HHS Public Access

Author manuscript

Eur Urol. Author manuscript; available in PMC 2017 March 01.

Published in final edited form as:

Eur Urol. 2016 March ; 69(3): 428–435. doi:10.1016/j.eururo.2015.06.046.

A Contemporary Prostate Cancer Grading System: A Validated Alternative to the Gleason Score

Jonathan I. Epstein[a,*], **Michael J. Zelefsky**[b], **Daniel D. Sjoberg**[b], **Joel B. Nelson**[c], **Lars Egevad**[d], **Cristina Magi-Galluzzi**[e], **Andrew J. Vickers**[b], **Anil V. Parwani**[c], **Victor E. Reuter**[b], **Samson W. Fine**[b], **James A. Eastham**[b], **Peter Wiklund**[d], **Misop Han**[a], **Chandana A. Reddy**[e], **Jay P. Ciezki**[e], **Tommy Nyberg**[d], and **Eric A. Klein**[e]

[a]The Johns Hopkins Medical Institutions, Baltimore, MD, USA

[b]Memorial Sloan Kettering Cancer Center, New York, NY, USA

[c]University of Pittsburgh Medical Center, Pittsburgh, PA, USA

[d]Karolinska Institute, Stockholm, Sweden

[e]Cleveland Clinic, Cleveland, OH, USA

Abstract

Background—Despite revisions in 2005 and 2014, the Gleason prostate cancer (PCa) grading system still has major deficiencies. Combining of Gleason scores into a three-tiered grouping (6, 7, 8–10) is used most frequently for prognostic and therapeutic purposes. The lowest score, assigned 6, may be misunderstood as a cancer in the middle of the grading scale, and 3 + 4 = 7 and 4 + 3 = 7 are often considered the same prognostic group.

*Corresponding author. The Johns Hopkins Hospital, The Weinberg Building, Room 2242, 401 N. Broadway Street, Baltimore, MD 21231, USA. Tel.: + 1 410 955 5043; Fax: +1 443 287 3818. jepstein@jhmi.edu.

Presented in part at the meeting of the United States and Canadian Academy of Pathology (USCAP), Boston, MA, USA, 2015.

Author contributions: Jonathan I. Epstein had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: Epstein.
Acquisition of data: Zelefsky, Nelson, Egevad, Magi-Galluzzi, Sjoberg, Vickers, Parwani, Reuter, Fine, Eastham, Wiklund, Han, Reddy, Epstein, Ciezki, Nyberg, Klein.
Analysis and interpretation of data: Zelefsky, Sjoberg, Epstein, Nyberg, Reddy
Drafting of the manuscript: Epstein.
Critical revision of the manuscript for important intellectual content: Zelefsky, Nelson, Egevad, Magi-Galluzzi, Sjoberg, Vickers, Parwani, Reuter, Fine, Eastham, Wiklund, Han, Epstein, Reddy, Ciezki, Nyberg, Klein.
Statistical analysis: Zelefsky, Sjoberg, Nyberg, Reddy.
Obtaining funding: None.
Administrative, technical, or material support: None.
Supervision: None.
Other (specify): None.

Financial disclosures: Jonathan I. Epstein certifies that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (eg, employment/affiliation, grants or funding, consultancies, honoraria, stock ownership or options, expert testimony, royalties, or patents filed, received, or pending), are the following: None.

Objective—To verify that a new grading system accurately produces a smaller number of grades with the most significant prognostic differences, using multi-institutional and multimodal therapy data.

Design, setting, and participants—Between 2005 and 2014, 20 845 consecutive men were treated by radical prostatectomy at five academic institutions; 5501 men were treated with radiotherapy at two academic institutions.

Outcome measurements and statistical analysis—Outcome was based on biochemical recurrence (BCR). The log-rank test assessed univariable differences in BCR by Gleason score. Separate univariable and multivariable Cox proportional hazards used four possible categorizations of Gleason scores.

Results and limitations—In the surgery cohort, we found large differences in recurrence rates between both Gleason 3 + 4 versus 4 + 3 and Gleason 8 versus 9. The hazard ratios relative to Gleason score 6 were 1.9, 5.1, 8.0, and 11.7 for Gleason scores 3 + 4, 4 + 3, 8, and 9–10, respectively. These differences were attenuated in the radiotherapy cohort as a whole due to increased adjuvant or neoadjuvant hormones for patients with high-grade disease but were clearly seen in patients undergoing radiotherapy only. A five–grade group system had the highest prognostic discrimination for all cohorts on both univariable and multivariable analysis. The major limitation was the unavoidable use of prostate-specific antigen BCR as an end point as opposed to cancer-related death.

