World J Urol (1986) 4: 147-152

World Journal of Urology

Problems in grading of prostatic carcinoma: interobserver reproducibility of five different grading systems

Fiebo J. W. ten Kate[1]*, Maarten P. W. Gallee[1], Paul I. M. Schmitz[2], Adriaan C. Joebsis[4], Roy O. van der Heul[1], M. Eric F. Prins[1] and Jan H. M. Blom[3]

Departments of [1]Pathology, [2]Biostatistics and [3]Urology, Erasmus University, NL-3000 DR Rotterdam and Department of [4]Pathology, Bergweg Hospital, NL-3000 DR Rotterdam, The Netherlands

Summary. In order to investigate the reproducibility of grading systems for prostatic carcinoma currently in use, a comparative histological grading study was done. These studies were carried out on tissue sections from radical prostatectomy specimens ($N = 50$) stained with hematoxylin and eosin. Five pathologists with varying professional experience participated in the study, using five different grading systems: those of Broders, Brawn, Gleason (for statistical compilation the modified version), Mostofi, and a modified Mostofi grading method recently described by Schroeder and Mostofi. Weighted kappa coefficients ranged from 0.21 to 0.52. None of the systems investigated demonstrated a high degree of reproducibility ($k > 0.70$). Reproducibility of the systems described by Broders and Brawn was reasonably good ($k = 0.52$ and 0.41, respectively). With the modified Gleason method (rearrangement of Gleason scores into 3 grades), a considerable difference was noted between the numerical agreement score (among at least 3 observers) and the measured kappa value (100% and 0.30, respectively). The methods described by Mostofi and Schroeder-Mostofi revealed only limited reproducibility ($k = 0.21$ and 0.34, respectively).

Introduction

Prostatic carcinoma is a major cause of death among the male population in most developed countries [8, 36]. Its prevalence and incidence is even higher, making this malignancy the most common neoplasm of the male urogenital tract [36]. The discrepancy between mortality and morbidity rates reflects the wide scope of its clinical behavior, ranging from neoplasms found only incidentally at post mortem examination (nearly 10%) to those that metastasize early and eventually lead to death [8, 36].

In order to predict the clinical behavior and aggressiveness of prostatic carcinoma so that the most appropriate therapeutic regimen can be chosen and the prognosis determined, several grading systems have been developed by pathologists [10, 11, 13, 15, 19, 21]. Nevertheless, much controversy still exists as to the most reliable grading method [1, 12, 17, 28].

Böcking et al. [5] have described the main objectives that a grading system should fulfill:

1. Each microscopic diagnostic criterion has to correlate with biological behavior and prognosis.
2. It must display sufficient reproducibility.
3. Grading done on random biopsies should, whenever possible, be representative of the tumor as a whole.

In some systems, grading is performed upon the least differentiated areas; in others, a selection of the most predominant, growth patterns is made. Thus far, approximately 30 grading systems have been described. Some use low-power microscopy, taking into consideration only histological growth characteristics and their relationships to the surrounding stroma [5, 20, 39]. However, in most, prognostic significance has been attributed to cytological features [9, 27, 30, 38]. Combinations of both histological and cytological features form the basis of yet other systems. Also, in some studies, cytological features are examined by use of morphometry [9, 12].

To investigate the degree of reproducibility, we have examined the interobserver variation of five grading systems in current use. Two methods, described by Broders [7] and Brawn (M. D. Anderson Hospital) [6], were chosen as their prognostic value is predominantly based upon histological criteria. The method described by Gleason [16, 18, 20, 22, 25, 37] was included for its general reputation and the attention given to both growth characteristics and interaction of tumor with the surrounding stroma. To incorporate a system that utilizes cytological features as a prognostic indicator, the method of Mostofi was selected [27, 28]. Finally, a recently introduced method described by Schroeder and Mostofi was included [33, 34, 35]. The latter system was based upon a retrospective multivariate analysis of a large number of histological and cytological criteria.

