

Prostate Cancer and the Will Rogers Phenomenon

Peter C. Albertsen, James A. Hanley, George H. Barrows, David F. Penson,
Pam D. H. Kowalczyk, M. Melinda Sanders, Judith Fine

Background: Information on tumor stage and grade are used to assess cancer prognosis and to produce standardized comparisons of end results over time. Changes in the interpretation of classification schemes can alter the apparent distribution of cancer stage or grade in the absence of a true biologic change. Since the introduction of prostate-specific antigen testing, the reported incidence of low-grade prostate cancer has declined. To determine whether this decline is in part a result of Gleason score reclassification during the same time period, we documented the potential impact of reclassification between 1992 and 2002 on clinical outcomes. **Methods:** A population-based cohort of 1858 men who were ≤ 75 years of age at diagnosis of prostate cancer in 1990–1992 was assembled retrospectively from the Connecticut Tumor Registry. Histology slides of the diagnostic prostate tissue were retrieved and reread in 2002–2004 by an experienced pathologist blinded to the original Gleason score readings. Prostate cancer mortality rates for the cohort calculated using the original Gleason score readings were compared with those calculated using the contemporary Gleason score readings. Statistical tests were two sided. **Results:** The contemporary Gleason score readings were statistically significantly higher than the original readings (mean score increased from 5.95 to 6.8; difference = 0.85, 95% confidence interval = 0.79 to 0.91; $P < .001$). Consequently, the Gleason score–standardized contemporary prostate cancer mortality rate (1.50 deaths per 100 person-years) appeared to be 28% lower than standardized historical rates (2.08 deaths per 100 person-years), even though the overall outcome was unchanged. This apparent improvement in mortality held for all Gleason score categories. **Conclusions:** In this population, a decline in the reported incidence of low-grade prostate cancers appears to be the result of Gleason score reclassification over the past decade. This reclassification resulted in apparent improvement in clinical outcomes. This finding reflects a statistical artifact known as the Will Rogers phenomenon. [J Natl Cancer Inst 2005;97:1248–53]

Clinicians classify patients with newly diagnosed cancer by stage and grade to assess prognosis. This classification is particularly important for men with prostate cancer because of the extraordinary variability in the potential for disease progression. Tumor grade, stage, and the presence of competing medical hazards are the most powerful predictors of survival (1). As a result of the widespread testing of patients for prostate-specific antigen (PSA) over the past decade, most patients with prostate cancer now present with clinically localized disease, and their tumors are rarely graded with Gleason scores < 6 (2,3).

Many researchers cite improvements in 5- and 10-year biochemical recurrence–free survival after surgery or radiation. Results are usually reported according to patients' diagnostic Gleason scores as evidence of the effectiveness of PSA testing (4,5). PSA testing has advanced the time of diagnosis of prostate cancer by as much as 5–10 years (6). This fact alone will yield dramatic survival rate improvements because patients will appear to live an additional 5–10 years with their diagnosis, even in the absence of any treatment intervention (7).

Clinical outcomes can also be influenced by another factor—Gleason score shift (8–12). Clinicians treating contemporary populations with newly diagnosed prostate cancer rarely encounter men with Gleason score 2–5 disease, whereas two decades ago pathologists used these classifications routinely. Two hypotheses have been proposed to explain this phenomenon. One explanation presumes that PSA testing identifies men with more aggressive tumors; another presumes that pathologists are more hesitant to assign low Gleason scores to contemporary prostate needle biopsy specimens because these scores are frequently upgraded after review of the entire surgical specimen (13). If the latter explanation is correct, the resulting shift in Gleason scores would lead to apparent improvements in survival when, in fact, no such improvements occurred.

To determine whether a Gleason score shift has occurred over the past decade, we asked an experienced pathologist to assign Gleason scores to a large series of prostate cancer biopsies that were performed over a decade earlier. We then compared these contemporary Gleason score assignments with the Gleason scores assigned at the time the biopsy was performed. We also investigated the impact of reclassification on prostate cancer mortality rates.

PATIENTS AND METHODS

Patient Cohort

We assembled information concerning the clinical outcomes of Connecticut residents diagnosed with prostate cancer between January 1, 1990, and December 31, 1992, from the Connecticut

Affiliations of authors: University of Connecticut Health Center, Farmington, CT (PCA, MMS, JF); McGill University, Montreal, PQ, Canada (JAH); St. Francis Hospital and Medical Center, Hartford, CT (GHB, PDHK); University of Southern California, Los Angeles, CA (DFP).

