# Concordance Between Biopsy and Radical Prostatectomy Specimen Gleason Score in Internal and External Pathology Facilities

ANGELICA A.C. GRASSO, GABRIELE COZZI, CARLOTTA PALUMBO, GIANCARLO ALBO and BERNARDO ROCCO

*Urology Department, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, University of Milan, Milan, Italy*

**Abstract.** *Background: Biopsy Gleason score (bGS) is an important tool for staging and decision making in patients with prostate cancer. Therefore, the data from biopsy should be both reproducible across different pathologists and predictive of the true underlying tumour. We evaluated the agreement between bGS with prostatectomy Gleason score (pGS) comparing patients who underwent prostate biopsy at our hospital with those who did it at an outside facility. Materials and Methods: We retrospectively analyzed patients who underwent robot-assisted radical prostatectomy at our Hospital in 2011 and 2012. Patients were divided depending on the site of prostate biopsy. We calculated a weighted κ statistic to evaluate the concordance from bGS and pGS in the two groups and to evaluate the Gleason score (GS) concordance comparing the proportion of positive cores at biopsy. Results: A total of 124 patients with completed data were identified (70 patients performed biopsy at our institution and 54 at an outside facility). The weighted κ score for GS agreement was 0.40 for our Institution and 0.27 for other facilities. The weighted κ score stratified by biopsy hospital for patients with at least 30% of positive cores was 0.46 for our hospital and 0.42 for other facilities. Conclusion: Internal prostate biopsy predicted better pGS than outside facility biopsy reports. When the percentage of biopsy-positive cores increases, the agreement between bGS and pGS is similar between the two groups. For certain cases in which an outside laboratory biopsy results in equivocal clinical decision, biopsy re-evaluation by internal pathologists can help reveal the true underlying tumor architecture and extension.*

*Correspondence to:* Angelica A.C. Grasso, Urology Department, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, University of Milan, via della Commenda 15, Milano, Italy. Tel: +39 3474535881, e-mail: angelica_grasso@yahoo.it

*Key Words:* Gleason score, prostate cancer, radical prostatectomy.

Gleason score (GS), consisting of primary and secondary grades, is the predominant grading system for prostate cancer (PCa). This system, based on the glandular architecture, defines five histological patterns or grades with decreasing differentiation (1).

Biopsy Gleason score (bGS) is an important tool for staging and decision-making in patients with PCa. In fact, bGS represents a powerful predictor of tumor biology and treatment outcome (2) together with presenting serum prostatic specific antigen (PSA), digital rectal examination and volume of cancer in the biopsy (3-4). GS of 7-10 is associated with worse prognoses, while tumours with GS of 5-6 are associated with lower progression rates after definitive therapy (1).

GS is especially important for patients considering a management alternative to radical prostatectomy (RP), such as active surveillance (AS), brachytherapy, cryotherapy or external beam radiotherapy (5). In these cases, the only tissue available is the one from the biopsy; therefore, any discrepancy between bGS and the actual cancer grading can result in under- or overtreatment (6).

Despite this fact, current evidence showed poor agreement between bGS and RP specimen Gleason score (pGS); particularly, an upgrade in pGS has been demonstrated in 50% of patients with low bGS, while a downgrade has been shown in up to 80% of patients with high bGS (7).

In order to allow for accurate decision-making, the data from biopsy should be both reproducible across different pathologists and predictive of the true underlying tumor.

We, thus, compared the agreement between bGS and pGS in patients who had biopsy in our institution, which is a tertiary referral centre, or in a referring institution, in order to investigate any significant difference in the rate of agreement.

## Materials and Methods

We retrospectively analyzed the records of patients who underwent robot-assisted radical prostatectomy (RARP) at our hospital in 2011 and 2012. We evaluated the agreement between bGS and pGS, comparing patients who underwent prostate biopsy in our Institution

(group 1) with those who had it in a referring institution (group 2). We also evaluated the correlation between agreement of GS and proportion of positive biopsy cores (cut-off value: 30%). Biopsy Gleason score were grouped as ≤6, 7 (subgrouping in 3+4 and 4+3) and ≥8. We calculated a weighted κ coefficient (that assigns different weights to subjects for whom the raters differ by categories) to evaluate the concordance between bGS and pGS in the two groups and to evaluate the GS concordance comparing the proportion of positive cores at biopsy. A value of the κ coefficient, adjusted for agreement expected by chance, of 0 indicates no agreement, <40% marginal agreement, 40-75% good agreement and >75% excellent agreement.