Materials and methods

Patients and materials

Out of 464 patients of the original series of Belt and Schroeder (1930–1970), who had had total perineal prostatectomy for limited prostatic

* To whom correspondence should be addressed

Table 1. Major grading criteria of the five systems evaluated in this study

Broders (1926) [7]

Grade 1 100%–75% glandular differentiation; 0%–25% undifferentiated
Grade 2 75%–50% glandular differentiation; 25%–50% undifferentiated
Grade 3 50%–25% glandular differentiation; 50%–75% undifferentiated
Grade 4 0%–25% glandular differentiation; 100%–75% undifferentiated

Gleason (1966) [18, 19, 20]

Pattern 1 well-differentiated, small, closely packed, uniform glands in essentially circumscribed masses
Pattern 2 similar to pattern 1 with moderate variation in shape and size of glands and more atypia in the tumor cells; more loosely arranged though still circumscribed
Pattern 3 similar to pattern 2 with marked irregularity in size and shape of glands with small glands or individual cells invading the stroma
Pattern 4 raggedly infiltrating, fused glandular tumor, frequently with pale cells, may resemble hypernephroma of kidney
Pattern 5 anaplastic carcinoma with minimal glandular differentiation, diffusely infiltrating prostatic stroma

Mostofi (1975) [27]

Grade 1 glandular differentiation with slight nuclear anaplasia
Grade 2 glandular differentiation with moderate nuclear anaplasia
Grade 3 glandular differentiation with marked nuclear anaplasia or undifferentiated tumor

Brawn (M. D. Anderson, 1982) [6]

Grade 1 75%–100% glandular differentiation; 0%–25% of the tumor does not form glands. Excluded are cribriform-papillary tumors
Grade 2 50%–75% glandular differentiation; 25%–50% of the tumor does not form glands. Included are tumors consisting of 50% or more of a cribriform-papillary pattern
Grade 3 25%–50% glandular differentiation; 50%–75% of the tumor does not form glands
Grade 4 0%–25% glandular differentiation; the remainder is undifferentiated

Schroeder/Mostofi (1985) [33, 34, 35]

Class 1 glandular differentiation, absence of mitoses, slight nuclear anaplasia
Class 2 glandular differentiation, absence of mitoses, moderate nuclear anaplasia
or: glandular differentiation, mitotic activity, slight nuclear anaplasia
or: undifferentiated tumor or cribriform growth variant, absence of mitosis and slight nuclear anaplasia
Class 3 glandular differentiation and mitotic activity combined with moderate or marked nuclear anaplasia
or: glandular differentiation, absence of mitoses, marked nuclear anaplasia
or: undifferentiated or cribriform growth variant, absence of mitoses, moderate anaplasia
or: undifferentiated or cribriform growth variant, presence of mitoses, slight nuclear anaplasia
Class 4 undifferentiated tumor or cribriform growth variant, absence of mitoses, marked nuclear anaplasia
or: undifferentiated or cribriform growth variant, presence of mitoses, moderate nuclear anaplasia
Class 5 undifferentiated tumor or cribriform growth variant, presence of mitoses, marked anaplasia

cancer, 50 patients were randomly selected [4, 32]. Average age at the time of surgery was 66 years, ranging from 48 to 80 years. The mean period of follow-up lasted for 125 months, ranging from 8 to 317 months. The survival curves and curves of time interval until first recurrence for the 50 patients included in this grading study did not differ significantly from that of a much larger random selection in the series of Belt and Schroeder. None of the 50 patients of this study were lost to follow-up. In some of these cases the histology was not optimal. These 50 patients are a random selection of those previously graded by Mostofi and Schroeder [33, 34, 35], one of the few series to date with a follow-up of more than 10 years. The data recorded by Mostofi were also utilized for statistical analysis.

Prostatectomy specimens were fixed in a buffered 4% formaldehyde solution, and sectioned in a stepwise fashion. All sections were routinely stained with hematoxylin and eosin and an average of 4 slides per case were available for histological examination. The number of slides per prostatectomy specimen varied from 2 to 11. The number of slides in which tumor was present varied from 1 to 9 (mean: 3).