Correspondence to: Peter C. Albertsen, MD, Division of Urology, University of Connecticut Health Center, 263 Farmington Avenue, Farmington, CT 06030–3955 (e-mail: Albertsen@nso.uhc.edu).

See “Notes” following “References.”

DOI: 10.1093/jnci/dji248

© The Author 2005. Published by Oxford University Press. All rights reserved. For Permissions, please e-mail: journals.permissions@oupjournals.org.

Tumor Registry. All men were 75 years of age or younger at the time of diagnosis. Men were excluded if they had a prior diagnosis of cancer (other than nonmelanoma skin cancer), if they were diagnosed with prostate cancer after radical cystoprostatectomy or at autopsy, or if they underwent biopsy outside of Connecticut (except Westerly, RI, near the Connecticut border). We chose the period 1990–1992 for two reasons: (1) to include men diagnosed primarily as a consequence of PSA testing and (2) to obtain a minimum of a 10-year follow-up on all patients included in the eligible cohort.

We initially identified 3739 men as being eligible for participation in the study. After obtaining all applicable state and local institutional review board approvals, we sought permission from Connecticut physicians to contact their patients. Medical records of men ($n = 2335$) who accepted the invitation to participate or who were enrolled via institutional review board waivers were abstracted in physicians' offices. Data collected included information on patients' initial diagnosis, pretreatment clinical stage, pretreatment PSA level, initial biopsy tumor grade, staging, comorbidities at the time of diagnosis assessed using the instrument developed by Charlson et al. (14), and initial treatment selected. Information on the hospital or laboratory that rendered the diagnosis on the original biopsy material, along with the corresponding pathology number, was also collected. Information concerning vital status was obtained from the Connecticut Tumor Registry.

We were able to retrieve original biopsy slides for 1988 (85%) of the men whose medical records were abstracted; original pathology reports with Gleason scores were available for 1858 (80%) of these men. During 2002–2004, a referee pathologist (GHB), who was blinded to the original readings and to the clinical baseline information and outcomes, re-read each of the slides. To ensure the reliability of the referee pathologist, two other pathologists, who are also experienced in the interpretation of prostate cancer, each read a 10% sample (i.e., 184) of the pathology slides.

By the time the re-readings were completed, a total of 308 men had died of prostate cancer among the cohort of 1858 patients.

Statistical Analysis

Mortality rate comparisons were restricted to the 1858 men for whom both original and contemporary Gleason score readings were available. Because no patients were assigned scores of 2 or 3 in the contemporary readings, Gleason scores 2–4 were grouped together, yielding seven Gleason score strata (i.e., 2–4, 5, 6, 7, 8, 9, and 10). Mortality rates were compared for patients in each Gleason score stratum. This type of analysis is best understood by examining a row of Table 3 or a panel of Fig. 2. For example, we compared mortality outcomes of the 454 patients who were classified as having Gleason score 6 in the original readings with those of the 814 patients who were classified as having Gleason score 6 tumors by contemporary readings. For each Gleason score stratum, two cause-specific survival curves for 12 years of follow up were constructed using the Kaplan and Meier method (15), one curve based on original Gleason score readings and one based on contemporary readings. Patients who had died of causes other than prostate cancer were censored at the time of their death.

We also used both regression and nonregression techniques to calculate summary prostate cancer mortality ratios for outcomes

based on the original Gleason score classification and the contemporary Gleason score classification. In each analysis, the dataset contained a total of 3716 records (i.e., 1858 cases \times 2). The survival of each patient was entered twice: once according to his original Gleason score classification and once according to his contemporary Gleason score classification. For the regression approach, we used a Cox proportional hazards regression model in which Gleason score was treated as a categorical variable with seven strata (i.e., six indicator values). The assumption of proportionality of hazards was assessed using graphical methods. The mortality rate ratio was estimated by the exponential of the coefficient of the binary variable that indicated a contemporary rather than an original Gleason score classification. To calculate Gleason score-adjusted (i.e., histology-standardized) cause-specific survival curves, the distribution of each Gleason score was taken to be the average of the original and contemporary distributions of scores (see Fig. 2, H). A stratified Cox model was also fit to the data using the Gleason score classifications as strata.