## Results

A total of 124 patients were identified in our dataset. Seventy patients underwent the biopsy at our Institution (group 1), while 54 had it in a referring Institution (group 2). Patients' characteristics are reported in Table I. The mean age was 66 years in both cohorts; mean PSA was 7.2 in patients who underwent the biopsy at our Hospital and 10 in patients coming from other facilities; mean number of cores was 15.6 in group 1 and 14 in group 2.

The weighted κ score for GS agreement was 40% for our institution and 27% for referring institution (Table II).

The weighted κ score, stratified by biopsy hospital for patients with at least 30% of positive cores, was 46% for our hospital and 42% for referring institutions (Table III).

## Discussion

Our results showed a quite good overall agreement between bGS and pGS in our institution, while only marginal agreement was reported when the biopsy was performed in referring institutions. A good agreement has been reported in both cases when at least 30% of the biopsy cores were positive for malignancy; the increment of the proportion of positive cores increased the concordance between the two groups. Differences in the two groups might be possibly explained by differential expertise of pathologists, considering that some referring institutions are less likely to have specialized uropathologists (8). It also must be noted that, in our institution, the same team of pathologists read both the prostate biopsy and the final surgical specimen, possibly contributing to increase concordance (9). A proportion of positive cores >30% increased the concordance between the two groups. One explanation might be that smaller cancer volumes may lead to sampling error and subsequent upgrading (10).

Several studies have investigated the correlation between the bGS and pGS, demonstrating discrepancy and especially upgrading from biopsy to prostatectomy specimens (7). This fact is particularly significant in clinical practice, since an upgrade to a higher GS is a strong and independent predictor of biochemical recurrence after attempted local curative therapy (10).

Many factors can explain the disagreement between bGS and pGS.

PCa is a multifocal disease, which is poorly imaged by ultrasound (11). Thus, transrectal ultrasound (TRUS) guided biopsies of 6-12 cores represent a small, random sampling of the overall prostatic tissue (approximately 0.04% of the average volume of the gland) (1). Therefore, it is very plausible to miss a focus of PCa with a higher GS (12).

Epstein et al. reported some of the more common pathology errors in grading needle biopsy specimens: overcalling Gleason pattern 4 on tangentially sectioned small glands of pattern 3 that mimic poorly formed glands; undercalling cribriform Gleason pattern 4 as pattern 3; undercalling GS 9-10. Furthermore, there are borderline grades between patterns 3-4 and 4-5 (6).

A role may also be played by the presence of a tertiary GS pattern (containing less than 5% of the tissue), which is usually uncommon and was not included in the standard GS (13). When the worst Gleason grade is the tertiary pattern, it should influence the final bGS (1).

Downgrading may be due to the processing of the prostatic tissue or the subjective nature of the GS (7). On the other hand, more biopsy cores with cancer, higher PSA levels and obesity seem to be predictors of upgrading (14). Obese men have larger prostates, which may contribute to poorer sampling (15).

In a nationwide study in Norway, Kvåle et al. showed that the concordance between bGS and pGS was significantly higher in pathology departments that examined >40 RP specimens annually than in departments assessing fewer specimens (16).

The Gleason grading system was revised in 2005 at the consensus conference of the International Society of Urological Pathology (ISUP) (17). Among the consensus statements, some are particularly significant: bGS 2-4 should rarely, if ever, be diagnosed; small amounts (≤5%) of biopsy Gleason pattern 3 should be ignored in the presence of Gleason pattern 4 or 5; in the case of patterns 3, 4 and 5, the GS should be the summation of the most common and the highest-grade pattern. These changes may reduce the difference between high- and low-volume pathology units (16).

Nevertheless, GS upgrading is clinically significant. A bGS <6 is often used as one of the inclusion criteria for AS (18) but cancers with a bGS of 6, that are upgraded to a pGS of 7, have the same pathological stage, margin status and biochemical failure rate as do cancers with bGS and pGS of 7 (16).