Conclusions—The new PCa grading system has these benefits: more accurate grade stratification than current systems, simplified grading system of five grades, and lowest grade is 1, as opposed to 6, with the potential to reduce overtreatment of PCa.

Patient summary—We looked at outcomes for prostate cancer (PCa) treated with radical prostatectomy or radiation therapy and validated a new grading system with more accurate grade stratification than current systems, including a simplified grading system of five grades and a lowest grade is 1, as opposed to 6, with the potential to reduce overtreatment of PCa.

Keywords

Gleason grade; Gleason score

1. Introduction

The current prostate cancer (PCa) grading system was developed between 1966 and 1974 by Donald Gleason and the Veterans Administration Cooperative Urologic Research Group [1]. The system assigns histologic patterns 1 through 5, adding the most and second most common patterns with Gleason scores ranging from 2 to 10. Over the subsequent 40 yr, histologic and clinical diagnosis of PCa along with its treatment has evolved, leading to revisions of the Gleason system first codified in 2005 and more recently in 2014 [2,3]. The current application of Gleason grading differs dramatically from the original system. Scores 2–5 are currently no longer assigned, and certain patterns that Gleason defined as a score of 6 are now graded as 7, thus leading to contemporary Gleason score 6 cancers having a better prognosis than historic score 6 cancers.

There are significant deficiencies with the current application of the Gleason system that have had an impact on patient care. A Gleason score 7 can represent mostly well-differentiated cancer with a lesser component of more poorly differentiated cancer (Gleason 3 + 4 = 7) or mostly poorly differentiated cancer with a smaller component of well-differentiated cancer (4 + 3 = 7). Treatment decisions using a simplified single Gleason score of 7 fail to recognize that 3 + 4 = 7 and 4 + 3 = 7 are prognostically very different. Another critical weakness of the Gleason system is that in practice the lowest score is now assigned a 6, although it is on a scale of 2–10. This leads to a logical yet incorrect assumption on the part of patients that their cancer is in the middle of the scale, compounding the fear of a cancer diagnosis with the belief that the cancer is serious, thus leading to an expectation that treatment is necessary.

In 2013 a new grading system, based on data from Johns Hopkins Hospital, was proposed to address the confusion inherent in the Gleason system [4]. We proposed a five–grade group system based on the much revised original Gleason score: grade group 1 (Gleason score <6), grade group 2 (Gleason score 3 + 4 = 7), grade group 3 (Gleason score 4 + 3 = 7), grade group 4 (Gleason score 8), and grade group 5 (Gleason score 9–10). This new grading system beginning with grade group 1 has the potential benefit of reducing fear and may contribute to a decrease in the overtreatment of low-grade PCa detected by prostate-specific antigen (PSA) screening. The current study was conducted to verify that this new grading system accurately produces a smaller number of grades with the most significant prognostic differences, using multi-institutional and multimodal therapy data.

2. Materials and methods

Between 2005 and 2014, 20 845 consecutive men with clinically localized PCa were treated by radical prostatectomy (RP) at the Cleveland Clinic (Cleveland, OH, USA), Memorial Sloan Kettering Cancer Center (MSKCC; New York, NY, USA), Johns Hopkins Hospital (Baltimore, MD, USA), University of Pittsburgh (Pittsburgh, PA, USA), and Karolinska Institute (Stockholm, Sweden). Surgical specimens were totally embedded and step-sectioned at 3- to 5-mm intervals and evaluated by pathologists with genitourinary expertise at each institution. Secondary therapy was uncommonly administered in the absence of biochemical recurrence (BCR). Most of the initial diagnostic biopsies were performed and graded elsewhere. In four of the study institutions, these biopsies were regraded at their respective institution and are included in the analysis. BCR was defined based on any postoperative PSA value ≥0.2 ng/ml, except at the Karolinska, where it was defined as two consecutive postoperative PSA values ≥0.2 mg/ml. Preoperative needle biopsies grades on 16 176 of the patients from the four institutions were compared with post-RP BCR. Between 2005 and 2014, 5501 consecutive patients with localized PCa were treated with radiotherapy at the Cleveland Clinic and MSKCC with the biopsies reviewed at their respective institutions [5,6]. Biochemical progression was based on the nadir +2 Phoenix definition.