In the nonregression approach, we used the number of deaths and person-years of follow-up in each of the seven Gleason score strata, coupled with the same standard distribution of scores, to directly calculate the two Gleason score-standardized mortality rates and their ratio. A Mantel-Haenszel summary mortality rate ratio was also calculated from these person-time data.

All analyses used the same 308 deaths among the 1858 patients to create two patient cohorts, one classified by original Gleason score readings and one classified by contemporary Gleason score readings. We analyzed 500 randomly selected bootstrap samples to estimate the 95% confidence intervals (CIs) for the mortality rate ratio. We repeated the comparisons using nonoverlapping series. Specifically, we created and compared two randomly formed halves of the 1858 patient cohort, using original readings from one half and contemporary readings from the other half. All P values and confidence intervals are two-sided. The data were analyzed using SAS version 6.12.

RESULTS

Clinical Characteristics and Changes in Gleason Scores

The clinical characteristics of the study cohort are presented in Table 1. The mean age of the 1858 men in the cohort was 67 years; the men were treated primarily with either surgery or radiation. The distributions of Gleason scores of the 1858 prostate biopsy specimens according to their original reading and their contemporary reading are presented in Fig. 1, and the changes in Gleason scores are presented in Table 2.

Upward shifts in Gleason scores outnumbered the downward shifts by more than 4 to 1. Of the 1858 specimens, scores for 1028 (55%) were upgraded, for 251 (14%) were downgraded, and for 579 (31%) remained unchanged. Overall, the contemporary Gleason score readings were upgraded from an average of just under 6 to an average of 6.8. The average upgrade was 0.85 points (95% CI = 0.79 to 0.91; $P < .001$). Moreover, the re-readings of a 10% sample (i.e., 184 slides) performed by two other pathologists indicated that the referee pathologist who read slides for all 1858 patients reflects contemporary practice. One secondary reviewer assigned scores that were on average 0.66 points lower than those assigned by the referee pathologist and the other secondary reviewer assigned scores that were on average 0.4 points higher than those assigned by the referee pathologist.

Table 1. Clinical characteristics and vital status of 1858 men for whom both an original (1990–1992) and a contemporary (2002–2004) prostate biopsy reading were available*

Characteristic	Value
Average age, y	67 years
Initial PSA score, %	
0–3.9	9
4–9.9	36
10–19	25
20–49	19
50+	11
Origin of specimen, %	
Needle biopsy	93
TURP	7
Clinical impression, %	
Localized disease	80
More advanced disease	20
Initial treatment, %	
Radical prostatectomy	30
Radiation therapy	38
Other	32
Charlson score ≥ 2 †	8
Vital status, n‡	
Alive	1054
Dead of prostate cancer	308
Dead of other causes	496

*PSA = prostate-specific antigen; TURP = transurethral resection of prostate.

†Charlson score (14) refers to a comorbidity scale in which 0 = no comorbidity, 1 = minimal comorbidity, and ≥ 2 = moderate to severe comorbidity.

‡As of March 2004.

Impact of Reclassification on Clinical Outcomes

For each Gleason score stratum, the cause-specific survival of patients given that score in the original Gleason score readings was compared with that of patients given that score in the contemporary Gleason score readings (Fig. 2, A–G). The cause-specific survival curve for patients whose tumor was assigned a specific Gleason score on the contemporary reading was

