

Is there a difference in outcome after radical prostatectomy between patients with biopsy Gleason sums 4, 5, and 6? results from the SEARCH database

The SEARCH Database Study Group; SJ Freedland¹, CL Amling², MK Terris³, JC Presti Jr⁴, WJ Aronson^{1,5}, D Elashoff⁶ & CJ Kane^{7,8}

¹Department of Urology, UCLA School of Medicine, Los Angeles, California, USA;

²Department of Urology, San Diego Naval Hospital, San Diego, California, USA;

³Division of Urology, Medical College of Georgia, Augusta, Georgia, USA; ⁴Department of Urology, Stanford University School of Medicine, Palo Alto, California, USA;

⁵Department of Surgery, Veterans Administration Greater Los Angeles Healthcare System, Los Angeles, California, USA; ⁶Department of Biostatistics, UCLA School of Medicine, Los Angeles, California, USA; and ⁷Urology Section, Department of Surgery, Veterans Administration Medical Center San Francisco, California, USA; ⁸Department of Urology, University of California, San Francisco

Purpose: Fewer patients newly diagnosed with prostate cancer today have biopsy Gleason sums <6 compared to several years ago. Several tables and nomograms for predicting disease recurrence after definitive therapy provide little or no discrimination between biopsy Gleason sums 4, 5, and 6. We sought to examine the significance of biopsy Gleason sum for predicting biochemical failure following radical prostatectomy (RP) for men with biopsy Gleason sums of 4, 5, and 6.

Materials and methods: We examined data from 988 men treated with RP between 1988 and 2002 who had biopsy Gleason sums of 4–6. Clinical and pathological variables as well as outcome information were compared between men with biopsy Gleason sums of 4–6. The log-rank and Cox proportional hazards analysis were used to determine whether biopsy Gleason sum provided unique prognostic information for men with low biopsy Gleason sums undergoing RP.

Results: There was statistically significant, but overall weak correlation between biopsy Gleason sum and Gleason sum of the RP specimen (Spearman's $r = 0.277$, $P < 0.001$). As biopsy Gleason sum increased from 4 to 5 to 6, there was a steady rise (HR = 1.31 for each one point increase in Gleason sum, Cox's model) in the risk of PSA failure ($P = 0.025$, log-rank). On multivariate analysis comparing biopsy Gleason sum, preoperative PSA, clinical stage, year of surgery, percent of biopsy cores positive, and age for their ability to predict time to biochemical recurrence, only PSA (HR 2.09, CI 1.56–2.80, $P < 0.001$) and biopsy Gleason sum (HR 1.33, CI 1.05–1.70, $P = 0.019$) were significant independent predictors of PSA failure.

Conclusions: Despite weak correlation between biopsy and pathologic Gleason sum among men with biopsy Gleason sum 4–6 tumors, grade was a significant independent predictor of PSA failure following RP. In the range of 4–6, biopsy Gleason sum acted as a continuous variable for predicting PSA failure. The routine use of Gleason sums 4 and 5 to grade prostate needle biopsy specimens should not be abandoned.

Prostate Cancer and Prostatic Diseases (2003) 6, 261–265. doi:10.1038/sj.pcan.4500673

Keywords: radical prostatectomy; PSA recurrence; Gleason sum; needle biopsy

*Correspondence: SJ Freedland, Department of Urology, UCLA School of Medicine, Box 951738, Los Angeles, CA 90095-1738, USA.
E-mail: sfreedla@ucla.edu

Received 12 March 2003; revised 13 May 2003; accepted 16 May 2003

Introduction

Biopsy grade is an important predictor of outcome following therapy for prostate cancer.^{1,2} As originally described, Gleason sum is determined by combining scores from the predominant and second most prevalent cell architectures, with possible grades from 1 to 5, generating sums from 2 to 10.³ However, primary grade 1 is infrequently used, resulting in few Gleason sum 2 or 3 tumors. Moreover, there has been a recent trend towards using Gleason grade 2 less, resulting in fewer Gleason sums 4 and 5 and an increased incidence of Gleason sum 6 tumors.⁴⁻⁷ Whether this reflects a changing natural history or simply different pathologic interpretation of similar tumors is unclear.⁸ Since some pathologists recommend that a diagnosis of Gleason sums 2-4 cancer not be made on needle biopsy specimens,⁹ systematic upgrading may be the cause of this upward grade migration. Reflecting this, nomograms and tables for predicting adverse pathology or biochemical recurrence following radical prostatectomy (RP) have provided either no risk separation^{1,10} between patients with biopsy Gleason sums 4-6, or essentially combined Gleason sums 5 and 6.^{2,11} However, other studies found that biopsy Gleason sums 4-6 do portend different cancer-specific survival risks.¹² Thus, it is unclear whether this trend toward decreasing use of Gleason sums 4 and 5 in favor of 6 on needle biopsy specimens is justified.