Grading systems

The 50 cases were evaluated by five pathologists using five grading systems in common use. The major criteria as well as the scoring systems are summarized briefly in Table 1 [6, 7, 18, 25, 27, 33, 34, 35].

The final Gleason score was assessed by the total of the scores of the two quantitatively predominating growth patterns. In those tumors in which only one growth pattern was recognized, the value of the growth pattern selected was doubled. For statistical compilation, as proposed in the literature [2, 14], Gleason scores were rearranged in 3 main grades (or groups) in the following manner: specimens assigned as Gleason score 2, 3, and 4 tumors were included in grade 1; tumors with Gleason scores 5, 6, and 7 in grade 2, and those with Gleason scores 8, 9, and 10 in grade 3. This is referred to as the modified Gleason grading method.

Statistical analysis. Data was analyzed by calculating a measure of agreement among the five pathologists for each grading system. For most grading methods included in this study, "standard" scores

Table 2. Distribution of grades (1–5) and Gleason scores (1–10) tabulated separately for each grading system as assessed by the participating pathologists (A–E) and by Mostofi himself[a]

Grading systems	Grade										Total no. of scores
	1	2	3	4	5	6	7	8	9	10	
Broders	144 (58%)	63 (25%)	31 (12%)	12 (5%)	–	–	–	–	–	–	250
M. D. Anderson	102 (41%)	108 (43%)	26 (10%)	14 (6%)	–	–	–	–	–	–	250
Mostofi	22 (9%)	120 (48%)	108 (43%)	–	–	–	–	–	–	–	250
Mostofi[a]	1 (2%)	32 (64%)	17 (34%)	–	–	–	–	–	–	–	50
Gleason	–	0 (0%)	4 (2%)	24 (10%)	45 (18%)	117 (47%)	27 (11%)	21 (8%)	6 (2%)	6 (2%)	250
Modified Gleason	28 (11%)	189 (76%)	33 (13%)	–	–	–	–	–	–	–	250
Mostofi-Schroeder	5 (2%)	34 (14%)	91 (36%)	84 (34%)	36 (14%)	–	–	–	–	–	250
Mostofi-Schroeder[a]	1 (2%)	3 (6%)	21 (42%)	17 (34%)	8 (16%)	–	–	–	–	–	50

Table 3. Absolute and percentual measure of agreement and kappa values for each grading method tabulated separately

	Broders 4-grade system	*Anderson* 4-grade system	*Mostofi* 3-grade system	Mod. *Gleason* 3-grade system	*Mostofi-Schroeder* 5-grade system
Agreement among:					
5 observers	13 (26%)	12 (24%)	6 (12%)	18 (36%)	1 (2%)
4 observers	14 (28%)	16 (32%)	19 (38%)	15 (30%)	8 (16%)
3 observers	18 (36%)	15 (30%)	20 (40%)	17 (34%)	29 (58%)
At least 3 observers	45 (90%)	43 (86%)	45 (90%)	50 (100%)	38 (76%)
A difference of more than 2 grades	14 (28%)	16 (32%)	12 (24%)	2 (4%)	29 (58%)
Total no. of specimens	50 (100%)	50 (100%)	50 (100%)	50 (100%)	50 (100%)
Kappa values	0.52	0.41	0.21	0.30	0.34

assigned by referee pathologists were not available. Therefore, weighted kappa coefficients were calculated as the best parameter of interobserver agreement. In order to incorporate the ordinal scales of the grading systems and to make the agreement of assessments for each system comparable, the so-called weighted kappa (k) coefficients were used together with quadratic disagreement weights [31].

Results

Grading results obtained for all five methods are summarized in Tables 2 and 3. Table 2 presents the number of scores assigned by all pathologists to each grade for the five different methods evaluated in this study. A total of 250 scores per method were available (5 pathologists and 50 specimens). For the Mostofi and Mostofi-Schroeder methods, referee scores recorded by Mostofi himself were available and are included in this table.

Table 3 presents data related to the measure of agreement and disagreement for each grading method.