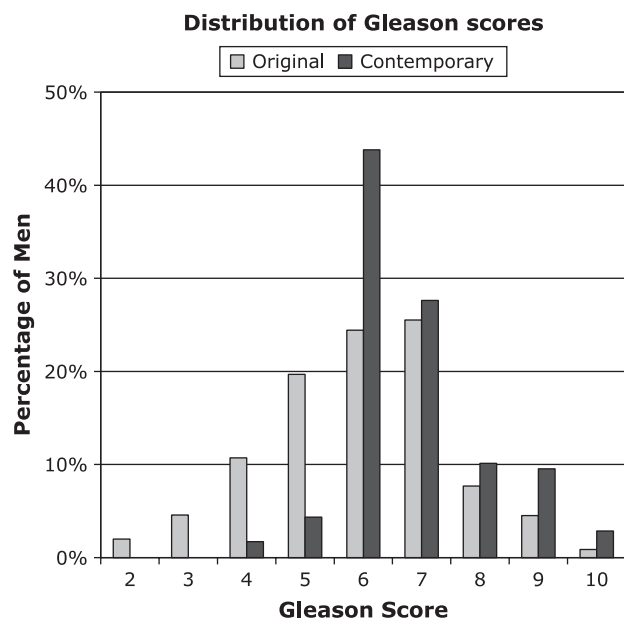


Fig. 1. Distribution of the original and contemporary Gleason score readings for the 1858 men for whom both prostate biopsy specimen readings were available.

consistently better than the cause-specific survival for patients whose tumor was assigned the same Gleason score on the original reading, for each of the Gleason score strata.

In addition, when the score-specific comparisons were aggregated across Gleason scores, a statistically significant improvement in cause-specific survival was observed when patients were classified according to contemporary Gleason scores as compared with original Gleason scores (Fig. 2, H). In the Cox model, the ratio of Gleason score–specific prostate cancer mortality rates for contemporary relative to original scores was 0.74 (bootstrap 95% CI = 0.69 to 0.80; $P < .001$). An identical mortality rate ratio of 0.74 (95% CI = 0.63 to 0.88), corresponding to a 26% reduction in mortality, was also obtained after 500 comparisons using the Gleason scores as strata in the stratified Cox regression model (median P value = .012).

The numbers of prostate cancer deaths and the numbers of man-years of follow-up according to how men were classified by either original or contemporary Gleason score readings are presented in Table 3. We compared the clinical outcomes of these two populations as if they were independent samples and standardized the original and contemporary series of patients by histology (i.e., for the potential differences in the distributions of Gleason scores). The resulting directly standardized mortality rates were 1.5 deaths per 100 person-years for the contemporary series and 2.08 deaths per 100 person-years for the original series. The analysis suggests a 28% [(2.08 – 1.50)/2.08] reduction in mortality. A similar result, i.e., an apparent 26% reduction in mortality, was obtained if the adjusted mortality rate ratios were estimated using the Mantel-Haenszel summary rate ratio.

DISCUSSION

Prostate cancer is now the leading cancer diagnosis in the United States and the second leading cause of cancer death among American men (16). As a consequence, many clinicians and patient advocates promote aggressive testing for serum PSA with the hope of identifying patients with early and potentially more curable prostate cancer. Many tertiary medical centers tout their success in this endeavor by reporting impressive biochemical relapse–free survival rates, defined as either undetectable or stable PSA levels after treatment (4,5).

Unfortunately, several statistical artifacts may be producing a false sense of therapeutic accomplishment. Stage migration and

Table 2. Distribution of the contemporary Gleason score readings for men with each original Gleason score

Original Gleason score (no.)	Contemporary Gleason score									
	2	3	4	5	6	7	8	9	10	
10 (16)	0	0	0	0	1	0	2	6	7	
9 (84)	0	0	0	0	3	15	19	35	12	
8 (143)	0	0	0	1	10	46	30	43	13	
7 (474)	0	0	0	10	107	212	72	62	11	
6 (454)	0	0	3	19	261	110	38	18	5	
5 (366)	0	0	9	25	205	100	17	8	2	
4 (199)	0	0	9	12	146	19	7	4	2	
3 (85)	0	0	5	10	60	6	2	1	1	
2 (37)	0	0	6	4	21	5	1	0	0	