We sought to determine the significance of Gleason sum in predicting biochemical recurrence following RP among men with biopsy Gleason sums 4-6. To accomplish this, we utilized the multicenter Shared Equal Access Regional Cancer Hospital (SEARCH) Database¹³ of men treated with RP.

Materials and methods

After obtaining Institutional Review Board approval from each institution, data from 1670 consecutive patients (excluding patients treated with preoperative androgen deprivation or radiation therapy) treated with RP from 1988 to 2002 at the West Los Angeles, Palo Alto, San Francisco, and Augusta, Georgia VAMCs, and the San Diego Naval Medical Center were combined into the SEARCH Database. Patients with biopsy Gleason sums ≥ 7 ($n=459$), <4 ($n=32$), or unknown ($n=191$) were excluded, resulting in a study population of 988, categorized as white ($n=578$), black ($n=253$), Hispanic ($n=51$), Asian (including Filipino, $n=15$), or other or unknown ($n=89$).

All biopsy samples were taken via ultrasound guidance. Percent of cores positive was calculated by dividing the number of cores positive by the total number of cores. A total of 178 patients did not have data on the percent of cores positive. Data on the percent of cores positive for patients with <4 cores obtained ($n=43$) were considered missing, to eliminate the impact of lesion directed biopsies. These 221 patients were included for comparing outcomes among men with various biopsy Gleason sums, but not for comparing Gleason sum to other clinical variables for their ability to predict PSA failure. The remaining patients had between

4 and 30 cores obtained, of which 44% had sextant (six cores) biopsies performed.

The prostatectomy specimens were sectioned per each institution's protocol.¹⁴⁻¹⁷ Patients were followed to determine PSA recurrence. Recurrence was defined as a single PSA >0.2 ng/ml or two values at 0.2 ng/ml. Patients with no follow-up data ($n=46$) were included for evaluating differences in preoperative and pathological characteristics, but not biochemical recurrence.

Statistics

Clinical and pathological characteristics were compared among patients with biopsy Gleason sums 4-6 tumors using an analysis of variance model (ANOVA) for continuous variables or χ^2 for categorical variables. Correlation between variables was performed using the Spearman correlation coefficient (r). PSA, age, year of surgery, and percent of cores with cancer were examined as continuous variables while clinical stage (T1 vs T2/T3) was examined as a categorical variable. PSA was examined using the logarithmic transformation of PSA plus one to avoid large negative values for patients with very low serum PSA values. Time to recurrence was compared across the groups using a log-rank survivorship analysis and a Cox proportional hazards model. For multivariate analysis, a forwards-stepwise Cox proportional hazards model was used with $P<0.15$ determining which variables should be entered into the model at each step. The variable with the highest P -value was successively deleted until only variables with $P<0.1$ remained. All clinical (age, clinical stage, PSA, and biopsy Gleason sum) and pathological variables (surgical Gleason sum, pathological stage, extracapsular extension rates, margin status, seminal vesicle invasion, and lymph node positivity) were similar between the centers, and therefore data were combined for analysis. All statistical analyses were performed using STATA 7.0 (Stata Corp., College Station, TX, USA).

