

Gleason grading of prostate cancer in needle core biopsies: a comparison of general and urologic pathologists

Jaudah Ahmed Al-Maghrabi,^a Nasir A. Bakshi,^b Hasan M.A. Farsi^c

From the ^aDepartment of Pathology, Faculty of Medicine, King Abdulaziz University, Jeddah, Saudi Arabia; ^bPathology and Laboratory Medicine, King Faisal Specialist Hospital and Research Centre, Riyadh, Saudi Arabia; ^cDepartment of Surgery, King Faisal Specialist Hospital and Research Centre, Jeddah, Saudi Arabia

Correspondence: Dr. Jaudah Ahmed Al-Maghrabi · Department of Pathology, Faculty of Medicine, King Abdulaziz University, PO Box 80205 Jeddah 21589, Saudi Arabia · T: +966-2-640-1000 Ext. 17069, F: 966-2-640-8433 · jalmaghrabi@hotmail.com

Ann Saudi Med 2013; 33(1): 40-44

DOI: 10.5144/0256-4947.2013.40

BACKGROUND AND OBJECTIVES: The Gleason grading of prostate carcinoma (PCa) in needle core biopsies is a major determinant used in management planning. The objective of this study was to evaluate the concordance between general pathologists Gleason grading and that of a urologic pathologist in our community.

DESIGN AND SETTING: Retrospective review conducted at three tertiary care hospitals in Jeddah and Riyadh for all prostatic biopsies with carcinoma from January 2002 to January 2011.

METHODS: Gleason scores assigned by the original pathologist were compared with that of the reviewing urologic pathologists. Biopsies were originally obtained and diagnosed at different referring hospitals and independent laboratories. The kappa test was used to evaluate agreement between the original and review scores.

RESULTS: For 212 biopsies the exact concordance of the Gleason score assigned by the original pathologist and the reviewer was 38.7% (82/212). However, when grouped into the main four-score categories of 2-4, 5-6, 7, and 8 or greater, disagreement was noted in 88 (41.5%) biopsies; 87 were upgraded and 1 was downgraded on review. When grouped into two-score categories of low grade (≤ 6) and high grade (≥ 7), disagreement was noted in 32 (15%) of the biopsies.

CONCLUSION: Gleason grade score shows that there was only slight to fair agreement between outside and review scoring (kappa=0.43). When using only low versus high grade categorization, there was good agreement (kappa=0.69). Almost all of the cases with score disagreement were upgraded on review.

Prostate carcinoma (PCa) is a very complex disease, and the decision to use radical treatment is challenging because it requires a fine balance among expected clinical benefit and potential treatment-related adverse events. Gleason grading is a fundamental factor in making that decision. Gleason grading is also an important prognostic factor in PCa. The correlation between Gleason scores of prostate needle core biopsies and that of prostatectomy specimens is generally good.¹⁻⁷ Several studies have highlighted that the interobserver agreement among pathologists for Gleason scoring varies considerably.⁸⁻¹¹ In the United States, United Kingdom and Japan, poor correlation was demonstrated between general pathologists and urologic pathologists on Gleason scoring.^{8,11,12} The objective of this study is to determine interobserver varia-

tion between the original Gleason score made by non-urologic pathologists at different outside institutions and the score by a urologic pathologist at our institution.

METHODS

A retrospective review was conducted of all prostate biopsies with PCa diagnosis referred to three tertiary care hospitals in Saudi Arabia (January 2002 to January 2011). Biopsies were originally obtained and diagnosed at different hospitals and independent laboratories in the country. These biopsies were reviewed at our institutions either because the patient was referred for management or a pathological consultation was requested. The prostate biopsies included in this study were those reviewed by urologic pathologists at the tertiary care centers. A total of 212 biopsies were included in this study.

The number of cores was variable from different referral institutions, ranging between 2 and 12 cores.

