

COMPARATIVE STUDY BETWEEN PROSTATIC CARCINOMA G1-G3 AND GLEASON GRADING SYSTEMS

B. ZAHARIA*, I. E. PLEȘEA*, S. D. ENACHE**, P. TOMESCU***, A. PĂNUȘ***, O. T. POP*,
P. BADEA****, A. KOŽOKIĆ*****

Department of Pathology, University of Medicine and Pharmacy of Craiova; **Department of Pathology, Emergency County Hospital, Craiova; *Department of Urology, University of Medicine and Pharmacy of Craiova; ****Department of Biostatistics, University of Medicine and Pharmacy of Craiova; *****City Hospital, Kladovo, Republic of Serbia and Montenegro*

Summary. The need of an accurate and prognostically valuable method for grading the prostatic carcinoma resulted in, in the last decades, a large number of systems, none of them succeeding in being unanimously accepted by the pathologists family. The authors selected and reassessed using Gleason system a group of 221 cases with prostatic adenocarcinoma previously diagnosed using a combined G1-G3 system. The results were converted from Gleason system to G1-G3 system and then compared with the initial diagnostics. The group was divided, following the type of surgical procedure used for drawing the tissue, into three groups: transvesical prostatectomies (TVP), transurethral resections (TUR) and needle biopsies (B). The concordance between the two systems was good, even the number of TUR and B specimens was significant (85% of the studied group) showing that any of them can be successfully used in current practice, depending only on pathologist's experience.

Key words: prostate, carcinoma, grading system, Gleason.

INTRODUCTION

The histopathologic diagnostic of prostatic adenocarcinoma is established using one or more of the following assessment criteria, which can be grouped in two main categories:

- **morphologic features of malignant cells** (cytoplasm features, nuclear changes which define the grade of anaplasia and ultrastructural characteristics).
- **features of tumoral architecture.**

The grading of prostatic epithelial malignancies is still difficult because of at least two reasons:

- most of the tumors show **more than one single histologic aspect**, this unique diversity being the result of a wide range of histologic patterns.
- the assessment criteria for cellular anaplasia and glandular differentiation are not clearly defined yet.

A lot of grading systems were proposed during the last three decades.

Some of these systems took into consideration only the presence and the amount of well-defined tumoral glands and not the cellular morphology.

Thus, the M.D. Anderson's grading system divided prostatic adenocarcinomas in: grade 1 tumors, with 75–100% gland formation, grade 2 tumors, with 50–75% of tumoral areas consisting of glandular structures (including also tumors containing papillary and/or cribriform patterns regardless the amount), grade 3 tumors, with glands in only 25–50% of tumoral areas and, finally, grade 4 tumors, also named solid, where the presence of gland formation is less than 25% (Brawn *et al.*, 1982).

U.I.C.C. (*Union Internationale de lutte Contre le Cancer*) proposed a very well known and used classification in four grades, G1 – well-differentiated tumors; G2 – moderately differentiated tumors; G3 – poorly differentiated tumors; G4 – anaplastic tumors, using as grading criterion only malignant cell morphology. Other systems used as grading criteria both nuclear and architectural features. For instance, Gaeta designed a grading system with four grades, combining glandular pattern with nuclear morphology (Gaeta *et al.*, 1980).

Broders designed a grading system for urogenital carcinomas, with also four grades using the following criteria: acinar structure and cell morphology, including nuclear and cytoplasmatic features, nucleoli presence, mytotic activity and invasiveness (Broders, 1972).

Finally, another well known system, Mostofi system (1–2–3) has three grades: grade 1 – tumors with well differentiated glands and mild nuclear atypia; grade 2 – tumors with glands lined with cells showing moderate nuclear atypia and grade 3 – tumors with marked nuclear anaplasia (Mostofi, 1975).

The Gleason system is the internationally recognized standard for grading prostate cancer, due mainly to its strong prognostic capability (De la Taille *et al.*, 2003).

This diagnostic grading method is based entirely on the histologic pattern of arrangement of carcinoma cells in Haematoxylin–Eosin-stained sections (Humphrey, 2004).

Dr Donald F. Gleason and members of the Veterans Administration Cooperative Urological Research Group devised it in the 1960's and 1970's. It uses five patterns of tumoral architecture, from well differentiated to poorly differentiate. The grade is scored by the sum of the two most important architectural patterns of the tumor (Gleason, 1977).

Concluding, most of grading systems are biased and cannot be generalized. Anyway, prognostic success of all these systems is based on a specific feature of prostatic carcinoma, compared with other neoplasia, i.e. behaviour of a certain tumor is reflected with an uncommon accuracy in its morphologic feature.