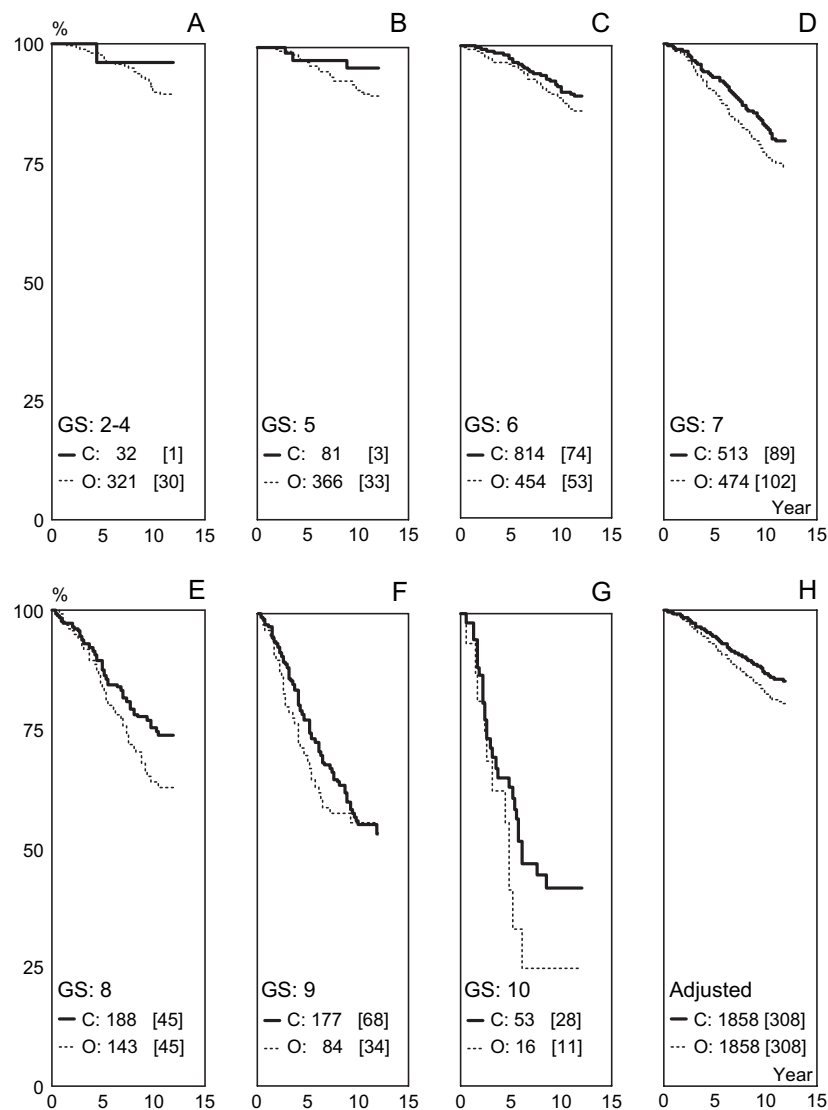


Fig. 2. Cause-specific survival curves for patients whose tumors was assigned the indicated biopsy Gleason score (GS) on the original (O) reading and on the contemporary (C) reading. **Horizontal axis** shows years since diagnosis, and **vertical axis** shows percentage cause-specific survival. Curves for the original reading are shown in **black dashed lines**, and those for the contemporary reading are shown in **black solid lines**. **Numbers** in each panel represent numbers of patients; numbers of deaths from prostate cancer are shown in **square brackets**. **A–G)** Kaplan–Meier cause-specific curves for survival among patients with

Gleason score 2–4 (**A**), Gleason score 5 (**B**), Gleason score 6 (**C**), Gleason score 7 (**D**), Gleason score 8 (**E**), Gleason score 9 (**F**), Gleason score 10 (**G**). **H)** Cause-specific survival curves of the entire contemporary series of 1858 patients (**black solid line**) and the entire historical series (i.e., the same set of 1858 patients, **black dashed line**) from a Cox proportional hazards model that adjusted for the differences in the distribution of Gleason scores between the two series. The contemporary series and the historical series were each standardized to the average of the two distributions of Gleason scores.

grade shift have had particularly profound impacts on prostate cancer outcomes assessment. PSA testing has produced a dramatic stage migration (2,3). Contemporary patients in the United States rarely present with advanced disease. Consequently, contemporary survival analyses include a lead time associated with earlier diagnosis that has been estimated to be between 5 and 10 years when results are compared with historical series (6). Epidemiologists have described this phenomenon as “zero-time shift” or “lead-time bias” (17). Patients appear to have an extension of their survival after cancer diagnosis when they may in fact have experienced no prolongation of their lives.

An equally important but subtler bias that may also be operating to improve apparent prostate cancer outcomes is the Will Rogers phenomenon (18). This term was coined by Feinstein et al. (18), who often quoted a Will Rogers joke that “when the Okies moved to California, the IQ of both states went up.” This

phenomenon can occur when patients are reclassified, as often happens after the introduction of more sensitive staging tools or changes in classification systems. In their original description of the phenomenon, Feinstein et al. focused on stage migration among men with newly diagnosed lung cancer. The phenomenon, however, can occur whenever patients are reclassified—as seen, for example, in the changes in stage-specific survival after the adoption of the 2003 American Joint Committee on staging recommendations for breast cancer (19).

