International Urology and Nephrology 23 (5), pp. 449–454 (1991)

Histologic Grading of Prostatic Adenocarcinoma: Intraobserver Reproducibility of the Mostofi, Gleason and Böcking Grading Systems

M. L. CINTRA, A. BILLIS

Departamento de Anatomia Patológica, Universidade Estadual de Campinas, Brazil

(Received October 8, 1990)

Intraobserver variation of three grading systems — Mostofi, Gleason and Böcking — is examined. No significant difference was noted between the histological grades found in the two examinations by any of the three methods used. Neither the type of surgical procedure nor the number of slices with tumour influenced the reproducibility of histological grading within each system studied. In the Gleason system the intraobserver highest disagreement would not have resulted in change of therapy choice, but in 2% of tumours graded according to the Mostofi system this would have occurred if the choice of therapy would depend on the grading results.

Introduction

The importance of histologic grading as prognostic factor to different tumours has been the subject of considerable study [12]. Its application is particularly suitable in prostatic adenocarcinoma due its high prevalence and variable clinical behaviour. Cancer of the prostate ranks third as a cause of death in American men [10] and fourth in incidence of malignant tumours in men in Brazil [5]. The natural history of prostatic cancer is characterized by extreme variations in its biologic course [7], and the possibility of predicting clinical evolution is important in therapeutic planning of localized prostatic cancer [3, 13]. Recognizing the relationship between histologic appearance of prostatic tumours and clinical evolution, many investigators have been searching for histologic classifications that fulfil the requirements of prognostic accuracy, reliability and objectivity. A limitation of most grading systems is that histologic grading is a subjective procedure. Systems based upon sophisticated technics are highly reproducible but require expensive equipment and are time-consuming, which restricts their application. The inconsistency in histologic grading may invalidate its use in treatment decision. While one pathologist recommends a conservative treatment, another prefers an aggressive one. In this way, the reproducibility has the same significance as the predictive character of prognosis. It is necessary to find the most reliable method before histologic grades are incorporated routinely into the diagnosis of prostatic carcinoma. The two most frequently used histologic grading systems at the moment are those proposed by Gleason [10] and Mostofi [14].

VSP, Utrecht
Akadémiai Kiadó, Budapest

The Gleason classification has received remarkable acceptation due to its prognostic accuracy and relative simplicity [1, 16, 17]. The Mostofi system has been adopted by the World Health Organization and the International Union Against Cancer [11] but it does not have the clinical significance established by Gleason. Another method, proposed by Böcking et al. [4] using diagnostic parameters according to WHO classification, showed a good correlation with survival probabilities and the presence of metastases. In the present investigation, the intraobserver reproducibility of these three histologic grading systems is examined.

Material and methods

A set of histological sections of 139 prostatic carcinomas stained with haematoxylin and eosin was graded twice by one and the same pathologist on two different occasions, and was classified using the grading systems of tumour differentiation as described by Mostofi, Gleason and Böcking et al. The cases were subdivided according to the type of surgical procedure (transurethral resection, TUR, or open prostatectomy) and to the number of fragments containing tumour. Clinical stage was determined in 129 cases according to Whitmore [20]. The intraobserver variation was assessed using a standard statistical analysis package (stepwise discriminant analysis).

Results

We examined from the material obtained by open prostatectomy a score of 8.37 slices, of which 3.64 had neoplasm, and from that obtained by TUR, a score of 11.65 slices, of which 7.80 had neoplasm. The clinical stages were grouped into three classes: A, B and C/D. The distribution curves of these three classes were normal, justifying the paired observations that have been done afterwards. Tumours obtained by TUR showed higher clinical stages when compared with those obtained from open prostatectomy (Table 1). The intraobserver reproducibility and variation of the grades assigned according to each of the three methods

Table 1

Group	Tumour stage (%)		
	A	B	C/D
Prostatectomy	47.06	35.29	17.65
TUR	10.52	47.37	42.11

Chi square: 21.607; D. F. = 2;
prob. = 2.033; $p < 0.001$

are illustrated in Table 2. No significant difference was noted between the histologic grades found in the two examinations by any of the three methods used (Table 3). Neither the type of surgical procedure nor the number of slices with tumour influenced the reproducibility of histologic grading within each system studied. If the tumours were graded only according to the nuclear features, either into the Mostofi or the Böcking system, there would have been significant intraobserver differences, but such disagreements did not affect reproducibility of the final grades (Table 4). Mostofi, Gleason and Böcking grades correlated strongly