Results

Table 1 demonstrates the clinical and pathological characteristics of the study population separated by biopsy Gleason sum. Over time, there was a steady increase in the number of men with biopsy Gleason sum 6 tumors and a decrease in the percentage of men with biopsy Gleason sum 4 and 5 tumors (Spearman, year of surgery vs Gleason sum $r=0.466$, $P<0.001$). Thus, the overall changing demographics of patients undergoing RP likely explains the fact that patients with higher biopsy Gleason sums (more recently treated patients) were younger and had a higher rate of clinical T1 disease. Patients with higher biopsy Gleason sums had higher pathological Gleason sums, and there was a trend for a higher rate of extracapsular penetration and higher pathological stage. The correlation between biopsy Gleason sum and pathological Gleason sum, while statistically significant, was weak (Spearman $r=0.277$, $P<0.001$). Among men with biopsy Gleason sum 4, only 16% had the same grade in the RP specimen and 24% had a pathological Gleason sum ≥ 7 .

Table 1 Clinical and pathological features of men undergoing radical prostatectomy

	Biopsy Gleason sum			P-value
	4	5	6	
No. patients	159	262	567	
Mean age±s.d. (y)	64.0±7.0	62.9±6.6	61.8±6.7	0.001
PSA (ng/ml)				
Median	7.7	6.9	7.5	0.532
Mean±s.d.	9.6±8.0	9.1±8.1	9.7±7.8	
Percent of cores positive (%)				<0.001
Median	25	33	25	
Mean±s.d.	30.2±17.0	39.8±23.4	31.6±21.5	
Clinical stage (n)				<0.001
T1	59 (39%)	96 (38%)	296 (53%)	
T2	92 (61%)	156 (62%)	258 (47%)	
T3	0 (0%)	1 (0%)	0 (0%)	
Pathological Gleason sum (n)				<0.001
2–3	2 (1%)	2 (1%)	3 (1%)	
4	23 (16%)	7 (3%)	10 (2%)	
5	56 (39%)	80 (32%)	51 (9%)	
6	28 (20%)	84 (33%)	278 (50%)	
7	30 (21%)	70 (28%)	188 (34%)	
8–10	4 (3%)	8 (3%)	24 (4%)	
Median	6	6	7	
Pathological stage				0.076
T2	119 (76%)	207 (80%)	400 (73%)	
T3	30 (19%)	46 (18%)	158 (25%)	
T4	7 (4%)	6 (2%)	13 (2%)	
Positive surgical margins	47 (30%)	84 (32%)	170 (31%)	0.840
Capsular penetration	28 (18%)	43 (17%)	125 (23%)	0.096
Seminal vesicle invasion	8 (5%)	13 (5%)	38 (7%)	0.488
Lymph node involvement	2 (1%)	7 (3%)	6 (1%)	0.238

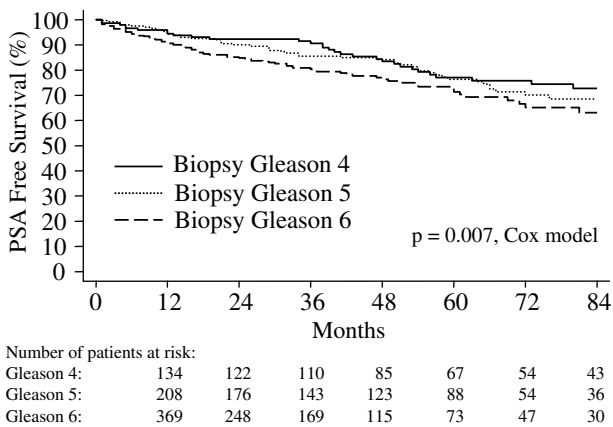


Figure 1 Actuarial 7-y biochemical recurrence rates of patients treated with radical prostatectomy segregated by biopsy gleason sum.

Median (mean) follow-up by Gleason sum was 61 (64) months for Gleason sum 4, 47 (49) months for Gleason sum 5, and 23 (31) months for Gleason sum 6 patients. Median year of surgery by Gleason sum was 1993 for Gleason sum 4, 1994 for Gleason sum 5, and 1997 for Gleason sum 6 patients. During follow-up, 187 patients (20%) developed a biochemical recurrence. Using a Cox proportional hazards model, PSA recurrence rates were compared between men with biopsy Gleason sums 4–6 (Figure 1, Table 2). Men with Gleason sum 6 had significantly higher recurrence rates than men with Gleason sum 4. The differences in PSA recurrence

Table 2 Evaluation of biopsy Gleason sum as a predictor of time to PSA failure following radical prostatectomy