In this study we tried to compare the results of two of these grading systems in the same group of prostatic carcinomas.

MATERIAL AND METHODS

We studied retrospectively 221 patients hospitalised in the Urology Department of Craiova Emergency County Hospital between 1992 and 1999, whose pathologic diagnostic, established in the Pathology Department of the same hospital, was prostate adenocarcinoma.

Using the type of surgical procedure for drawing prostatic tissue as a distribution criterion, the selected cases were divided into three groups as follows:

- **group 1:** transvesical prostatectomy (TVP) for benign nodular hyperplasia – 33 cases.
- **group 2:** transurethral resection (TUR) – 82 cases.
- **group 3:** prostatic needle biopsy (B) – 106 cases.

The materials were represented two different data sources from the Pathology Department archives: histopathologic records and histopathological slides and paraffin blocks of each case.

All tissue specimens were initially processed using the classical histopathologic techniques (fixation and paraffin embedding) and then stained with Haematoxylin-Eosin.

We used, before 2000, in our current practice, an evaluation system for the degree of differentiation resulted from combining the UICC system (G1-G4), Mostofi system (1–2–3) and the older american system proposed by Broders.

The histopathological analysis was performed using the following algorithm:

- **registration** of the histopatologic diagnosis established using the G1-G3 system.
- **reassessment** of the histopathologic diagnosis using the Gleason system.
- **conversion** of the Gleason established diagnosis to G1-G3 system using Epstein's method.
- **comparison** between the initial and obtained results.

For comparing the two series of data: those obtained after conversion and those registered initially we used the following algorithm:

Each series, considered as a classification, had three classes: G1, G2 G3.

We designed a matrix for establishing the errors in grading between the two classifications (Table 1).

We put “0” because for the cases in these cells there was no error of grading between classifications.

We multiplied by “1” the number of cases falsely graded in the initial assessment, i.e. we considered one error for each falsely classified case because the difference in grading was of only one class.

We multiplied by “2” the number of cases falsely graded in the initial assessment, i.e. we considered two errors for each falsely classified case because the difference in grading was of two classes.

We made the sum of errors on each line. Then, we calculated the total number of errors for the studied group, noted as “E”.

We noted the total number of each group with “T”. The E/T ratio, named “Distance between classes” measures with how much mistakes one classification in comparison with the other.

When the two series have three classes, the accepted error value is >15%.

Table 1
Matrix for assessment of “Distance between classes”

Grade		Initial Assessment			No. of Errors
		G1	G2	G3	
Conversion	G1 2	G1	0	...x 1	...x 2
	G1 3		0	...x 1	...x 2
	G1 4		0	...x 1	...x 2
	G1 5	G2	...x 1	0	...x 1
	G1 6		...x 1	0	...x 1
	G1 7		...x 1	0	...x 1
	G1 8	G3	...x 2	...x 1	0
	G1 9		...x 2	...x 1	0
	G1 10		...x 2	...x 1	0
	Total number of errors / Group				
Total number of cases / Group					T
Mean distance between classes					$E \times 100/T$

RESULTS AND DISCUSSIONS

The first step was the registering of histopatologic diagnosis established using the G1-G3 system for each case in all three-study groups (Table 2, Figure 1).

Table 2
Case distribution following G1-G3 system in the whole group and the three subgroups

Grade	Entire Group		I (TVP)		II (TUR)		III (B)	
	Cases	%	Cases	%	Cases	%	Cases	%
G 1	21	9.5	6	18	8	10	7	7
G 2	87	39.5	14	42	32	39	41	39
G 3	113	51	13	40	42	51	58	54
Cases	221	100	33	100	82	100	106	100

The well-differentiated forms of adenocarcinoma were present in a small percent in groups II (TUR) and III (B). However, on transvesical prostatectomy samples, they accounted for almost one fifth of the cases (Figure 2).

The percentage of moderately differentiated forms was almost constant, around 40% of the cases in each study group (Figure 3).

The poorly differentiated forms of adenocarcinoma accounted for half of the cases in groups II and III, while at group I they were only slightly exceeded by the moderately differentiated forms (Figures 4 and 5).

The second step was the reassessment of histopathologic diagnosis for each one of the 221 cases, using the Gleason system. Thus, we established the dominant and secondary patterns for each case and calculated the Gleason score (Table 3, Figure 6). The distribution following the Gleason assessment showed some dispersion of different score types but with an overall and subgroup prevalence of scores higher than “5”, the most frequent being scores “6” and “8”.