Kaplan-Meier curves were used to illustrate BCR after treatment. The log-rank test was used to assess univariable differences in BCR by Gleason score. Separate univariable and multivariable Cox proportional hazards models were built using four possible categorizations of Gleason scores. The pretreatment models were adjusted for the log of

pretreatment PSA and clinical stage (T1 vs T2 vs T3/4); post-treatment models were adjusted for log preoperative PSA, surgical margin status, and pathology stage (pT2 vs pT3a vs pT3b vs pT4). Harrell's C-index was used to assess discrimination of the alternative Gleason grading schemes models, corrected for optimism using 10-fold cross validation. The models we planned to assess were the currently popular three-group risk stratification (6 vs 7 vs 8–10), splitting grade 7 (6 vs 3 + 4 vs 4 + 3 vs 8–10), adding grade 9 (6 vs 7 vs 8 vs 9–10), and a five-grade system (6 vs 3 + 4 vs 4 + 3 vs 8 vs 9–10).

3. Results

Patient characteristics are shown by treatment in Table 1 and by institution in Supplementary Table 1. The median follow-ups without BCR for the RP and radiation cohorts were 3.0 and 3.1 yr, respectively. The number of men followed without BCR at 5 yr was 6008 and 1258 for the RP and radiation therapy cohorts, respectively. When we compared the four different scoring schemes, we found large differences between both Gleason 3 + 4 versus 4 + 3 and Gleason 8 versus 9 (Table 2 and 3). In the RP cohort, hazard ratios (HRs) for Gleason 4 + 3 disease were generally threefold higher than for 3 + 4; the HR for Gleason 9–10 was about twice as high than for 8. The 5-yr BCR-free progression probabilities for RP Gleason scores 6, 3 + 4, 4 + 3, 8, and 9–10 were 96% (95% confidence interval [CI], 95–96), 88% (95% CI, 85–89), 63% (95% CI, 61–65), 48% (95% CI, 44–52), and 26% (95% CI, 23–30), respectively (Fig. 1–4). Differences between 3 + 4 and 4 + 3 were smaller for the radiotherapy cohort. There is excellent separation between the five grade groups for surgery, whereas for radiotherapy there is an overlap between 4 + 3 and 4 + 4. This appears to be a result of the association between grade and hormone use. The proportion of patients undergoing adjuvant or neoadjuvant hormonal therapy among radiotherapy patients with Gleason 6, 3 + 4, 4 + 3, 8, and 9–10 disease was 15%, 26%, 45%, 84%, and 95%, respectively. Figure 4 shows results for radiotherapy patients treated without hormone therapy, demonstrating a clear separation between the five grade groups. Among the Gleason 8 scores, the vast majority were 4 + 4, so we did not consider it useful to separate these out any further into 3 + 5 or 5 + 3 based on the frequencies alone.

On univariate analysis, the C-index for a five grade group system was 0.02 to 0.05 higher than for the common three-group approach, and the highest of any approach (Table 3). Differences were smaller on multivariable analysis, but the five grade group system retained its advantage. The increment in the C-index by including 9–10 separately from 8 ranges from 0.001 to 0.003. This is largely because the prevalence of Gleason 9–10 disease is low, only about 5% of the cohort. Nonetheless, as can be seen in Table 2 and Figure 1 and 2, Gleason 9–10 has a markedly poorer prognosis than Gleason 8. Similarly, HRs for Gleason 4 + 3 disease were generally threefold higher than for 3 + 4 and are not accounted for in the three-group approach.