Statistical analysis was performed using SPSS software program, version 16. The Gleason score of the biopsies reported by urologic pathologists was used as a gold standard. The kappa test was used to evaluate agreement between outside pathologists and urologic pathologists. The interpretation of the agreement by kappa value was done by the intervals: kappa <0.20=poor agreement; kappa 0.21-0.40=slight agreement, kappa 0.41-0.60=fair agreement, kappa 0.61-0.80=good agreement, kappa 0.81-1.00=very good.

RESULTS

The Gleason score of all the cases is summarized in **Table 1**. The overall concordance between the original pathologist and the reference pathologist for exact Gleason score for the entire study was 38.7% (82/212). Overall disagreement noted in 130 (61.3%) biopsies. Using the commonly used grade compression model which includes groups 2 through 4, 5 and 6, 7, and 8 through 10, we found that disagreement noted in 88 (41.5%) biopsies; 87 upgraded and 1 downgraded on review. When the Gleason grade score was 2-4, there was total disagreement between the two pathologists, with no cases reported by urologic pathologists. In contrast, outside pathologists reported 39 cases in this grade score (**Table 2**). Thirty-seven cases out of 68 (54%) were given the same Gleason score 5-6, 28 (61%) were given the same Gleason score of 7 and 59 cases (100%) were given the same Gleason score of 8 and more by both pathologists. Overall, Gleason score shows that there was slight to fair agreement between original and review scores (kappa=0.43). When we use only low grade (score 6 and less) versus high grade (score 7 and more) to identify the major disagreement with major clinical significance, we found a disagreement in 32 (15%) biopsies; 31 upgraded and 1 downgraded on review. There was good agreement in this regards between outside and review pathologists (kappa=0.69).

DISCUSSION

Gleason grading is a unique grading system for PCa that was created by Donald Gleason in 1966.^{13,14} This grading system has evolved significantly since its initial description and recently it was updated at a 2005 consensus conference of international experts in urological pathology, under the auspices of the International Society of Urological Pathology (ISUP).¹⁵ The Gleason score in needle biopsies strongly influences decisions for appropriate management for a patient with a new diagnosis of PCa and helps determine many aspects of care; such as whether a patient

Table 1. Initial and review Gleason score for prostatic biopsies with adenocarcinoma.

Gleason score	Number of cases (%) identified by outside pathologist	Number of cases (%) identified by review pathologists
2	5 (2.4)	0 (0)
3	19 (9)	0 (0)
4	15 (7)	0 (0)
5	24 (11.3)	2 (1)
6	44 (20.8)	75 (35.4)
7	46 (21.7)	58 (27.3)
8	27 (12.7)	28 (13.2)
9	25 (11.8)	36 (17)
10	7 (3.3)	13 (6.1)
Total	212 (100)	212 (100)

Table 2. Agreement between general pathologists and urologic pathologists according to grade scheme Gleason score for prostatic biopsies with adenocarcinoma.

	Gleason score	Score by review pathologists				Total
		2- 4	5- 6	7	≥8	
Score by outside pathologist	2-4	0	39	0	0	39
	5-6	0	37	30	1	68
	7	0	1	28	17	46
	≥8	0	0	0	59	59
	Total	0	77	58	77	212

Table 3. Agreement between general pathologists and urologic pathologists using two categories (low grade and high grade) for prostatic biopsies with adenocarcinoma.