Table 2

Intraobserver reproducibility (%)

System	Total values*	Disagreement within		Architectural	Nuclear	Primary
		one grade unit	two grade units		pattern	
Mostofi	79.85	17.98	2.14	90.64	77.69	—
Gleason	62.58	28.05	9.35	—	—	78.41
Böcking	78.41	20.86	0.71	79.88	65.47	—

* Mostofi final grade; Gleason pattern score (sum of primary and secondary Gleason patterns); Böcking combined grade

Table 3

Difference between means (paired observations/intraobserver reproducibility)

System	Prostatectomy group			TUR group		
	Mean	Std. error	T	Mean	Std. error	T
Mostofi	0.0270	0.0611	0.4423	0.0588	0.0554	1.0613
Gleason	0.0270	0.1313	0.2058	0.0196	0.0812	0.2414
Böcking	0.0811	0.0598	1.3568	0.0392	0.0520	0.7243
	N = 37; D. F. = 36			N = 102; D. F. = 101		

Table 4

Intraobserver reproducibility/nuclear pattern

System	Mean	Std. error	T
Mostofi	0.0863	0.0375	2.3021
Böcking	0.1223	0.0490	2.4959
	N = 139; D. F. = 138		

Table 5

Combined Böcking grades (%)*

Prostatectomy	I	II	III
Prostatectomy	4.31	11.51	10.79
TUR	5.75	20.86	46.76

Chi square: 6.266; D. F. = 2;
prob. = 0.0436; $p < 0.05$
* Values obtained on the first occasion of two observations.

with clinical staging at admission (correlation matrix, with clinical stages grouped in three classes: A, B and C/D as fixed variable). However, only Böcking grades displayed a higher incidence of advanced clinical stages in tumours resected via the transurethral route as compared to those obtained by enucleation (Table 5).

Discussion

Open prostatectomy was indicated for clinical but localized tumours (radical type) and treatment for nodular hyperplasia (tissue resected for obstruction), and in this case cancer was found unexpectedly. That is why this group displayed smaller clinical stages when compared with tumours obtained from TUR.

Employing his own system, Gleason estimated his intraobserver reproducibility rate to be 80% [15] which is better than the reproducibility found by us (62%) and by other workers (65% and 42%) [18]. We had agreement within one score unit of 90% which is similar to theirs (87%) [18]. Considering that Gleason grade sum (pattern score) ranges from 2 to 10, as compared to the other two methods that assign only three final grades to prostatic adenocarcinoma, we think that this is an excellent level of intraobserver reproducibility. Furthermore, in the Gleason system the intraobserver highest disagreement would not have resulted in change of therapy choice if it would depend on the grading results. For the Mostofi system exact agreements were seen in 90%, but in 2% of the cases change of treatment decision would have occurred (disagreement of two grade units). There was exact agreement for Mostofi and Böcking nuclear anaplasia in 77% and 65%, respectively, which is worse than the intraobserver reproducibility found by Böcking (73.5%), and there was significant intraobserver difference.

Structural details are more reproducible than cytological [19], so that nuclear features, while useful as diagnostic and prognostic factors, may not be used alone. Gaeta et al. [9] observed that glandular grade was more related to biologic behaviour than nuclear grade. Considering that nuclear grade sums an element always present in subjectivity, we disagree with the authors who recommend that nuclear characteristics be included to further the discriminative capabilities of

the Gleason system [14, 17]. Within the Böcking system, nuclear grade disagreement did not modify the level of the combined final grade reproduciblity as compared with the architectural pattern alone. It is possible that the same may occur within the Gleason system if nuclear characteristics are to be considered in prospective studies. Böcking found his own reproducibility rate with his system to be 87.5% [4] which is better than ours (78.4%) and that of others [18]. They emphasized as an influencing factor in system reproducibility the sampling of pathological material. According to Böcking, all systems, like Gleason's and Mostofi's that consider the quantitatively predominant growth patterns, attach too much weight to sample size and introduce another subjective criterion: the quantification of different architectural/cytological pattern areas. Our work is based on material selected retrospectively and so it was not possible to keep a fixed number of histological slices. Furthermore, the TUR material displayed smaller and more numerous slices as compared with enucleated prostates. Meanwhile we did not observe histological sampling size or surgical procedure type interference in the reproducibility of any system studied.