4. Discussion

The Gleason score has been the single most powerful predictor of PCa prognosis. Although Gleason scores range from 2 to 10, there are 25 potential scores (eg, 1 + 1, 1 + 2, 1 + 3, 1 + 4, 1 + 5, 2 + 1). In the literature and for therapeutic purposes, various scores have been

grouped together based on the assumption that they have a similar prognosis. Analyzing some of the highest impact recent articles on PCa reveals considerable diversity of Gleason score groupings: 2–4, 5–7, 8–10 (Prostate Cancer Outcomes Study) [7]; 2–6, 7, 8–10 (Scandinavian Prostate Cancer Group Study); and 2–6, 7–10 (Prostate Cancer Prevention Trial and Prostate Cancer Intervention versus Observation Trial) [8,9]. The most common risk stratification for PCa is the D'Amico classification, also used by the National Comprehensive Cancer Network [10]. It stratifies PCa based on serum PSA values, clinical stage, and biopsy score into low-, intermediate-, and high-risk groups incorporating Gleason scores into a three-tiered Gleason score grouping (2–6, 7, and 8–10).

In addition to the lack of uniformity of the various score groupings, precluding meaningful comparisons between studies, the combinations used have significant flaws. Gleason scores 2–4 virtually never exist on current biopsy material. Many of these cases in Gleason's era, predating the use of modern immunohistochemistry, probably represent adenosis, a mimic of cancer. Studies combining Gleason 6 and 7 span tumors with an almost uniformly excellent prognosis (3 + 3) to those with a substantial likelihood of progressing following therapy (4 + 3). All of the classification systems just described consider Gleason 7 as a single score without distinguishing 3 + 4 versus 4 + 3, despite studies, including the current one, showing a significantly worse prognosis for the latter [11,12]. Combining scores 7–10 includes cases with an excellent prognosis (3 + 4) along with those that have a high PCa-specific lethality (5 + 5). Even within the high score group of 8–10, Gleason scores 9–10 have a much poorer prognosis than Gleason score 8, an observation also supported by prior studies [13].

The current study distills grades of PCa down to the lowest number of grades, each with a unique prognosis. The grading system used in the current study has as its underpinning the Gleason grading, but it bears little resemblance to the original scoring. While retaining the practice of combining the two most common patterns, there have been many changes, first codified in 2005 and more recently at a 2014 conference of experts [2,3].

Two of the more prominent changes have been the disappearance of Gleason scores 2–5 from clinical practice and the more restrictive definition of Gleason score 6. In Gleason's original data, 28% of the cases were Gleason scores 2–5, and Gleason pattern 4 (which includes Gleason scores 7–8 and some of 9–10) was present in only 12% of the cancers [14]. In a study from Danneman et al, Gleason scores 2–5 decreased from 27% in 1998 to 1% in 2011, and Gleason score 7 accounted for 38% of cases in 2011 [15]. The increase in Gleason score 7 tumors in large part reflects that poorly formed glands and some cribriform glands were considered as Gleason pattern 3 in the original system, now upgraded to Gleason pattern 4 in the modified system [2]. In the original Gleason system, large cribriform glands that in current practice are pattern 4 were typically graded as Gleason pattern 3 (Fig. 5) [16,17]. Numerous studies have demonstrated the adverse prognosis of cribriform glands such that all cribriform patterns were accepted as Gleason pattern 4 at the 2014 grading conference [3,18].

As a result of significant differences in criteria and reporting compared with Gleason's original grading system, we have regarded the newly proposed grades as a "new grading system," although one could also consider it as a "novel grouping" of a much modified

original Gleason grading system. The histologic definitions of the five grade groups in the new grading system are listed in Table 4.

The major consequence of this shift in grading has been the better prognosis associated with Gleason score 6 cancer because patterns associated with more aggressive behavior have been shifted to Gleason score 7. Historically, a diagnosis of Gleason score 6 cancer was not as predictive of good behavior, with a higher rate of progression and some men even dying of PCa [17,19]. In the current study, Gleason score 6 cancer at RP has a 96% cure rate at 5 yr, even including cases with extraprostatic extension and positive margins. One large multi-institutional study using the 2005 revised Gleason system demonstrated that a pure Gleason 6 cancer at surgery has no potential for metastatic behavior [17].

Some have questioned whether Gleason score 6 should even be called cancer, proposing alternative terms such as *indolent lesion of epithelial origin* because of the fear associated with the term *cancer* [20]. Contributing to this fear is a perception of a more serious cancer implicit in a system assigning a Gleason score of 6 to a cancer out of a grading scale of 2–10, although 6 is the lowest score currently assigned. Although there are numerous morphologic, molecular, and clinical reasons why the term *cancer* should be retained for Gleason 6 tumors, Esserman's contention that "Changing the language we use to diagnose various lesions is essential to give patients confidence that they don't have to aggressively treat every finding in a scan" is sensible [21,22].