In this analysis, we have demonstrated that a tumor grade shift occurred during the 1990s for men with prostate cancer. Although the Gleason scoring system itself has not changed since the mid-1980s, its application has. Several factors, including the introduction of PSA testing, transrectal ultrasonography, the spring-loaded biopsy gun, and the dramatic increase in the performance of radical prostatectomy, have conspired to produce a

Table 3. Prostate cancer mortality rates among 1858 men classified according to contemporary and original Gleason scores, along with score-specific and standardized mortality rate ratios

Gleason scores	Contemporary Gleason readings		Original Gleason readings		Mortality rate ratio*
	No. of prostate cancer deaths	Man-years (no.)	No. of prostate cancer deaths	Man-years (no.)	
2–4	1	289.7 (32)	30	3225.3 (321)	0.37
5	3	797.3 (81)	33	3675.7 (366)	0.42
6	74	8065.8 (814)	53	4380.8 (454)	0.76
7	89	4840.7 (513)	102	4185.1 (474)	0.75
8	45	1644.0 (188)	45	1127.6 (143)	0.69
9	68	1305.6 (177)	34	592.4 (84)	0.91
10	28	322.1 (53)	11	78.3 (16)	0.62
All	308	17265 (1858)	308	17265 (1858)	1.00
Standardized mortality rate (deaths/100 man-years)†	1.50		2.08		0.72
Rate ratio (95% confidence interval)‡	0.74 (0.69 to 0.80)				

*Mortality rate ratios were calculated by dividing the mortality rate from the original Gleason score readings by the mortality rate calculated from the contemporary Gleason score readings.

†Adjusted for the average Gleason score distribution of the study cohorts.

‡Calculated from a proportional hazards model that includes the Gleason score.

statistically significant upgrading in biopsy Gleason scores, which has, in turn, produced a statistically significant apparent survival improvement in our study cohort. In a traditional comparison of data from two different case series, the impact of this upgrading is often difficult to quantify. In our analysis, by contrast, each person served as his own control, perfectly matched on age, anatomic stage, treatment received, method and duration of follow-up, and the recording of the cause of death. The series differed only by how the Gleason score was applied to the reading of the tumor biopsy.

Researchers conducting an outcomes analysis comparing two separate series of patients usually use several standard statistical tools to adjust for differences in the distribution of Gleason scores to ensure that they are comparing “apples with apples.” Without an adjustment, researchers would be comparing the clinical outcomes of patients with varying levels of tumor aggressiveness rather than the impact of a newer treatment. We applied three of these standard statistical adjustment/standardization techniques to our study cohort that was classified according to historical and contemporary Gleason score readings to determine what impact this would have on the clinical outcomes. In the usual comparison of clinical outcomes involving two separate series of patients it would be difficult to assess how much of any observed reduction in standardized mortality rates is real and how much is an artifact arising from changes in the application of the histology scales, such as the Gleason score. Our results, which are based on the same series of patients, but with their tumors reclassified, demonstrate that contemporary Gleason score readings can yield an apparent statistically significant improvement when clinical outcomes are compared against those of patients classified according to historical Gleason score readings.

Several authors (9–12) have expressed similar concerns about the effect of Gleason score upgrading and the potential impact on mortality rates. Chism et al. (9) reviewed outcomes of 983 prostate cancer patients treated with conformal radiation therapy at the Fox Chase Cancer Center in Philadelphia. They found a systematic Gleason score upgrading of specimens between 1992 and 1997 that led them to suggest that a Gleason score shift may partially explain a statistically significant 5-year improvement in

biochemical relapse-free survival from 68% to 82%. Smith et al. (10) noted a similar phenomenon among men treated with radical surgery. They found that a reinterpretation in 2000 of the original pathology slides by the original pathologist, whose first reading of the slides was in 1989–1991, resulted in a statistically significant Gleason score upgrading of the specimens. Schellhammer et al. (11) also found statistically significant increases in the assigned Gleason score when reevaluating specimens of men who had undergone brachytherapy 15 years earlier. The extent of the upgrading observed in these studies and the magnitude of the changes in cause-specific survival appear to be similar to what we have observed (12).