In men choosing radiotherapy, the presence of a GS 6 or 7 might affect the type and/or location of the radiotherapy and the possible addition of hormonal therapy. Furthermore, men with upgraded disease were more likely to have adverse pathologic features and biochemical progression (14).

Main limitations of our study were the small sample size, which makes it difficult to generalize the results, and its retrospective nature.

Table I. *Patients' characteristics*.

| | | Our institution (95% CI) | | Referring institution (95%CI) | |
|---|---|---|---|---|---|
| Age | Number | 70 | | 54 | |
| | Mean | 66.63 | (65.12; 68.13) | 66.92 | (65.29; 68.54) |
| | Standard deviation | 6.304 | | 5.952 | |
| | Range | 45-78 | | 54-77 | |
| PSA | Mean | 7.26 | (6.043; 8.475) | 10.04 | (6.43; 13.64) |
| | Standard deviation | 5.063 | | 13.12 | |
| | Range | 1.82;31.81 | | 0.04;92 | |
| Number of cores | Mean | 15.63 | (14.99; 16.26) | 14.05 | (12.89; 15.21) |
| | Standard deviation | 2.6 | | 4.257 | |
| | Range | 8-26 | | 4-36 | |
| Gleason score | 3+3 | 64.28% | | 55.55% | |
| | 3+4 | 15.71% | | 22.22 % | |
| | 4+3 | 12.86% | | 12.96% | |
| | 4+4 | 7.14% | | 7.41% | |
| | 5+4 | --- | | 1.85% | |

CI, Confidence interval; PSA, prostatic specific antigen.

Table II. *Percentage of agreement, κ score and weighted κ score stratified by biopsy hospital*.

| bGS *vs*. pGS | No of patients | No of cores | Agreement | κ (%) | 95% CI | Weighted κ (%) | 95% CI |
|---|---|---|---|---|---|---|---|
| Our institution | 70 | 15.6 (2.5) | 50.7% | 28.7 | (14.2-43.1) | 40.5 | (25.5-55.4) |
| Referring institutions | 54 | 14.0 (4.2) | 46.4% | 17.8 | (0.2-35.4) | 26.9 | (5.9-47.8) |

bGS, biopsy Gleason score; pGS, pathological Gleason score; CI, confidence interval.

Table III. *Percentage of agreement, weighted κ score stratified by group for patients with at least 30% of positive cores*.

| bGS *vs*. pGS | No of patients | No of cores | Agreement (%) | κ (%) | 95% CI | Weighted κ (%) | 95% CI |
|---|---|---|---|---|---|---|---|
| Our institution | 32 | 15.2 (2.1) | 53.1 | 36.8 | 16.3-57.4 | 45.7 | 24.3-67.0 |
| Referring institutions | 34 | 12.9 (2.38) | 50.0 | 27.1 | 6.1-50.1 | 41.9 | 17.8-66.0 |

bGS, Biopsy Gleason score; pGS, pathological Gleason score; CI, confidence interval.

## Conclusion

The pathological GS was better-predicted when prostate biopsy was performed in a referral center. When the percentage of biopsy-positive cores increases (>30%), the agreement between bGS and pGS is similar between the two groups.

In cases in which biopsy results obtained from a referring Center can lead to doubts about the clinical management, specimen re-evaluation by referral Center uropathologists can help revealing the underlying tumour architecture and extension.

## References

1 Montironi R, Mazzuccheli R, Scarpelli M, Lopez-Beltran A, Fellegara G and Algaba F: Gleason grading of prostate cancer in needle biopsies or radical prostatectomy specimens: contemporary approach, current clinical significance and sources of pathology discrepancies. BJU Int *95(8)*: 1146-1152, 2005.

2 Divrik RT, Eroglu A, Sahin A, Zorlu F and Ozen H: Increasing the number of biopsies increases the concordance of Gleason scores of needle biopsies and prostatectomy specimens. Urol Oncol *25(5)*: 376-382, 2007.

3 Rajinikanth A, Manoharan M, Soloway CT, Civantos FJ and Soloway MS: Trends in Gleason score: concordance between biopsy and prostatectomy over 15 years. Urology *72(1)*: 177-182, 2008.

4 D'Amico AV, Desjardin A, Chung A, Chen MH, Schultz D, Whittington R, Malkowicz SB, Wein A, Tomaszewski JE, Renshaw AA, Loughlin K and Richie JP: Assessment of outcome prediction models for patients with localized prostate carcinoma managed with radical prostatectomy or external beam radiation therapy. Cancer 15; *82(10)*: 1887-1896, 1998.