	Referent group	
	Biopsy Gleason sum 4	Biopsy Gleason sum 5
Biopsy Gleason sum 5		
Hazard ratio	1.31	—
P-value*	0.227	—
Biopsy Gleason sum 6		
Hazard ratio	1.71	1.31
P-value*	0.010	0.113

*P-value and Hazard ratio from Cox's proportional hazards analysis.

between biopsy Gleason sum 5 relative to 4 and 6 relative to 5 were not significant. However, both comparisons showed a trend for an identical increased risk of recurrence among higher Gleason sums (HR = 1.31). There were no differences in outcome when Gleason sum 5 was separated into 2+3 vs 3+2 ($P = 0.701$, Cox model). The continuous nature of Gleason sum as a predictor for biochemical failure was further confirmed using log-rank analyses, which demonstrated that as biopsy Gleason sum increased from 4 to 5 to 6, the number of observed PSA failures relative to the number expected steadily increased ($P = 0.025$, Table 3).

Given that biopsy Gleason sums between 4, 5, and 6 gave significant prediction for PSA failure in univariate analysis, we sought to compare grade to other clinical

Table 3 Log-rank analysis of biopsy Gleason sum for predicting biochemical recurrence following radical prostatectomy

Biopsy Gleason sum	No. of patients	Observed events	Expected events	Ratio of observed to expected events
4	159	32	44.42	0.72
5	262	56	60.01	0.93
6	567	99	82.57	1.20

Log-rank for all groups, $P=0.025$; Gleason 4 vs 5, $P=0.200$; Gleason 4 vs 6, $P=0.019$; Gleason 5 vs 6, $P=0.111$.

variables including PSA, clinical stage, year of surgery, percent of biopsy cores positive and age to determine the best preoperative predictors of biochemical recurrence. On multivariate analysis, PSA (HR 2.09, CI 1.56–2.80, $P<0.001$) and biopsy Gleason sum (HR 1.33, CI 1.05–1.70, $P=0.019$) were the only significant independent predictors of PSA failure. There was a trend for higher percent of cores positive to be associated with recurrence (HR 1.90, CI 0.93–3.90, $P=0.079$), which did not reach statistical significance.

As mean Gleason sum changed over the 14-y period of the current study, we examined the value of Gleason grading for predicting time to PSA recurrence during each 3-y span since the beginning of the study period in 1988. During each 3-y span from 1988 to 1990, 1991 to 1993, and 1994 to 1996, there was a consistent trend for increasing biopsy Gleason sums from 4 to 5 to 6 to be associated with an increased risk of PSA recurrence (Cox model, HR = 1.28–1.50). Owing to limited patient numbers, this was only statistically significant from 1991 to 1993 ($P=0.021$). From 1997 to 2000 and 2001 to 2002, biopsy Gleason sum was not correlated with PSA outcomes. However, it should be noted that from 1997 to 2000, 68% of patients had Gleason 6 sum tumors, and from 2001 to 2002, 92% of patients had Gleason sum 6 tumors.

Discussion

There has been a recent trend towards a decreasing incidence of biopsy Gleason sum 4 and 5 tumors, and an increasing incidence of Gleason sum 6 among men undergoing RP.^{4–7} This probably reflects a trend in the pathology community away from assigning Gleason grade 2 towards assigning grade 3.^{8, 9} However, whether this reduction in the number of tumors graded as Gleason sum 4 or 5 relative to 6 is justified by the finding that outcomes among men are similar across the range of Gleason sum 4–6 is controversial.^{2, 10, 12, 18} To test whether men with low biopsy Gleason sums have different outcomes following RP, we examined data from a large multicenter database, the SEARCH Database. Among men with biopsy Gleason sums 4–6 tumors, we found that biopsy Gleason sum was a significant predictor of PSA failure following RP. Specifically, there was a steady increase in the risk of biochemical recurrence as biopsy grade increased from 4 to 5 to 6. After controlling for other preoperative clinical variables, biopsy Gleason sum remained a significant predictor of PSA failure. However, due to the systematic upgrading

that has occurred over the past several years, the significant risk stratification provided by Gleason sums 4 and 5 has been lost.