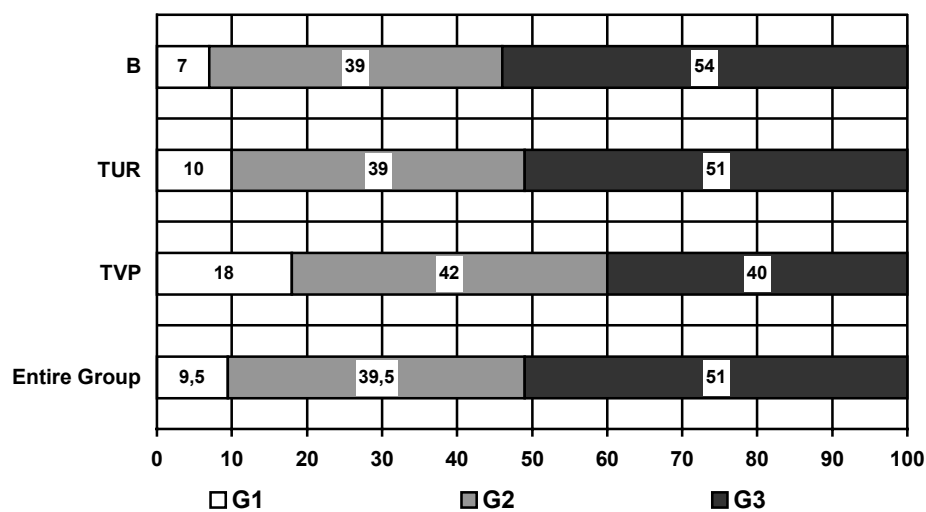


Figure 1 – Histogram of case distribution following G1-G3 (%)

Table 3

Case distribution following Gleason system in the whole group and the three subgroups

Gleason Score	Entire Group		I (TVP)		II (TUR)		III (B)	
	Cases	%	Cases	%	Cases	%	Cases	%
Gleason 2	1	0.5	0	0	1	1.2	0	0
Gleason 3	4	2	3	9	1	1.2	0	0
Gleason 4	12	5.5	1	3	6	7.3	5	4.7
Gleason 5	11	5	3	9	6	7.3	2	2
Gleason 6	55	25	8	24.3	16	19.5	31	29.3
Gleason 7	26	12	5	15.2	11	13.4	10	9.4
Gleason 8	51	23	6	18.2	20	24.4	25	23.6
Gleason 9	32	14	7	21.3	9	11	16	15
Gleason 10	29	13	0	0	12	14.7	17	16
Cases	221	100	33	100	82	100	106	100

Then, we converted Gleason system to G1-G3 system, using Epstein's method, i.e. (Figures 7, 8, 9, and 10):

Gleason scores from 2 to 4 Grade 1

Gleason scores from 5 to 7 Grade 2

Gleason scores from 8 to 10 Grade 3

After conversion, we obtained the results listed in Table 4 and Figure 11.

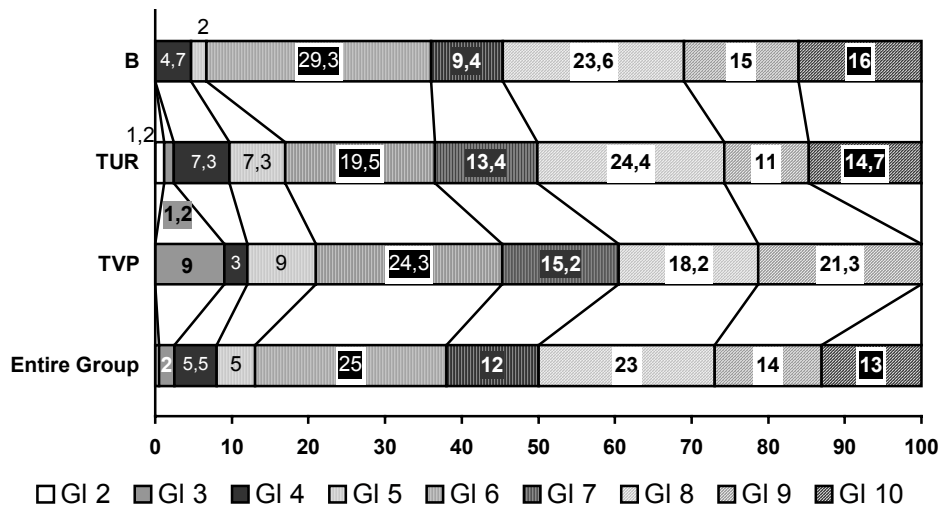


Figure 6 – Gleason scores rates in the three studied groups

Table 4

Case distribution following conversion from Gleason to G1-G3 system in the whole group and the three subgroups