Broders grading method

Tumors in which glandular formation predominates (i.e. at least 50% consists of glandular formation) relatively outnumber tumors in which a minor part is differentiated. Of the 250 gradings performed, 207 (83%) were recorded as grade 1 or 2 (Table 2). Most agreement was also observed in the lower grades, i.e., 62% and 18% agreement for grades 1 and 2, respectively, by at least 3 observers (data not shown). The weighted kappa coefficient measured for the Broders method was 0.52 (Table 3). In 14 cases, the recorded scores differed by more than one grade. In 5 cases, only 2 pathologists agreed with each other (10%). The latter demonstrates almost total disagreement among the participating pathologists.

M. D. Anderson grading method

Of the total of 250 scores, 210 were conferred upon tumors predominantly characterized by a glandular or cribriform-papillary growth pattern (84% (Table 2). The kappa coefficient calculated for the Anderson system was 0.41 (Table 3). Most agreement was found for grade 1 and 2. Agreement existed between at least 3 pathologists in 78% related to grade 1 and 2 tumors. In 16 cases (32%), more than one grade difference was observed. Concordance between only 2 pathologists was noticed in 7 cases (14%).

Mostofi grading method

Glandular differentiation combined with slight nuclear anaplasia was only recorded in 22 scores out of a total of 250 (9%) (Table 2). The scores assessed earlier by Mostofi for the same material are also summarized in Table 2. The overall weighted kappa score was calculated as 0.21 (Table 3). Complete agreement among all pathologists was never found for grade 1 tumors (data not shown). More than one grade difference per specimen was observed in 12 cases (24%). Only in 5 cases was a total lack of concordance seen.

To obtain insight into the measure of agreement (or disagreement) of each participating pathologist with the grading results obtained by Mostofi himself, an individual kappa score was assessed. None of the 5 pathologists attained a weighted kappa coefficient of agreement above 0.40.

Gleason grading method

Grading results determined according to the traditional and modified Gleason methods are shown in Table 2. None of the tumors received a traditional Gleason score of 2. In Table 3, the measure of agreement is presented only for the modified results of the rearranged Gleason scores [2, 14). The weighted kappa score was calculated as 0.30. In all 50 cases (100%), at least 3 investigators agreed with each other concerning the chosen modified Gleason grade. As group/grade 2 was most frequently chosen, maximal agreement was found in this category. More than one grade disagreement of the modified system was found in only 2 cases.

Mostofi-Schroeder grading method

In Tables 2, 3 and 4, results using the Mostofi-Schroeder method are shown. Besides the distribution of scores obtained by the 5 participating pathologists, Table 2 also includes the results obtained by Mostofi. The weighted kappa coefficient as a measure of agreement among pathologists was calculated as 0.34. An overall percentage for all prognostic classes of 76% was a measure of agreement of at least 3 pathologists. In only 5 cases was grade 1 recorded. Regarding grade 5 tumors, agreement among 3 pathologists was only achieved in 2 cases. Disagreement reflected in the number of cases in which pathologists differed by more than one grade was recorded in 29 cases. Using the data recorded by Mostofi himself as a "standard", the individual weighted kappa coefficient as a measure of agreement for each pathologists was assessed (Table 4). With the Mostofi-Schroeder method, unlike the Mostofi method, 2 of the 5 pathologists scored a kappa value above 0.40.

Discussion

Kappa values for each method separately indicate that none of the grading systems had a high degree of concordance ($k>0.70$). Kappa values reflecting fair to reasonable reproducibility (kappa ranging from 0.40 to 0.70) were only found by the Broders and Anderson grading methods (Table 3). However, in contrast to these values, in both the Broders system and the Anderson grading method (in 14 and 16 cases, respectively), disagreement among observers of more than two grading steps was noted (Table 3). Such lack of agreement may be partly attributed to impaired tissue preservation and fixation. Furthermore, when the distribution of grades is considered for both methods, by far the most tumors were assessed as grade 1 or 2 (Table 2). Preponderance of such a relatively large number of well-differentiated tumors may be attributed to the inclusion of only prostatectomy specimens in this study.