5 Moussa AS, Li J, Soriano M, Klein EA, Dong F and Jones JS: Prostate biopsy clinical and pathological variables that predict significant grading changes in patients with intermediate and high grade prostate cancer. BJU Int *103(1)*: 43-48, 2009.

6 Epstein JI, Feng Z, Trock BJ and Pierorazio PM: Upgrading and downgrading of prostate cancer from biopsy to radical prostatectomy: incidence and predictive factors using the modified Gleason grading system and factoring in tertiary grades. Eur Urol *61(5)*: 1019-1024, 2012.

7 Cohen MS, Hanley RS, Kurteva T, Ruthazer R, Silverman ML, Sorcini A, Hamawy K, Roth RA, Tuerk I and Libertino JA: Comparing the Gleason prostate biopsy and Gleason prostatectomy grading system: the Lahey Clinic Medical Center experience and an international meta-analysis. Eur Urol *54(2)*: 371-381, 2008.

8 Allsbrook WC Jr., Mangold KA, Johnson MH, Lane RB, Lane CG, Amin MB, Bostwick DG, Humphrey PA, Jones EC, Reuter VE, Sakr W, Sesterhenn IA,Troncoso P, Wheeler TM and Epstein JI: Interobserver reproducibility of Gleason grading of prostatic carcinoma: urologic pathologists. Hum Pathol *32(1)*: 74-80, 2001.

9 Cookson MS, Fleshner NE, Soloway SM and Fair WR: Correlation between Gleason score of needle biopsy and radical prostatectomy specimen: accuracy and clinical implications. J Urol *157(2)*: 559-562, 1997.

10 Corcoran NM, Hong MK, Casey RG, Hurtado-Coll A, Peters J, Harewood L, Goldenberg SL, Hovens CM, Costello AJ and Gleave ME: Upgrade in Gleason score between prostate biopsies and pathology following radical prostatectomy significantly impacts upon the risk of biochemical recurrence. BJU Int *108(8 Pt 2)*: E 202-210, 2011.

11 May F, Treumann T, Dettmar P, Hartung R and Breul J: Limited value of endorectal magnetic resonance imaging and transrectal ultrasonography in the staging of clinically localized prostate cancer. BJU Int *87(1)*: 66-69, 2001.

12 Douglas TH, McLeod DG, Mostofi FK, Mooneyhan R, Connelly R, Moul JW and Sesterhenn IA: Prostate-specific antigen-detected prostate cancer (stage T1c): an analysis of whole-mount prostatectomy specimens. Prostate 15; *32(1)*: 59-64, 1997.

13 Mosse CA, Magi-Galluzzi C, Tsuzuki T and Epstein JI: The prognostic significance of tertiary Gleason pattern 5 in radical prostatectomy specimens. Am J Surg Pathol *28(3)*: 394-398, 2004.

14 Freedland SJ, Kane CJ, Amling CL, Aronson WJ, Terris MK and Presti JC Jr, SEARCH Database Study Group: Upgrading and downgrading of prostate needle biopsy specimens: risk factors and clinical implications. Urology *69(3)*: 495-499, 2007.

15 Freedland SJ, Platz EA and Presti Jr JC: Obesity, serum prostate specific antigen and prostate size: implications for prostate cancer detection. J Urol *175(2)*: 500-504, 2006.

16 Kvåle R, Møller B, Wahlqvist R, Fosså SD, Berner A, Busch C, Kyrdalen AE, Svindland A, Viset T, Halvorsen OJ: Concordance between Gleason scores of needle biopsies and radical prostatectomy specimens: a population-based study. BJU Int *103(12)*: 1647-1654, 2009.

17 Epstein JI, Allsbrook WC Jr., Amin MB and Egevad LL, ISUP Grading Committee: The 2005 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma. Am J Surg Pathol *29(9)*: 1228-1242, 2005.

18 Klotz L: Active surveillance for prostate cancer: for whom? J Clin Oncol 10; *23(32)*: 8165-8169, 2005.

*Received June 11, 2014*
*Revised July 15, 2014*
*Accepted July 16, 2014*