Grade	Entire Group		I (TVP)		II (TUR)		III (B)	
	Cases	%	Cases	%	Cases	%	Cases	%
G 1	17	8	4	12	8	10	5	4.7
G 2	92	42	16	49	33	40	43	40.7
G 3	112	50	13	39	41	50	58	54.6
Cases	221	100	33	100	82	100	106	100

Conversion into G1-G3 system highlights, as the initial evaluation, the prevalence of well and moderately differentiated carcinomas in TVP cases, equal percentages of poorly differentiated and moderately differentiated carcinomas in TUR cases and the prevalence of poorly differentiated carcinomas in B cases.

In the latter group, there were very few well-differentiated carcinomas that, together with the moderately differentiated ones, accounted for less than half of the cases.

Finally we compared the initial diagnosis and those obtained after converting the Gleason system for each subgroup and for the entire study group.

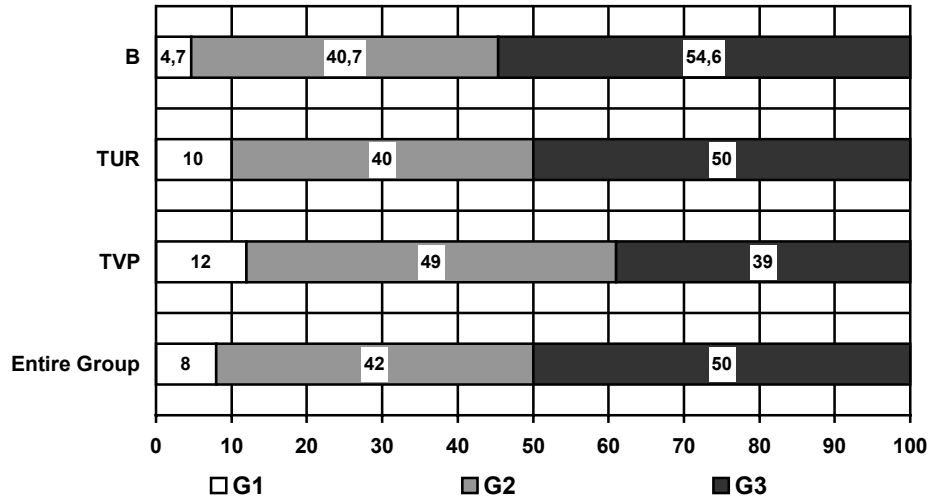


Figure 11 – Histogram of case distribution G1-G3 system after conversion (%)

Group 1 (TVP)

In this group, of the 6 cases initially diagnosed as G1, 2 have been undergraded, probably because of the misinterpretation of Gleason 3a (in one case) as Gleason 2 pattern and of high amount of Gleason 2 pattern as dominant pattern in the other case (Table 5).

Table 5
Comparison between the two grading systems in TVP cases

Gleason System		G1-G3 System				
Score	Cases	Conversion		Initial Assessment		
		Grade	Cases	G1	G2	G3
G1 2	0	G1	4	0	0	0
G1 3	3			3	0	0
G1 4	1			1	0	0
G1 5	3	G2	16	2	1	0
G1 6	8			0	8	0
G1 7	5			0	3	2
G1 8	6	G3	13	0	2	4
G1 9	7			0	0	7
G1 10	0			0	0	0
Cases		33		6	14	13

Other 2 undergraded cases were initially diagnosed as G2, and after conversion, they had G3. One of them had as dominant pattern Gleason 3c and small amounts of Gleason 5b as secondary pattern.

In the other, probably, the uniform pattern 4a was misinterpreted as pattern 3c. Finally, 2 of the 13 cases initially diagnosed as G3, were overgraded, due, in both cases, probably to the large amounts of pattern 4a as dominant pattern (Table 5).

The mean distance between classes, calculated for this group was 18.2% (Table 6).

Table 6
Error Value for TVP group

Grade		Initial Assessment			No. of Errors
		G1	G2	G3	
Conversion	G1	0	0	0	0
		0	0	0	0
		0	0	0	0
	G2	2 × 1	0	0	2
		0	0	0	0
		0	0	2 × 1	2
	G3	0	2 × 1	0	2
		0	0	0	0
		0	0	0	0
	Total number of errors (E)				
Total number of cases (T)					33
Mean distance between classes (E x 100/T)					18.2%

Even the value is higher than threshold value (15%) the errors could be explained either by the close resemblance between neighbouring patterns in Gleason scale or by the high amount of the dominant pattern. Both situations could result in undergrading or overgrading with a class using the G1-G3 system.