	Gleason score	Score by review pathologists		Total
		≤6	≥7	
Score by outside pathologist	≤6	76	31	107
	≥7	1	104	105
	Total	77	135	212

is a candidate for surgery, and whether lymph nodes have to be sampled during surgery or before prostatectomy.¹⁶ The literature suggests that there is discordance between general pathologists and urologic pathologists with respect to grading of PCa.^{8,11,12} Previous studies indicated that general pathologists or community pathologists have a tendency to undergrade prostate cancer.^{8,9,11,12,17,18}

Different schemes for grade “compression” of prostate cancer have been applied to simplify management algorithms.^{16,19-21} However, the commonly used grade compression model includes groups 2 through 4, 5 and 6, 7, and 8 through 10. In the current study same Gleason score was seen in 82 (38.7%) of the cases. Even in those biopsies with agreement on the summation score, there was some discrepancy in identifying the primary pattern. For example in 28 biopsies, the same score of 7 was given by both original and review pathologists, but in 11 there was a discrepancy regarding primary versus secondary patterns in high-grade cancer. Eight biopsies have been initially called score 7 (3+4) and were reassigned to score 7 (4+3) on review and 3 cases have been initially scored as 7 (4+3) and rescored as 7 (3+4) on review.

From the review of the biopsies with score discrepancy, we speculate that the problematic areas include: 1) difficulty of recognition that small microscopic foci of tumor are not equivalent to low-grade carcinoma; 2) difficulty of recognition of Gleason pattern 3 with reference to the importance of invasive growth between benign prostatic acini; 3) confusion between patterns 3 and 4; 4) inappropriate reporting primary versus secondary patterns.

The first problematic area is the misinterpretation of very small foci of tumor as grade 1 or 2. Undergrading of the outside pathologist for Gleason score 6 was the major area of discrepancy. It is well known that a small foci of tumor in prostate needle core biopsy is often challenging to diagnose as cancer; however, after a diagnosis of cancer is made, these are usually assigned to Gleason pattern 3. This low volume cancer was frequently scored as low-grade cancer in the original scoring (Gleason pattern 1 and 2) and the majority of them rescored on review as grade 3 in this study (Figure 1A-C), which is in keeping with a previous study.¹⁸ Pathologists have to be aware that a small focus of tumor is not necessarily a low-grade cancer and the same criteria of Gleason system are applicable to them.

The second problematic area is the misdiagnosis of infiltrative tumor as a low-grade prostate cancer (Gleason pattern 1 and 2; combined score, 2-4) even in the presence of large areas of cancer in the biopsies. For the last two decades, education on Gleason grading has centered on this concept and several articles have appeared in the pathology literature addressing this issue.¹⁸⁻²⁰ The conclusion of the education is that Gleason pattern 1 and 2 is very unusual in prostate needle biopsies. In the 2005 consensus conference of international experts in urological pathology, the consensus was that a Gleason score 4 on needle biopsy should be made rarely, if ever.¹⁵ The consensus conference cautioned

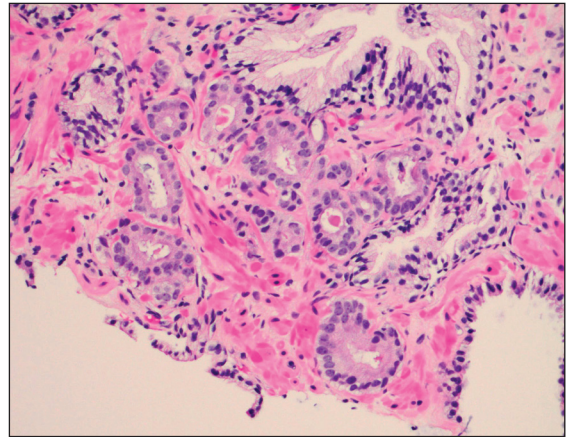


Figure 1A. Microscopic section of prostate biopsy shows a small focus of prostate carcinoma that was scored initially as 1+1 and rescored on review as 3+3, the number of acini is limited, however they are infiltrative and of variable size (hematoxylin-eosin, original magnification 200×).