In contrast to the low kappa score obtained by the Mostofi method, the percentual agreement was markedly better (Table 3). In fact, in 90% of the cases, at least 3 observers shared the same opinion about the assignment of tumor grade (Table 3). However, when this agreement per tumor grade was analyzed (data not shown), concordance was achieved most frequently for grade 2 tumors. Discrepancy of kappa values and percentual agreement score may be explained by the fact that observers are often inclined to classify tumors most frequently as grade 2 when a three-step grading method is used. In addition, inadequate preservation of tissue slides has undoubtedly influenced precise judgement of cytological features such as degree of anaplasia. Furthermore, lack of sufficient experience and poor definability of slight, moderate and marked nuclear anaplasia may also contribute to this low kappa value.

Applying the modified Gleason method, a maximal numerical agreement score was achieved. In all specimens investigated, at least 3 observers agreed after rearrangement of results into a three-step grading system (100%). However, a low kappa value was a measure reflecting the low degree of accuracy. The latter may be explained by an unbalanced distribution, since grade 1 and 3 tumors are scarcely represented (Table 2). Although a kappa value was not determined for the traditional Gleason system, our results seem to be in contrast to those presented by others [16, 22, 24, 26, 29, 37].

Table 4. Measure of agreement of each participating pathologist (A–E) compared to the scores of Mostofi (M) using the Mostofi-Schroeder method (1–5)

Pathologist A

	1(M)	2(M)	3(M)	4(M)	5(M)	
1(A)	0	0	0	0	0	0
2(A)	1	2	2	2	1	8
3(A)	0	0	10	9	3	22
4(A)	0	0	9	5	4	18
5(A)	0	1	0	1	0	2
	1	3	21	17	8	50

Weighted kappa coefficient A k = 0.14

Pathologist B

	1(M)	2(M)	3(M)	4(M)	5(M)	
1(B)	0	1	0	1	0	2
2(B)	0	1	4	1	1	7
3(B)	1	1	10	4	1	17
4(B)	0	0	6	8	5	19
5(B)	0	0	1	3	1	5
	1	3	21	17	8	50

Weighted kappa coefficient B k = 0.35

Pathologist C

	1(M)	2(M)	3(M)	4(M)	5(M)	
1(C)	1	0	0	0	0	1
2(C)	0	1	2	0	0	3
3(C)	0	1	6	3	2	12
4(C)	0	1	10	9	2	22
5(C)	0	0	3	5	4	12
	1	3	21	17	8	50

Weighted kappa coefficient C k = 0.46

Pathologist D

	1(M)	2(M)	3(M)	4(M)	5(M)	
1(D)	0	0	0	0	0	0
2(D)	1	0	0	1	0	2
3(D)	0	3	10	3	2	18
4(D)	0	0	8	7	2	17
5(D)	0	0	3	6	4	13
	1	3	21	17	8	50

Weighted kappa coefficient D k = 0.42

Pathologist E

	1(M)	2(M)	3(M)	4(M)	5(M)	
1(E)	0	0	1	1	0	2
2(E)	1	2	7	3	1	14
3(E)	0	1	11	8	2	22
4(E)	0	0	1	4	3	8
5(E)	0	0	1	1	2	4
	1	3	21	17	8	50

Weighted kappa coefficient E k = 0.35

Harada et al. [23] stated that, for the primary growth pattern, agreement was found in 64% of cases; for the secondary pattern, agreement was 44%. When the primary and secondary patterns were compared to those of Gleason, agreement was recorded in only 38% [23]. Gleason [23] postulates that the margin of error of reproducibility from one institution to another could be as much as 50% and probably reflects the degree of experience of the particular observer. Bain et al. [3] stated in their study concerning the reproducibility of the Gleason method that agreement was reached among 7 pathologists in 74%–93% of the cases studied ($N = 58$; kappa ranged from 60.5% to 83.6%). However, grading in this study was performed predominantly upon transurethral resection specimens.