The success of surgical treatment for cancer is dependent upon complete removal of the tumour while it is still confined to the primary site. However, biological nature and not only tumour volume will define the result of precocious treatment. Tumour grade is to some extent predictive of tumour stage [2, 6, 8]. In this work all the three methods studied were predictive of clinical stage on admission. This relationship emphasizes the accuracy of histologic grades in predicting mortality rates since poorly differentiated tumours are generally more advanced. The only system that considered the clinical stage differences between the two groups (TUR and prostatectomy) was the Böcking system.

As a result of this study, the following regimen is proposed: histologic grading system must be incorporated routinely into the diagnosis of prostatic adenocarcinoma.

A unique method must be used to further self-confidence by the pathologist and facilitate the memorization and automation of the residents of pathology. We suggest that one of the two methods studied (Gleason's or Böcking's) should be chosen with close co-operation between hospital urologists and pathologists.

Acknowledgements

The authors are indebted to Prof. Luis Alberto Magna and Prof. Walter Pinto Jr. (Departamento de Genética, UNICAMP) for statistical assistance.

References

1. Bain, G. O., Koch, M., Hanson, J.: Feasibility of grading prostatic carcinomas. *Arch. Pathol. Lab. Med.*, *106*, 265 (1982).
2. Billis, A.: Graduação histológica do carcinoma da próstata correlação com estadiamento clínico. *Br. J. Urol.*, *9*, 88 (1983).

3. Billis, A., Ferreira, U., Castilho, L. N., Baccili, C.: Carcinoma incidental (estádio A) da próstata em material de patologia cirúrgica. Importância do grau histológico (sistemas Mostofi, Böcking e Gleason) e da extensão tumoral. *Br. J. Urol., 14,* 84 (1988).
4. Böcking, A., Kiehn, J., Heinzel-Wach, M.: Combined histologic grading of prostatic carcinoma. *Cancer, 50,* 288 (1982).
5. Brumini, R.: Câncer no Brasil: dados histopatológicos. Rio de Janeiro, Ministério da Saúde, Campanha Nacional de Combate ao Câncer, 1982.
6. Catalona, W. J.: Grading and staging of prostate cancer. *J. Urol., 128,* 747 (1982).
7. Denis, L. J.: Clinical cancer of prostate. Influence of tumor stage and grade of five year survival. *Acta Urol. Belg., 40,* 126 (1972).
8. Faul, P., Eisenberger, F., Elsässer, E.: Metastatic involvement of pelvic lymph node in relation to the morphological differentiation grade and clinical status of prostatic cancer. *Urologe, 24,* 326 (1985).
9. Gaeta, J. F., Englander, L. C., Murphy, G. P.: Comparative evaluation of national prostatic cancer treatment group and Gleason systems for pathologic grading of primary prostatic cancer. *Urology, 27,* 306 (1986).
10. Gleason, D. F.: Histologic grading and clinical staging of prostatic carcinoma. In: M. Tannenbaum (ed.): Urologic Pathology: The Prostate. Lea and Febiger, Philadelphia 1977.
11. Harada, M., Mostofi, F. K., Corle, D. K., Byar, D. P., Trump, B. F.: Preliminary studies of histological prognosis in cancer of the prostate. *Cancer Treat. Rep., 61,* 223 (1977).
12. Henson, D. E.: The histological grading of neoplasm. *Arch. Pathol. Lab. Med., 112,* 1091 (1988).
13. Huben, R. P., Murphy, G. P.: Prostate cancer, an update. CA, *36,* 274 (1986).
14. Mostofi, F. K.: Grading of prostatic carcinoma. *Cancer Chemother. Res., 59,* 111 (1975).
15. Murphy, G. P., Whitmore, W. F. Jr.: A report of the workshops on the current status of the histologic grading of prostate cancer. *Cancer, 44,* 1490 (1979).
16. Paulson, D. F., Piserchia, P. V., Gardner, W.: Predictors of lymphatic spread in prostatic adenocarcinoma: uro-oncology research group study. *J. Urol., 123,* 697 (1980).
17. Sogani, P. C., Israel, A., Lieberman, P. H., Lesser, M. L., Whitmore, W. F. Jr.: Gleason grading of prostate cancer: a predictor of survival. *Urology, 25,* 223 (1985).
18. Svanholm, H., Mygind, H.: Prostatic carcinoma reproducibility of histologic grading. *Acta Pathol. Microbiol. Immunol. Scand., 93,* 67 (1985).
19. Voeth, C., Droese, M., Steuer, G.: Results from cytologic grading of prostatic carcinoma. *Urologe, 17,* 367 (1978).
20. Whitmore, W. F. Jr.: The natural history of prostatic cancer. *Cancer, 32,* 1104 (1973).