Group 2 (TUR)

In this group, of the 8 cases initially diagnosed as G1, 2 have been undergraded for the same reason as in group 1: the high amount of Gleason 2 pattern as dominant pattern.

From the 32 cases initially graded as G2, 2 were overgraded. Even both of them had a uniform Gleason pattern 2, it is likely that lot of tumoral areas were graded as moderately differentiated because of their close resemblance with Gleason pattern 3 a.

Other 4 cases were undergraded. The explanation could be the small amounts of tumoral tissue on the examined specimens, revealing areas with small glands penetrating between stromal smooth muscle fibers which were probably interpreted as moderate differentiation (Table 7).

From the 42 cases initially graded as G3, 6 were overgraded. In one case, the only Gleason pattern 3b with cordonal small glands between stromal components

was interpreted as an aggressive G3 aspect. In the other 5, the Gleason 4a or 4b component of the small tumoral areas resulted in initial G3 overgrading (Table 7).

Table 7
Comparison between the two grading systems in TUR cases

Gleason System		G1-G3 System				
Score	Cases	Conversion		Initial Assessment		
		Grade	Cases	G1	G2	G3
G1 2	1	G1	8	1	0	0
G1 3	1			1	0	0
G1 4	6			4	2	0
G1 5	6	G2	33	2	4	0
G1 6	16			0	15	1
G1 7	11			0	6	5
G1 8	20	G3	41	0	3	17
G1 9	9			0	0	9
G1 10	12			0	1	11
Cases		82		8	32	42

Table 8
Error Value for TUR group

Grade		Initial Assessment			No. of Errors
		G1	G2	G3	
Conversion	G1	0	0	0	0
		0	0	0	0
		0	2 × 1	0	2
	G2	2 × 1	0	0	2
		0	0	1 × 1	1
		0	0	5 × 1	5
	G3	0	3 × 1	0	3
		0	0	0	0
		0	1 × 1	0	1
	Total number of errors (E)				14
Total number of cases (T)				82	
Mean distance between classes (E x 100/T)				17%	

The mean distance between classes, calculated for this group was 17% (Table 8). This value could be explained too either by the close resemblance between neighbouring patterns in Gleason scale or by the high amount of the dominant pattern but also by the small fragments of prostatic tissue and, often, the small amounts of tumoral tissue on the specimens.

Group 3 (B)

Three of the 7 initially G1 graded cases of this group were undergraded. In one case, the mistake was due to the misinterpretation of small amounts of 3a

secondary pattern as pattern 2 and these, an undergrading of one class. In the other 2 cases, the reduced amount of a single pattern 4, with closely packed small glands, between stromal components, was interpreted initially as well-differentiated areas, resulting in an undergrading with two classes (Table 9).

In the group of initially graded G2 cases, we found an overgraded case from the same reason as previously, in TUR group: lot of tumoral areas with Gleason pattern 2, were graded as moderately differentiated because of their close resemblance with Gleason pattern 3a. We also found 2 undergraded cases, from the same reasons as the two cases of the G1 series of this group: the initial interpretation as well-differentiated areas of the single pattern 4, with closely packed small glands between stromal components, but, this time resulting in only one class undergrading (Table 9).

Of the 58 cases initially graded as G3, 4 were overgraded because of the presence of 3c cribriform areas in three cases and the 4a pattern as the dominant one in one case (Table 9).

Table 9
Comparison between the two grading systems in B cases

Gleason System		G1-G3 System				
Score	Cases	Conversion		Initial Assessment		
		Grade	Cases	G1	G2	G3
G1 2	0	G1	5	0	0	0
G1 3	0			0	0	0
G1 4	5			4	1	0
G1 5	2	G2	43	1	1	0
G1 6	31			0	29	2
G1 7	10			0	8	2
G1 8	25	G3	58	1	1	23
G1 9	16			1	1	14
G1 10	17			0	0	17
Cases		106		7	41	58

The mean distance between classes, calculated for this group was 11,3%, a normal value, even it contained two cases undergraded with two classes (Table 10) and in spite of small fragments of prostatic tissue and, often, the small amounts of tumoral tissue on the specimens.