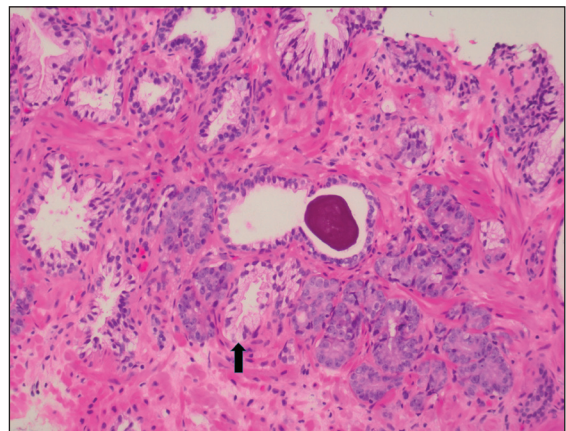


Figure 1B. Microscopic section of prostate biopsy shows that the glands are infiltrating and variable in size and shape. Benign glands (arrow) are seen in between the malignant acini. These are features of Gleason pattern 3. The case was initially scored as 1+2 (hematoxylin-eosin, original magnification 200×)

that a Gleason score 4 on a prostatic needle biopsy is a diagnosis that general pathologists should almost never make without consultation.¹⁵ It was also a consensus that also a Gleason score of 1+1=2 is a grade that should not be diagnosed regardless of the type of specimen (biopsies or prostatectomy), with extremely rare exception.¹⁵

Regarding the third problematic area, there was clear discrepancy between reporting grades 3 and 4. Practically the major area of confusion is the undergrading of fused glands as grade 3 instead of 4. Clearly, fused glands should be assigned for Gleason 4 according to Gleason grading system (Figure 1D). The reason for

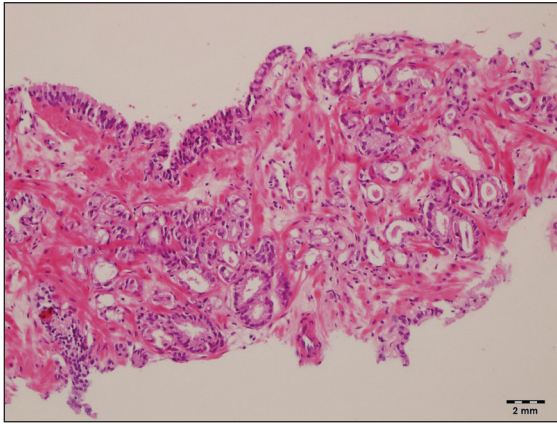


Figure 1C. Microscopic section of another prostate biopsy with initial score of 2+2 and rescored on review as 3+3, the malignant acini are of variable size and shape with an infiltrative pattern (hematoxylin-eosin, original magnification 200 \times).

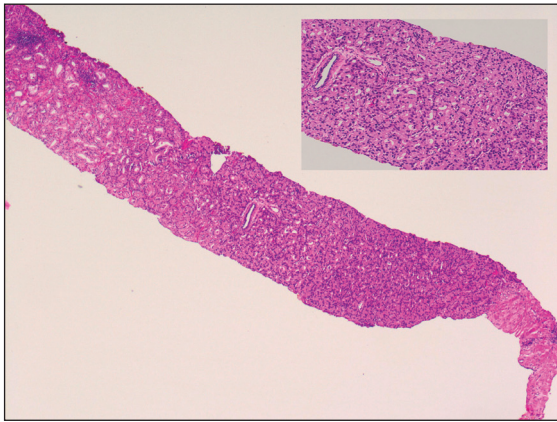


Figure 1E. Microscopic section of prostate biopsy with initial score 3+4 and rescored as 4+3, pattern 4 is predominating (represent >80% of the involved area) with many fused glands associated with a smaller component of pattern 3. The insert shows pattern 4 component. (hematoxylin-eosin, original magnifications 40 \times , inset 200 \times).

this downgrading is either due to the difficulty of recognizing this as pattern 4 or due to ignoring it because of its small percentage. However, it was recommended by many urologic pathologist and was also clearly stated in the ISUP consensus that a high-grade tumor of any quantity on needle biopsy should be included within the Gleason score, while in the setting of high-grade cancer lower-grade patterns should be ignored if they occupy less than 5% of the area of the tumor.¹⁵