Many studies examined the role of Gleason grade in predicting outcome among prostate cancer patients.^{2,10,12,18,19} Clearly, the presence and amount of high-grade tumor (Gleason grade 4 or 5) in the biopsy specimen is associated with worse outcomes following RP.¹⁹ However, whether the Gleason grading system maintains its prognostic significance among patients without high-grade tumor is unclear. Albertsen *et al*,¹² found that men with Gleason sum 6 tumors were at higher risk of prostate cancer death than men with Gleason sum 5 tumors who in turn were at higher risk than men with Gleason sum 4. These results are consistent with our findings. However, others found no significant differences between Gleason sums 4–6.^{10,18} Yet other studies found Gleason sum 4 tumors behaved similarly to Gleason sum 5, but that Gleason sum 6 tumors were higher risk.² Finally, others found no differences between Gleason sum 5 and 6 tumors, but that Gleason sum 4 tumors were lower risk.¹¹ We found that as biopsy grade increased from 4 to 5 to 6, the risk of PSA failure following RP steadily rose. This steady rise in PSA recurrence risk argues that in the range of 4–6, biopsy Gleason sum acts as a continuous variable with an equivalent increase in risk of recurrence for each successive rise in grade.

In the current analysis, we did not include men with biopsy Gleason sum 2 or 3 tumors, largely due to the limited number of patients ($n=32$). Alternatively, we could have combined them with Gleason sum 4 patients. However, whether it is valid to combine these groups, although often done,^{2,11,12} is unknown. Moreover, none of these 32 patients were treated in the past 4 years, indicating that trying to predict outcome following RP for a patient with a biopsy Gleason sum 2 or 3 is an uncommon clinical scenario today.

Some urologic pathologists have argued that Gleason sums 2–4 should not be made on prostate needle biopsy specimens for several reasons.⁹ First, the vast majority of biopsy Gleason sums 2–4 tumors diagnosed by community pathologists were upgraded when analyzed by urologic pathologists.²⁰ Second, there is poor reproducibility between biopsy and RP specimen grading and patients with low-grade tumors are often found to have higher-grade disease in the RP specimen. Third, there is often poor agreement among urologic pathologists in assigning a particular grade.²¹ Finally, there is concern that men with low-grade tumors would be encouraged to undergo expectant therapy or be improperly counseled as to the risk of tumor progression. While we agree that these concerns are valid, the true test of any grading system is whether it provides accurate preoperative prognostic information. Our data show that there are significant differences in the risk of PSA recurrence for each of these low Gleason sums and support the continued reporting of all Gleason sums between 4 and 6 on needle biopsy specimens.

Bostwick²² found that the greatest ‘error’ in assigning a grade occurred among small volume tumors. Thus, one could postulate that low Gleason sums are associated with lower volume disease resulting in improved PSA control rates confounding any analysis of Gleason sum. However, we saw no consistent trend in the relationship

between percent of cores positive, a proxy of tumor volume, and biopsy Gleason sum. Specifically, patients with biopsy Gleason sum 4 tumors had similar percent of cores positive as patients with Gleason sum 6. Even after controlling for percent of cores positive, lower Gleason sums had lower PSA recurrence rates.

Limitations to the current study were that it was retrospective and the mean follow-up was relatively short. In addition, noncentralized pathologic evaluation of biopsy specimens from multiple institutions may make pooling of data problematic. However, it should be noted that all patients were treated since the publication and widespread acceptance of uniform prostate cancer grading guidelines.²³ Moreover, our biopsy data are representative of what most practicing urologists can expect to see in their own practice and as such, may have broader applicability than single institutional series using a dedicated urologic pathologist.

Conclusions

As biopsy Gleason sum increased from 4 to 5 to 6, the risk of PSA failure following RP steadily rose. Biopsy Gleason sum in the range of 4–6 remained a significant independent predictor of PSA failure after controlling for other preoperative clinical variables. However, due to systematic upgrading that has occurred over the past several years, the significant risk stratification provided by Gleason sums 4 and 5 has been lost. The routine use of Gleason sums 4 and 5 to grade prostate needle biopsy specimens should not be abandoned. More frequent use of Gleason sums 4 and 5 on prostate needle biopsy may restore the significant risk stratification once provided for by biopsy tumor grade among patients with low-grade tumors.

