr>
<td>≥7</td>
<td>12 (22.1)</td>
<td>14 (26.9)</td>
<td>\n</td>
</tr>
<tr>
<td>6</td>
<td>29 (55.7)</td>
<td>30 (57.6)</td>
<td>0.810</td>
</tr>
</tbody>
</table>
| ≥7 | 23 (44.3) | 22 (42.4) | \n
\*Student test; \(\chi^2\) square test.

\(^1\)The last two patients were excluded from analysis to have an equal number to compare.

SD: standard deviation; PSA: prostate-specific antigen.

\[^{1}^{1}\text{The last two patients were excluded from analysis to have an equal number to compare.}\]
the needles (whatever gauge) are not sampling the relevant dominant tumour nodules or the higher grade component. In fact, Rubin and colleagues showed that the GS correlation rate is worse (52.4%) if prostate cancer is diagnosed in less than 1 mm or 5% of one biopsy core and only one GS discrepancy issue.

Although undergrading biopsy GS is the most common problem, overgrading biopsy GS may also occur; we did not evaluate this issue for 2 reasons. The first reason is that we did not consider biopsy GS less than 6 because it is generally considered the cutoff for low-dying risk from prostate cancer and because Gleason pattern one and two are usually rare and seen in the transition zone, so it is unlikely to find a biopsy GS less than 6. Moreover referring GS ≤4 in a biopsy report is not indicated. The second reason is that, although intermediate- and high-GS risk classes are described, we included them in only one because deferred treatment or brachytherapy for prostate cancer is not indicated in cases of biopsy GS ≥7. Moreover the World Health Organization consensus conference suggests reporting the worst GS pattern, even if it is not the predominant or secondary pattern.

No statistical difference was found in sample quality between 16G or 18G needles; this was not the aim of the study, but we needed to assess this to avoid bias. Fragmentation and short core may impede the possibility of correctly assessing the GS in the prostate biopsy sample. Finally, multivariate analysis did not show a correlation between GS concordance and age, prostate volume, PSA total and free serum level, regardless of the type of needle used. Smaller prostate volume was suggested by Sfoungaristos and colleagues as a predictor for upgrading GS, but in our cohort the mean prostate volume size was 57 and 52 cm³ for the both needle groups.

We must acknowledge some important study limitations. We did not evaluate the prostate cancer volume, its position and percent of core tumour involvement. Therefore, we cannot establish if more tissue samples may reduce the sampling error. In particular, the heterogeneous and multifocal nature of prostate cancer may lead to an under- or overgrading error according to where and how much tissue is sampled by the needle. Although our sample number is acceptable for a pilot study, it could be small to assess minute differences in the use of 16G or 18G needle for TRUS.

To date, our prospective study is the first to compare 16G and 18G needles in GS concordance using a standard extended biopsy scheme (and obtaining a quality sample) in patients undergoing their first prostate biopsy. A single biopsy operator and single pathologist, evaluating sample quality and using a 12-core biopsy have improved the strength of this study. Intra- and inter-observed errors were avoided and the use of an extended scheme increased the correlation rate.

Finally, multivariate analysis did not show a correlation between GS concordance and age, prostate volume, PSA total and free serum level, regardless of the type of needle used. Smaller prostate volume was suggested by Sfoungaristos and colleagues as a predictor for upgrading GS, but in our cohort the mean prostate volume size was 57 and 52 cm³ for the both needle groups.

We must acknowledge some important study limitations. We did not evaluate the prostate cancer volume, its position and percent of core tumour involvement. Therefore, we cannot establish if more tissue samples may reduce the sampling error. In particular, the heterogeneous and multifocal nature of prostate cancer may lead to an under- or overgrading error according to where and how much tissue is sampled by the needle. Although our sample number is acceptable for a pilot study, it could be small to assess minute differences in the use of 16G or 18G needle for TRUS.

To date, our prospective study is the first to compare 16G and 18G needles in GS concordance using a standard extended biopsy scheme (and obtaining a quality sample) in patients undergoing their first prostate biopsy. A single biopsy operator and single pathologist, evaluating sample quality and using a 12-core biopsy have improved the strength of this study. Intra- and inter-observed errors were avoided and the use of an extended scheme increased the correlation rate.

Finally, multivariate analysis did not show a correlation between GS concordance and age, prostate volume, PSA total and free serum level, regardless of the type of needle used. Smaller prostate volume was suggested by Sfoungaristos and colleagues as a predictor for upgrading GS, but in our cohort the mean prostate volume size was 57 and 52 cm³ for the both needle groups.

We must acknowledge some important study limitations. We did not evaluate the prostate cancer volume, its position and percent of core tumour involvement. Therefore, we cannot establish if more tissue samples may reduce the sampling error. In particular, the heterogeneous and multifocal nature of prostate cancer may lead to an under- or overgrading error according to where and how much tissue is sampled by the needle. Although our sample number is acceptable for a pilot study, it could be small to assess minute differences in the use of 16G or 18G needle for TRUS.

To date, our prospective study is the first to compare 16G and 18G needles in GS concordance using a standard extended biopsy scheme (and obtaining a quality sample) in patients undergoing their first prostate biopsy. A single biopsy operator and single pathologist, evaluating sample quality and using a 12-core biopsy have improved the strength of this study. Intra- and inter-observed errors were avoided and the use of an extended scheme increased the correlation rate.
Conclusions

In our experience, biopsy needle size does not influence the concordance between biopsy and pathological GS; the undergraduate biopsy GS=6 was the most frequent discordance error. Therefore, in the era of extended biopsy protocol and less invasive treatment for prostate cancer, urologists should be aware that GS upgrading after biopsy is still an unresolved issue. Further studies and technologies should be developed to minimize this possible bias associated with important implications for prostate cancer treatment.

Competing interests: None declared.

This paper has been peer-reviewed.

Acknowledgments: The authors acknowledge Kimberlee Manzi from the University “La Tuscia,” Viterbo, Italy, who provided English editing support.

References


Correspondence: Dr. Antonio Cicione, Urology Unit, Magna Graecia University of Catanzaro, Viale Europa, Germanofo, Catanzaro 88100, Italy; fax: 039-09613647184; acicione@libero.it