Although the percentual numerical scores of the Schroeder-Mostofi method demonstrate a nearly total lack of agreement (Table 3), the weighted kappa coefficient can be considered as relatively high (0.34). Again, lack of numerical concordance may be partly due to the fact that, in addition to histological criteria, cytological features, especially nuclear anaplasia and mitotic activity are incorporated in this grading system. Ordinal scale correction for this five-step grading method has undoubtedly influenced the kappa score favorably.

As the treatment decision for prostatic cancer is often influenced by the results of grading, accuracy and reproducibility of grading methods are of the utmost importance. However, even when sufficient tissue sections are available, accuracy of grading, as this study shows, is considered to be relatively low. Simple methods involving only histological features seem to have a better interobserver agreement than those systems using cytological features. However, low agreement scores found for the Mostofi method and the recently described system introduced by Schroeder and Mostofi may be attributed to lack of optimal fixation and preservation of tissue sections, resulting in difficulty of evaluating nuclear anaplasia. In addition, our results may also indicate that, in practice, pathologists are more accustomed to dealing with well-defined (histological) criteria as compared to less circumscribed criteria such as degree of nuclear anaplasia.

Calculation of the weighted kappa scores may hamper precise interpretation of the accuracy of grading scores, especially when the tumors selected are not equally divided among the diverse grades. Finally, reproducibility is largely dependent upon training and experience and is considerably facilitated by the elimination of ambiguity in the definition of predictive morphological

criteria. As yet, it remains to be determined to what extent morphometrical techniques will enhance the accuracy and reproducibility of grading methods.

Since the final aim of grading is to predict the clinical behavior of neoplasms, grading results obtained so far will be correlated with survival and recurrence rates and the data presented in a second report.

Acknowledgements. The authors are indebted to Miss M. Hanegraaff for preparation of this manuscript.