Several investigators (4,5) have suggested that the application of modern surgical and radiation techniques have resulted in improved outcomes for prostate cancer patients. Han et al. (4) and D’Amico et al. (5) have noted an improvement in biochemical relapse-free survival among contemporary patients compared with patients diagnosed a decade ago. Both of these studies suggest that the improvement noted results from a change in the biologic aggressiveness of prostate cancer at presentation as measured by the Gleason score. However, neither study controlled for changes in the application of Gleason scores over time. Therefore, it is likely that a shift in Gleason scores accounts for a portion or all of the observed time-related improvements.

The shift in the application of Gleason scores documented in our study reflects the change in how pathologists interpret prostate biopsy specimens. Epstein (13) concluded in a 2000 editorial that, because of frequent upgrading of Gleason scores when prostate biopsy specimens are compared with the subsequent surgical specimens, prostate biopsy specimens should not be given Gleason scores of 2–4. In addition, Pan et al. (20) recommended modifying the Gleason score system to reflect the presence of relatively small quantities of poorly differentiated prostate cancer. If adopted, these practices will further contribute to Gleason score shift.

Our study provides some of the strongest evidence to date that the Gleason score shift observed during the past decade is the result of a change in the interpretation of prostate biopsy specimens rather than a selective identification of more aggressive

tumors by PSA testing. The strength of our study stems from the fact that it includes patients who have undergone both surgery and radiation and that it is drawn from community practice.

The primary limitation of our study is that it does not provide a method for quantifying and correcting for a reclassification bias. Researchers reporting improved clinical outcomes when comparing contemporary results with historical case series need to recognize that a portion or all of the reported improvement may simply be the result of Gleason score reclassification. Researchers cannot assume that historical Gleason score readings will be interpreted in the same way by contemporary pathologists. Unless researchers are careful, some or all of an apparent improvement in clinical outcome that is observed when contemporary series are compared with historical series may reflect a statistical artifact—Will Rogers would probably not be amused.