The fourth problematic area is reporting primary versus secondary patterns. In this study, among cases that were given similar summation Gleason scores by both an original and reviewer pathologist, 11 cases

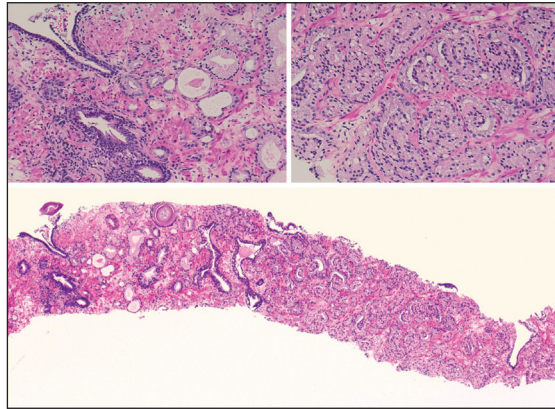


Figure 1D. Microscopic section of prostate biopsy with initial score of 3+3 and rescored on review as 4+3, the malignant acini reveal a clear fusion. The insert on left side represent the small focus that is consider as pattern 3 and the insert on the right side shows pattern 4, which is predominating (hematoxylin-eosin, original magnifications 40 \times , inset 200 \times).

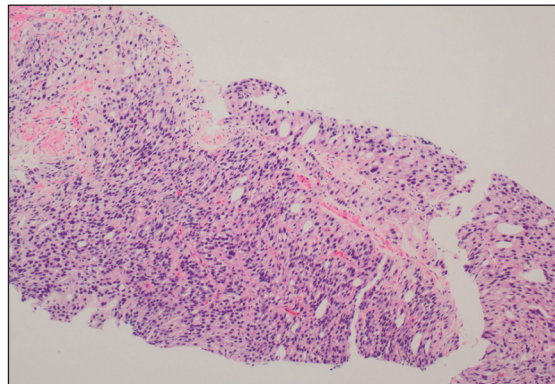


Figure 1F. Microscopic section of prostate biopsy with initial and review score of 4+4 (hematoxylin-eosin, original magnifications 200 \times).

showed a discrepancy regarding primary versus secondary patterns in high-grade cancer. Eight cases that have been called score 7 (3+4) were reassigned to score 7 (4+3) on review (**Figure 1E**). It has been demonstrated that for patients with Gleason score 7 PCa, those with 4+3 cancers have a worse prognosis than those with 3+4 cancers.²¹⁻²³

When using only low grade (score 6 and less) versus high grade (score 7 and more) to identify the major disagreement with major clinical significance, we found disagreement in 32 (15%) biopsies; 31 were upgraded and 1 was downgraded on review. There is good agreement in this regards between the referring and the reviewing pathologist (kappa=0.69). Interestingly, all the cases diagnosed with score >8 by a referring pathologist were

assigned to the same category on review (Figure 1F).

It is therefore clear that areas of discrepancy in Gleason grading system between the practicing pathologists and expert urologic pathologists are quite common.^{8,9,11,12,17,18} These problems of interpretation in Gleason grade usually leads to undergrading of prostate cancer by general pathologists.

In conclusion, the Gleason grade score shows that there was only a slight to fair agreement between outside and review pathologists ($\kappa=0.43$). It is clear that general pathologists in our region have a tendency to undergrade prostate cancer, as has been reported in other countries.^{8,11,12} Educational efforts to improve Gleason grading are recommended for all pathologists involved in evaluation of prostate biopsies. This edu-

cation can be achieved by participation in educational courses at meetings and by participation in tutorial programs including web-site programs.²⁴⁻²⁹ In addition, consultation from tertiary centers is recommended in difficult cases. Whole slide digital imaging is a possible way in which pathologists may seek a second opinion consultation on challenging cases.³⁰ It is also recommended for pathologists to follow the international guidelines of the Gleason grading system and to be updated on this issue to avoid underestimation of the cancer aggressiveness and subsequent mismanagement by urologists. If at all possible, and before embarking on major surgery with potential significant morbidity, the urologist is advised to have the prostate needle biopsy reviewed by a urologic pathologist.