Entire Group

In Table 11, are put together all the cases from the three groups. We found 31 cases with errors of diagnostic at the initial G1-G3 assessment in the whole group after assessment using the more descriptive system designed by Gleason and reconversion to G1-G3 system using Epstein's algorithm. That meant 14% of the all diagnosis. 16 of these errors (51.6% of all errors and 7.2% of all cases) consisted in undergrading of tumoral proliferation. This had mainly by two causes:

the first is the misinterpretation of neighboring architectural patterns as pattern 2/pattern 3a which can result in Gleason borderline score 5 instead of 4 or pattern 3c/pattern 4a which can result in Gleason borderline scores 8 instead of 7. The second is the insufficient amount of prostatic and also tumoral tissue, offered often by TUR and B procedures. The other 15 errors (48.4% of all errors and 6.8% of all cases) were overgradings, due mainly to the same causes mentioned above.

Table 10
Error Value for B group

Grade		Initial Assessment			No. of Errors
		G1	G2	G3	
Conversion	G1	0	0	0	0
		0	0	0	0
		0	1 × 1	0	1
	G2	1 × 1	0	0	1
		0	0	2 × 1	2
		0	0	2 × 1	2
	G3	1 × 2	1 × 1	0	3
		1 × 2	1 × 1	0	3
		0	0	0	0
	Total number of errors (E)				
Total number of cases (T)					106
Mean distance between classes (E x 100/T)					11.3%

Table 11
Comparison between the two grading systems in the entire group

Gleason System		G1-G3 System				
Score	Cases	Conversion		Initial Assessment		
		Grade	Cases	G1	G2	G3
G1 2	1	G1	17	1	0	0
G1 3	4			4	0	0
G1 4	12			9	3	0
G1 5	11	G2	92	5	6	0
G1 6	55			0	52	3
G1 7	26			0	17	9
G1 8	51	G3	112	1	7	43
G1 9	32			1	1	30
G1 10	29			0	1	28
Cases		221		21	87	113

We mention again the two cases from third group with a two classes undergrading because they modified a little the error value (Mean distance between classes) by doubling their index. The error percentage of 14.9% was considered a normal one, showing that the diagnosis concordance between the two grading

systems on our study group and subgroups was good. We can draw the conclusion that both systems are reliable if experienced pathologists use them (Table 12).

Table 12
Error value for the entire group

	Grade	Initial Assessment			No. of Errors
		G1	G2	G3	
Conversion	G1	0	0	0	0
		0	0	0	0
		0	3 × 1	0	3
	G2	5 × 1	0	0	5
		0	0	3 × 1	3
		0	0	9 × 1	9
	G3	1 × 2	7 × 1	0	9
		1 × 2	1 × 1	0	3
		0	1 × 1	0	1
	Total number of errors (E)				33
	Total number of cases (T)				221
	Mean distance between classes (E x 100/T)				14.9%

Our data are concordant with those from other studies. Thus, Fukagai *et al.* (2001) insisted on clarifying the histologic criteria for distinguishing each grade and especially between Gleason grades (patterns) 2 and 3 for an accurate grading. Coard and Freeman (2004) highlighted the idea that, in the Gleason scores of prostate needle biopsy specimens, the greatest discordance seems to be in distinguishing Gleason score 6 from 7, and is more frequent among biopsy specimens with lower tumor volumes, particularly among those with less than 30% involvement. And, finally, Shen *et al.* (2003) concluded that the potential for grading errors is greatest with well-differentiated tumors and in patients with a Gleason score of <7 on the needle biopsy. Even Gleason pointed out possible disagreements between a combined Gleason score and the 1–2–3 systems but he thought that their occurrence is rare, namely when there is a marked difference between the dominant and secondary Gleason patterns. For example if a tumor had a combined Gleason score of $2 + 5 = 7$, it would be graded as part of the middle group. In contrast, in the 1–2–3 system, if the tumor were undifferentiated, than it would be graded as part of the third group (Gleason, 1977).

CONCLUSIONS

On the whole, with only 16 undergraded cases and 15 overgraded cases and with an overall error value of 14.9%, the correlation between the two histopathologic grading systems was very good, even if the histopathologic material obtained by TUR and B was dominant, both systems proving to be reliable.