References

1. Albertsen P (1982) Histologic grading and the practicing urologist. The Prostate 3:333
2. Babaian RJ, Grunow WA (1985) Reliability of Gleason grading system in comparing prostate biopsies with total prostatectomy specimens. Urology 25:564
3. Bain GO, Koch M, Hanson J (1982) Feasibility of grading prostatic carcinomas. Arch Pathol Lab Med 106:265
4. Belt E, Schroeder FH (1972) Total perineal prostatectomy for carcinoma of the prostate. J Urol 107:91
5. Böcking A, Kiehn J, Heinzel-Wach M (1982) Combined histologic grading of prostatic carcinoma. Cancer 50:288
6. Brawn PN, Ayala AG, Von Eschenbach AC, Hussey DH, Johnson DE (1982) Histologic grading study of a new system and comparison with other methods - a preliminary study. Cancer 49:525
7. Broders AC (1926) Grading and practical application. Arch Path and Lab Med 1:376
8. Catalona WJ (1984) Prostatic Cancer. Grune and Stratton, Orlando
9. Diamond DA, Berry SJ, Umbricht C, Jewett HJ, Coffey DS (1982) Computerized image analysis of nuclear shape as a prognostic factor for prostatic cancer. The Prostate 3:321
10. Diamond DA, Berry SJ, Jewett HJ, Eggleston JC, Coffey DS (1982) A new method to assess metastatic potential of human prostate cancer: Relative nuclear roundness. J Urol 128:729
11. Emmett JL, Greene LF, Papantoniou A (1960) Endocrine therapy in carcinoma of the prostate gland: ten year survival studies. J Urol 83:471
12. Epstein JI, Berry SJ, Eggleston JC (1984) Nuclear roundness factor: A predictor of progression in untreated stage A2 prostate cancer. Cancer 54:1666
13. Epstein NA, Fatti LP (1976) Prostatic carcinoma: Some morphological features affecting prognosis. Cancer 37:2455
14. Fowler JE Jr, Mills SE (1985) Operable prostatic carcinoma: Correlations among clinical stage, pathological stage, Gleason histological score and early disease-free survival. J Urol 133:49
15. Gaeta JF (1981) Glandular profiles and cellular patterns in prostatic cancer grading. Urology [Suppl] 17:33
16. Gaeta JF, Englander LC, Murphy GP (1986) Comparative evaluation of national prostatic cancer treatment group and Gleason systems for pathologic grading of prostatic cancer. Urology 27:306
17. Gardner WA Jr (1982) Histologic grading of prostate cancer: A retrospective and prospective overview. The Prostate 3:555
18. Gleason DF (1966) Classification of prostatic carcinomas. Cancer Chemother Rep 50:125
19. Gleason DF, Mellinger GT, The Veterans Administration Cooperative Urological Research Group (1974) Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging. J Urol 111:58
20. Gleason DF, The Veterans Administration Cooperative Urological Research Group (1977) Histologic grading and clinical staging of prostatic carcinoma. In: Tannenbaum M (ed) Urologic Pathology: The Prostate, Lea & Febiger, Philadelphia, pp 177-198
21. Grayhack JT, Assimos DG (1983) Prognostic significance of tumor grade and stage in the patient with carcinoma of the prostate. The Prostate 4:13
22. Guinan P, Talluri K, Nagubadi S, Sharifi R, Ray V, Shaw M (1983) Evaluation of Gleason classification system in prostate cancer. Urol 21:458
23. Harada M, Mostofi FK, Corle DK, Byar DP, Trump BF (1977) Preliminary studies of histologic prognoses in cancer of the prostate. Cancer Treat Rep 61:223
24. Kramer SA, Spähr J, Brendler CB, Glenn JF, Paulson DF (1980) Experience with Gleason's histopathologic grading in prostatic cancer. J Urol 124:223
25. Mellinger GT, Gleason D, Bailar J (1967) The histology and prognosis of prostatic cancer. J Urol 97:331
26. Mills SE, Fowler JE (1986) Gleason histologic grading of prostatic carcinoma; correlations between biopsy and prostatectomy specimens. Cancer 57:346
27. Mostofi FK (1975) Grading of prostatic carcinoma. Cancer Chem Rep 59:111
28. Mostofi FK (1976) Problems of grading carcinoma of the prostate. Sem in Oncol 3:161
29. Murphey GP, Whitmore WF (1979) A report of the workshops on the current status of the histologic grading of prostate cancer. Cancer 44:1490
30. Myers RP, Neves RJ, Farrow GM, Utz D (1982) Nucleolar grading of prostatic adenocarcinoma: light-microscopic correlation with disease progression. The Prostate 3:42
31. Schouten HJA (1982) Measuring pairwise interobserver agreement when all subjects are judged by the same observers. Statistica Neerlandica 36:45
32. Schroeder FH, Belt E (1975) Carcinoma of the prostate: a study of 213 patients with stage C tumors treated by total perineal prostatectomy. J Urol 114:257
33. Schroeder FH, Blom JHM, Hop WCJ, Mostofi FK (1985) Grading of prostatic cancer: I. An analysis of the prognostic significance of single characteristics. The Prostate 6:81
34. Schroeder FH, Blom JHM, Hop WCJ, Mostofi FK (1985) Grading of prostatic cancer: II. The prognostic significance of the presence of multiple architectural patterns. The Prostate 6:403
35. Schroeder FH, Hop WCJ, Blom JHM, Mostofi FK (1985) Grading of prostatic cancer: III. Multivariate analysis of prognostic parameters. The Prostate 7:13
36. Smith PH, Robinson MRG, Coopers EH (1976) Carcinoma of the prostate, a focal point for divergent disciplines. Eur J Cancer 12:937
37. Sogani PC, Israel A, Lieberman PH, Lesser ML, Whitmore WF (1985) Gleason grading of prostate cancer: a predictor of survival. Urology 25:223
38. Tannenbaum M, Tannenbaum S, DeSanctis PN, Olsson CA (1982) Prognostic significance of nucleolar surface area in prostate cancer. Urology 19:546
39. Wilson JWL, Morales A, Bruce AW (1983) The prognostic significance of histological grading and pathological staging in carcinoma of the prostate. J Urol 130:481