REFERENCES

1. Arellano L, Castillo O, Metrebian E. [Concordance of Gleason histological scoring for prostatic cancer in needle biopsies and the surgical piece obtained during radical prostatectomy]. *Rev Med Chil* 2004 Aug;132:971-8.
2. Cookson MS, Fleshner NE, Soloway SM, Fair WR. Correlation between Gleason score of needle biopsy and radical prostatectomy specimen: accuracy and clinical implications. *J Urol* 1997 Feb;157:559-62.
3. Divrik RT, Eroglu A, Sahin A, Zorlu F, Ozen H. Increasing the number of biopsies increases the concordance of Gleason scores of needle biopsies and prostatectomy specimens. *Urol Oncol* 2007 Sep;25:376-82.
4. Kvale R, Moller B, Wahlqvist R, Fossa SD, Berner A, Busch C, et al. Concordance between Gleason scores of needle biopsies and radical prostatectomy specimens: a population-based study. *BJU Int* 2009 Jun;103:1647-54.
5. Nayyar R, Singh P, Gupta NP, Hemal AK, Dogra PN, Seth A, et al. Upgrading of Gleason score on radical prostatectomy specimen compared to the pre-operative needle core biopsy: An Indian experience. *Indian J Urol* 2010 Jan;26:56-9.
6. Antunes AA, Leite KR, Dall'Oglio MF, Cury J, Srougi M. The effect of the number of biopsy cores on the concordance between prostate biopsy and prostatectomy Gleason score: a prostate volume-controlled study. *Arch Pathol Lab Med* 2008 Jun;132:989-92.
7. May M, Brookman-May S, Lebentrau S, Gilfrich C, Loy V, Theissig F, et al. [Concordance of the Gleason score in prostate multibiopsy and definitive histology]. *Aktuelle Urol* 2010 May;41:184-92.
8. Allsbrook WC, Jr., Mangold KA, Johnson MH, Lane RB, Lane CG, Epstein JI. Interobserver reproducibility of Gleason grading of prostatic carcinoma: general pathologist. *Hum Pathol* 2001 Jan;32:81-8.
9. Allsbrook WC, Jr., Mangold KA, Johnson MH, Lane RB, Lane CG, Amin MB, et al. Interobserver reproducibility of Gleason grading of prostatic carcinoma: urologic pathologists. *Hum Pathol* 2001 Jan;32:74-80.
10. McLean M, Strigley J, Banerjee D, Warde P, Hao Y. Interobserver variation in prostate cancer Gleason scoring: are there implications for the design of clinical trials and treatment strategies? *Clin Oncol (R Coll Radiol)* 1997;9(4):222-5.
11. Oyama T, Allsbrook WC, Jr., Kurokawa K, Matsuda H, Segawa A, Sano T, et al. A comparison of interobserver reproducibility of Gleason grading of prostatic carcinoma in Japan and the United States. *Arch Pathol Lab Med* 2005 Aug;129:1004-10.
12. Renshaw AA, Schultz D, Cote K, Loffredo M, Ziemia DE, D'Amico AV. Accurate Gleason grading of prostatic adenocarcinoma in prostate needle biopsies by general pathologists. *Arch Pathol Lab Med* 2003 Aug;127:1007-8.
13. Bailar JC, III, Mellinger GT, Gleason DF. Survival rates of patients with prostatic cancer, tumor stage, and differentiation--preliminary report. *Cancer Chemother Rep* 1966 Mar;50:129-36.
14. Gleason DF. Classification of prostatic carcinomas. *Cancer Chemother Rep* 1966 Mar;50:125-8.
15. Epstein JI, Allsbrook WC, Jr., Amin MB, Egevad LL. The 2005 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma. *Am J Surg Pathol* 2005 Sep;29:1228-42.