REFERENCES

- BRAWN P.N., AYALA A.G., VON ESCHENBACH A.C. *et al.*, *Histologic grading study of prostate adenocarcinoma: The development of a new system and comparison with other methods – a preliminary study*, *Cancer*, 1982, 49:525.
- BRODERS A.C., *Epithelioma of the genito-urinary organs*, *Ann. Surg.*, 1972, 75:574–604.
- DE LA TAILLE A., VIELLEFOND A., BERGER N. *et al.*, *Evaluation of the interobserver reproducibility of Gleason grading of prostatic adenocarcinoma using tissue microarrays*, *Hum Pathol*, 2003, 34(5):444–449.
- EPSTEIN J.I., STEINBERG G.D., *The significance of low-grade prostate cancer on needle biopsy*, *Cancer*, 1990, 66:1927.
- GAETA J.F., ASIRWATHAM J.E., MILLER G., MURPHY G.P., *Histological grading of primary prostatic cancer: a new approach to an old problem*, *J Urol*, 1980, 123:689.
- GLEASON D.F., *Histologic grading and staging of prostatic carcinoma*. In: TANNENBAUM M. (ed), *Urologic pathology: The prostate*, Lea & Febiger, Philadelphia, 1977, 171–197.
- FUKAGAI T., NAMIKI T., NAMIKI H. *et al.*, *Discrepancies between Gleason scores of needle biopsy and radical prostatectomy specimens*, *Pathol Int*, 2001, 51(5):364–370.
- HUMPHREY P.A., *Gleason grading and prognostic factors in carcinoma of the prostate*, *Mod Pathol*, 2004, 17(3):292–306.
- MOSTOFI F.K., *Grading of prostatic carcinoma*, *Cancer Chemother Rep (Part I)*, 1975, 59:111–117.

Acknowledgements. The study was part of the Grant C.N.C.S.I.S code D107/2000, financed by Romanian Government and World Bank.

E-mail: Bogdan Zaharia, MD, emp@umfcv.ro or emp@euroweb.ro

Received: 14 September, 2004

Accepted: 23 November, 2004

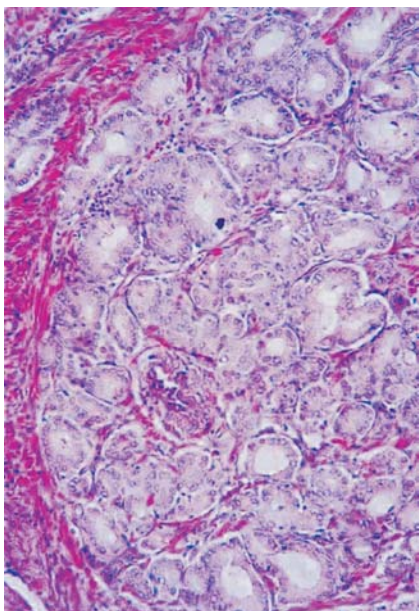


Figure 2 – Well-differentiated adenocarcinoma (G1)

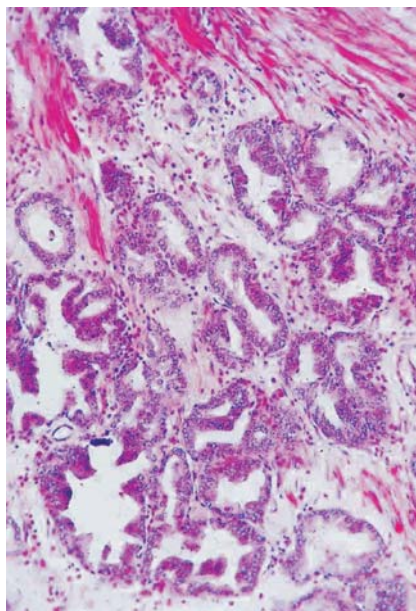


Figure 3 – Moderately differentiated adenocarcinoma (G2)