REFERENCES

- (1) Albertsen PC, Fryback DG, Storer DE, Kolon TF, Fine J. Long-term survival among men with conservatively treated localized prostate cancer. *JAMA* 1995;274:69–74.
- (2) Hankey BF, Feuer EJ, Clegg LX, Hayes RB, Legler JM, Prorok PC, et al. Cancer surveillance series: interpreting trends in prostate cancer—Part I: Evidence of the effects of screening in recent prostate cancer incidence, mortality, and survival rates. *J Natl Cancer Inst* 1999;91:1017–24.
- (3) Jhaveri FM, Klein EA, Kupelian PA, Zippe C, Levin HS. Declining rates of extracapsular extension after radical prostatectomy: evidence for continued stage migration. *J Clin Oncol* 1999;17:3167–72.
- (4) Han M, Partin AW, Piantadosi S, Epstein JI, Wash PC. Era specific biochemical recurrence-free survival following radical prostatectomy for clinically localized prostate cancer. *J Urol* 2001;166:416–9.
- (5) D'Amico AV, Chen M, Oh-Ung J, Renshaw AA, Cote K, Loffredo M, et al. Changing prostate specific antigen outcome after surgery or radiotherapy for localized prostate cancer during the prostate specific antigen era. *Int J Radiat Oncol Biol Phys* 2002;54:436–41.
- (6) Draisma G, Boer R, Otto SJ, van der Crujjsen IW, Damhuis RA, Schroder FH, et al. Lead times and overdiagnosis due to prostate-specific antigen screening: estimates from the European Randomized Study of Screening for Prostate Cancer. *J Natl Cancer Inst* 2003;95:868–78.
- (7) Welch HG, Schwartz LM, Woloshin S. Are increasing 5-year survival rates evidence of success against cancer? *JAMA* 2000;283:2975–8.
- (8) Gilliland FD, Gleason DF, Hunt WC, Stone N, Harlan LC, Key CR. Trends in Gleason score for prostate cancer diagnosed between 1983 and 1993. *J Urol* 2001;165:846–50.
- (9) Chism DB, Hanlon AL, Tronsco P, Al-Saleem T, Horowitz EM, Pollack A. The Gleason score shift: score four and seven years ago. *Int J Radiat Oncol Biol Phys* 2003;56:1241–7.
- (10) Smith EB, Frierson HF Jr, Mills SE, Boyd JC, Theodorescu D. Gleason scores of prostate biopsy and radical prostatectomy specimens over the past 10 years. *Cancer* 2002;94:2282–7.
- (11) Schellhammer P, Moriarty R, Bostwick D, Kuban D. Fifteen-year minimum follow-up of a prostate brachytherapy series: comparing the past with the present. *Urology* 2000;56:436–9.
- (12) Kondylis FI, Moriarty RP, Bostwick D, Schellhammer P. Prostate cancer grade assignment: the effect of chronological, interpretive and translation bias. *J Urol* 2003;170:1189–93.
- (13) Epstein JI. Gleason score 2–4 adenocarcinoma of the prostate on needle biopsy. *Am J Surg Pathol* 2000;24:477–8.
- (14) Charlson ME, Pompei P, Alex KL, MacKenzi CR. A new method of classifying prognostic co-morbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40:373–83.
- (15) Kaplan EL, Meier P. Non parametric estimation from incomplete observations. *J Am Stat Assoc* 1958;53:457–81.
- (16) Jemal A, Tiwari RC, Murray T, Ghafoor A, Samels A, Ward E, et al. Cancer statistics, 2004. *CA Cancer J Clin* 2004;54:8–29.
- (17) Hutchinson GB, Shapiro S. Lead time gained by diagnostic screening for breast cancer. *J Natl Cancer Inst* 1968;41:665–73.
- (18) Feinstein AR, Sosin DM, Wells CK. The Will Rogers phenomenon: stage migration and new diagnostic techniques as a source of misleading statistics for survival in cancer. *New Engl J Med* 1985;312:1604–8.
- (19) Woodward WA, Strom EA, Tucker SL, McNeese MD, Perkins GH, Schechter NR, et al. Changes in the 2003 American joint committee cancer staging for breast cancer dramatically affect stage-specific survival. *J Clin Oncol* 2003;21:3244–8.
- (20) Pan C, Potter SR, Partin AW, Epstein JI. The prognostic significance of tertiary Gleason patterns of higher grade in radical prostatectomy specimens. *Am J Surg Pathol* 2000;24:563–9.

NOTES

This study was funded by grants from the Agency for Healthcare Research and Quality (HS09578) and from the Catherine Weldon Donaghue Foundation (DF98–057). Dr. Penson is the recipient of a Veterans Affairs Health Services Research and Development Service Level I Career Development Award. The sponsors of the study had no role in design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

The authors thank the men who participated in this research, and Nancy Dittes, Nancy Hotchkiss, RN, and Susan Walters, RN, for their assistance. This research could not have been accomplished without the participation of the practicing urologists in Connecticut and Westerly, RI, and the staff of the following medical institutions located in Connecticut and Rhode Island: Hartford Hospital, Hartford; Yale New Haven Hospital, New Haven; St. Francis Hospital and Medical Center, Hartford; Bridgeport Hospital, Bridgeport; Waterbury Hospital, Waterbury; Hospital of St. Raphael, New Haven; Danbury Hospital, Danbury; New Britain General Hospital, New Britain; Norwalk Hospital, Norwalk; St. Vincent's Medical Center, Bridgeport; The Stamford Hospital, Stamford; Middlesex Hospital, Middletown; St. Mary's Hospital, Waterbury; Lawrence and Memorial Hospital, New London; Manchester Memorial Hospital, Manchester; The Greenwich Hospital Association, Greenwich; MidState Medical Center, Meriden; Griffin Hospital, Derby; Bristol Hospital, Bristol; John Dempsey Hospital, Farmington; William W. Backus Hospital, Norwich; Charlotte Hungerford Hospital, Torrington; Windham Community Memorial Hospital, Willimantic; Milford Hospital, Milford; Day Kimball Hospital, Putnam; Rockville General Hospital, Rockville; Bradley Memorial Hospital, Southington; The Sharon Hospital, Sharon; New Milford Hospital, New Milford; Johnson Memorial Hospital, Stafford Springs; and Westerly Hospital in Westerly, Rhode Island.

Manuscript received July 26, 2004; revised June 23, 2005; accepted June 30, 2005.