16. Maygarden SJ, Pruthi R. Gleason grading and volume estimation in prostate needle biopsy specimens: evolving issues. *Am J Clin Pathol* 2005 Jun;123 Suppl:S58-S66.
17. Griffiths DF, Melia J, McWilliam LJ, Ball RY, Grigor K, Harnden P, et al. A study of Gleason score interpretation in different groups of UK pathologists; techniques for improving reproducibility. *Histopathology* 2006 May;48:655-62.
18. Steinberg DM, Sauvageot J, Piantadosi S, Epstein JI. Correlation of prostate needle biopsy and radical prostatectomy Gleason grade in academic and community settings. *Am J Surg Pathol* 1997 May;21:566-76.
19. Epstein JI. Gleason score 2-4 adenocarcinoma of the prostate on needle biopsy: a diagnosis that should not be made. *Am J Surg Pathol* 2000 Apr;24:477-8.
20. Iczkowski KA, Bostwick DG. The pathologist as optimist: cancer grade deflation in prostatic needle biopsies. *Am J Surg Pathol* 1998 Oct;22:1169-70.
21. Chan TY, Partin AW, Walsh PC, Epstein JI. Prognostic significance of Gleason score 3+4 versus Gleason score 4+3 tumor at radical prostatectomy. *Urology* 2000 Nov 1;56:823-7.
22. Lau WK, Blute ML, Bostwick DG, Weaver AL, Sebo TJ, Zincke H. Prognostic factors for survival of patients with pathological Gleason score 7 prostate cancer: differences in outcome between primary Gleason grades 3 and 4. *J Urol* 2001 Nov;166:1692-7.
23. Rasiah KK, Stricker PD, Haynes AM, Delprado W, Turner JJ, Golovsky D, et al. Prognostic significance of Gleason pattern in patients with Gleason score 7 prostate carcinoma. *Cancer* 2003 Dec 15;98:2560-5.
24. Egevad L. Reproducibility of Gleason grading of prostate cancer can be improved by the use of reference images. *Urology* 2001 Feb;57:291-5.
25. Kronz JD, Silberman MA, Allsbrook WC, Epstein JI. A web-based tutorial improves practicing pathologists' Gleason grading of images of prostate carcinoma specimens obtained by needle biopsy: validation of a new medical education paradigm. *Cancer* 2000 Oct 15;89:1818-23.
26. Bova GS, Parmigiani G, Epstein JI, Wheeler T, Mucci NR, Rubin MA. Web-based tissue microarray image data analysis: initial validation testing through prostate cancer Gleason grading. *Hum Pathol* 2001 Apr;32:417-27.
27. Helin H, Lundin M, Lundin J, Martikainen P, Tammela T, Helin H, et al. Web-based virtual microscopy in teaching and standardizing Gleason grading. *Hum Pathol* 2005 Apr;36:381-6.
28. Kronz JD, Silberman MA, Allsbrook WC, Jr., Bastacky SI, Burks RT, Cina SJ, et al. Pathology residents' use of a Web-based tutorial to improve Gleason grading of prostate carcinoma on needle biopsies. *Hum Pathol* 2000 Sep;31:1044-50.
29. Mulay K, Swain M, Jaiman S, Gowrishankar S. Gleason scoring of prostatic carcinoma: impact of a web-based tutorial on inter- and intra-observer variability. *Indian J Pathol Microbiol* 2008 Jan;51:22-5.
30. Rodriguez-Urrego PA, Cronin AM, Al-Ahmadie HA, Gopalan A, Tickoo SK, Reuter VE, et al. Interobserver and intraobserver reproducibility in digital and routine microscopic assessment of prostate needle biopsies. *Hum Pathol* 2011 Jan;42:68-74.