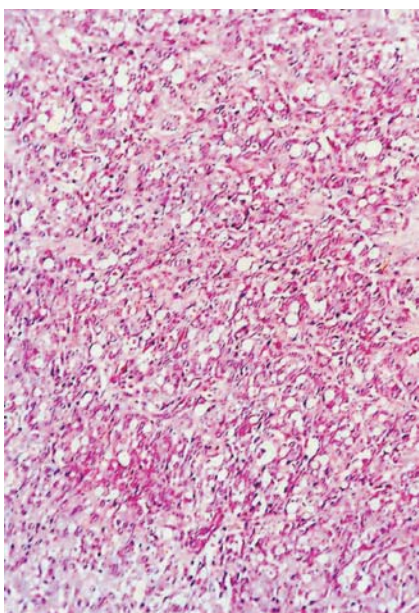


Figure 4 – Poorly differentiated adenocarcinoma (G3), solid type

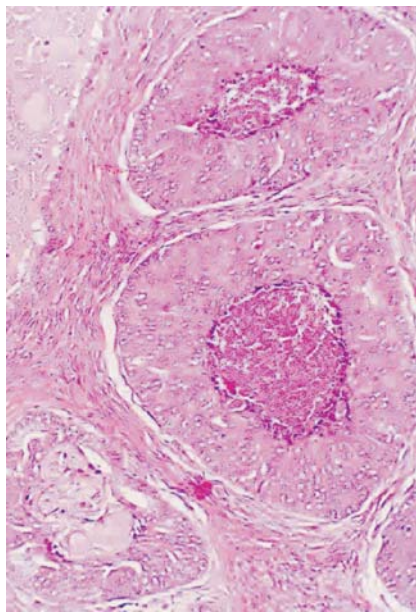


Figure 5 – Poorly differentiated adenocarcinoma (G3), comedo type

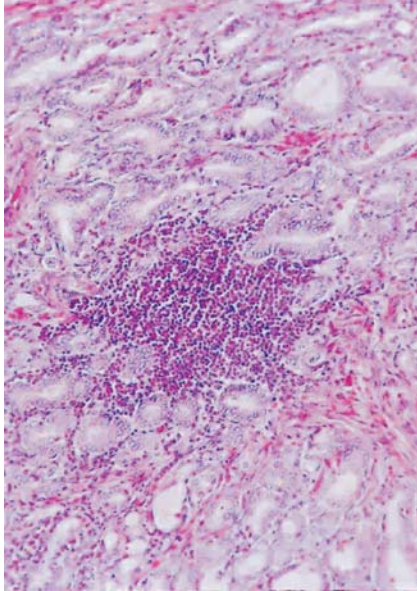


Figure 7 – Gleason 2 pattern, realising
a Gleason 4 score, converted as G1
Initially, OVERGRADED as G2

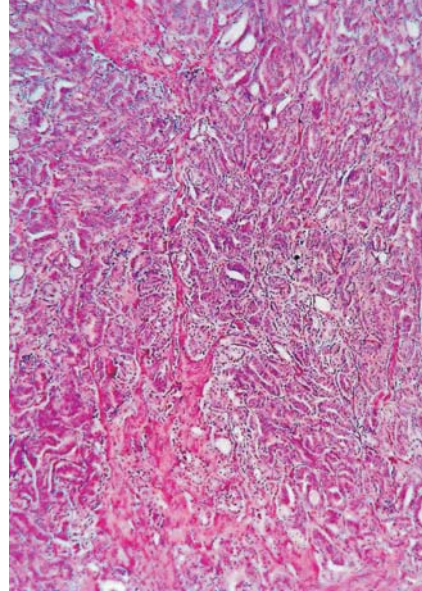


Figure 8 – Gleason 3b pattern, realising
a Gleason 6 score, converted as G2
Initially, OVERGRADED as G3

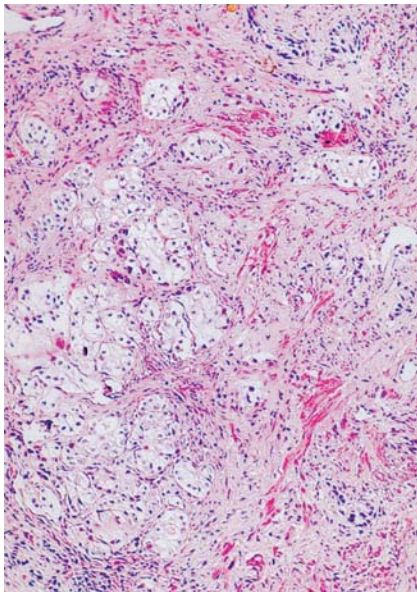


Figure 9 – Gleason 4b pattern, realising
a Gleason 8 score, converted as G3
Initially, UNDERGRADED as G1

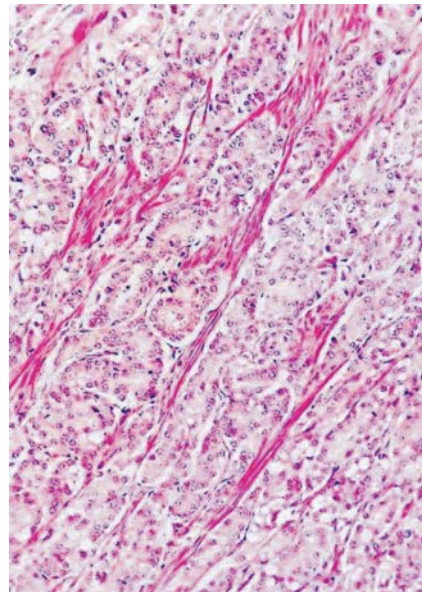


Figure 10 – Gleason 4A pattern, realising
a Gleason 8 score, converted as G3
Initially, UNDERGRADED as G2