Acknowledgements

This work was supported by the Department of Veterans Affairs and a Center for Prostate Disease Research (CPDR) grant from the United States Army Medical Research and Materiel Command.

References

- 1 D'Amico AV *et al*. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. *JAMA* 1998; **280**: 969–974.
- 2 Kattan MW *et al*. A preoperative nomogram for disease recurrence following radical prostatectomy for prostate cancer. *J Natl Cancer Inst* 1998; **90**: 766–771.
- 3 Gleason DF. Classification of prostatic carcinomas. *Cancer Chemother Rep* 1966; **50**: 125–128.
- 4 Soh S *et al*. Has there been a recent shift in the pathological features and prognosis of patients treated with radical prostatectomy? *J Urol* 1997; **157**: 2212–2218.

- 5 Schwartz KL, Grignon DJ, Sakr WA, Wood Jr DP. Prostate cancer histologic trends in the metropolitan Detroit area, 1982 to 1996. *Urology* 1999; **53**: 769–774.
- 6 Han M *et al*. Era specific biochemical recurrence-free survival following radical prostatectomy for clinically localized prostate cancer. *J Urol* 2001; **166**: 416–419.
- 7 Freedland SJ *et al*. Time trends in biochemical recurrence after radical prostatectomy: results of the SEARCH database. *Urology* 2003; **61**: 736–741.
- 8 Smith EB *et al*. Gleason scores of prostate biopsy and radical prostatectomy specimens over the past 10 years: is there evidence for systematic upgrading? *Cancer* 2002; **94**: 2282–2287.
- 9 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; **24**: 477–478.
- 10 Nelson CP *et al*. Preoperative parameters for predicting early prostate cancer recurrence after radical prostatectomy. *Urology* 2002; **59**: 740–745 discussion 745–746.
- 11 Partin AW *et al*. Contemporary update of prostate cancer staging nomograms (Partin Tables) for the new millennium. *Urology* 2001; **58**: 843–848.
- 12 Albertsen PC, Hanley JA, Gleason DF, Barry MJ. Competing risk analysis of men aged 55 to 74 years at diagnosis managed conservatively for clinically localized prostate cancer. *JAMA* 1998; **280**: 975–980.
- 13 Freedland SJ *et al*. Race as an outcome predictor after radical prostatectomy: results from the Shared Equal Access Regional Cancer Hospital (SEARCH) database. *Urology* 2002; **60**: 670–674.
- 14 Terris MK. Sensitivity and specificity of sextant biopsies in the detection of prostate cancer: preliminary report. *Urology* 1999; **54**: 486–489.
- 15 Borirakchanyavat S *et al*. Systematic sextant biopsies in the prediction of extracapsular extension at radical prostatectomy. *Urology* 1997; **50**: 373–378.
- 16 Freedland SJ *et al*. Race is not an independent predictor of biochemical recurrence after radical prostatectomy in an equal access medical center. *Urology* 2000; **56**: 87–91.
- 17 Borboroglu PG, Comer SW, Riffenburgh RH, Amling CL. Extensive repeat transrectal ultrasound guided prostate biopsy in patients with previous benign sextant biopsies. *J Urol* 2000; **163**: 158–162.
- 18 Zagars GK, Ayala AG, von Eschenbach AC, Pollack A. The prognostic importance of Gleason grade in prostatic adenocarcinoma: a long-term follow-up study of 648 patients treated with radiation therapy. *Int J Radiat Oncol Biol Phys* 1995; **31**: 237–245.
- 19 McNeal JE *et al*. Histologic differentiation, cancer volume, and pelvic lymph node metastasis in adenocarcinoma of the prostate. *Cancer* 1990; **66**: 1225–1233.
- 20 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; **21**: 566–576.
- 21 Allsbrook Jr WC *et al*. Interobserver reproducibility of Gleason grading of prostatic carcinoma: urologic pathologists. *Hum Pathol* 2001; **32**: 74–80.
- 22 Bostwick DG. Grading prostate cancer. *Am J Clin Pathol* 1994; **102**: S38–56.
- 23 Gardner Jr WA *et al*. A uniform histopathologic grading system for prostate cancer. Subcommittee on Diagnostic Nomenclature, Prostate Cancer Working Group, Organ Systems Program. *Hum Pathol* 1988; **19**: 119–120.