Rather than reclassifying current Gleason score 6 to a noncancerous term, a change to a new prostate cancer grade group 1 of 5 will help define the indolent nature of the cancer and reassure an initial strategy of active surveillance in the appropriate patient [23]. Given the issues associated with incomplete sampling with prostate biopsy, follow-up for men with grade group 1 (Gleason 3 + 3 = 6) cancer undergoing surveillance is still needed because approximately 20–30% of cases harbor unsampled higher grade cancer in the gland [24]. In addition to biopsy grade, deciding whether a patient is a candidate for surveillance is complex, and it factors in multiple clinical findings as well as the extent of cancer on biopsy.

With proper education, patients are not threatened when diagnosed with basal cell and squamous cell carcinomas of the skin; likewise, understanding the generally indolent behavior of grade group 1 PCa will permit more rational and less emotional decision making. Grade group 2 of 5 (as opposed to Gleason score 7 of 10) has a very good prognosis with rare metastases. Grade group 3 of 5 has a significantly worse prognosis than Grade 2 as opposed to Gleason score 7 that combines Gleason scores 3 + 4 and 4 + 3. Grade group 4 of 5 is not considered the highest grade (as opposed to Gleason scores 8–10) and has a significantly better prognosis than grade group 5 (Gleason scores 9–10). Finally, grade group 5 obviates the need to distinguish between Gleason scores 4 + 5, 5 + 4, and 5 + 5, just as grade group 1 makes irrelevant the distinction between Gleason scores 2 + 2, 2 + 3, 3 + 2, and 3 + 3.

The major limitation of the current study is the use of PSA BCR as an end point as opposed to cancer-related death. This limitation was unavoidable because the modified Gleason system was first introduced in 2005. It will take another 10–15 yr to have sufficient follow-

up to use metastasis or death as end points, and the need to change reporting of PCa grade is urgent. An alternative approach, retrospectively re-grading thousands of RP and needle biopsy specimens, was virtually impossible because most biopsy specimens have been returned to their original institutions.

Although PSA recurrence is an imperfect clinical end point, it is currently the strongest practical end point, driving almost all initial disease management decisions after primary treatment. Models developed to predict PCa-specific mortality in men with PSA recurrence show a strong association, particularly for those with a rapidly rising PSA [25,26]. The currently proposed grading system also has as its foundation the Gleason system, which has been correlated with death due to PCa [12,13,18].

We did not address tertiary grade patterns because the magnitude of its effect is not uniform across each grade, so a simple rule cannot be applied. We proposed in a prior study to incorporate the tertiary grade within the actual grade, and it was never accepted in practice [27]. Different institutions define and apply tertiary grades variably. A separate manuscript in the pathology literature will deal with how to handle tertiary grades. A minor pattern 5 component will be noted in a Gleason score 7 tumor at RP; in needle biopsy specimens, the grade is derived from adding the most common and highest grade patterns.

Strengths of the current study include multi-institutional and international data, the largest contemporary follow-up study after RP yet assembled, and similar methods for the processing of specimens. Although one institution used a slightly different definition of BCR, we have previously demonstrated that this would have minimal impact on the findings [28]. RP specimens were chosen as the primary source to correlate grade with outcome because the entire tumor was available for analysis, but the new grading system was also validated on biopsies. The current study also verifies the grading system in a large multi-institutional group of men treated with radiation. Data from many other institutions and cooperative groups with large radiation cohorts were not usable in the current study because Gleason score 7 was not separated into 3 + 4 and 4 + 3, and Gleason scores 8–10 were not subdivided. We noted an overlap between the Gleason scores 8 and 4 + 3 in the entire radiation cohort. The best explanation for this observation is that more Gleason score 8 patients received hormone therapy, a treatment that is effective at reducing recurrence risk. Note that progression rates between the surgical and radiation cohorts cannot be directly compared due to different definitions of progression [29,30].

5. Conclusions

The new grading system for PCa has obvious benefits: (1) more accurate grade stratification than the current Gleason system; (2) a simplified grading system of 5 as opposed to 25 scores, depending on various Gleason pattern combinations; and (3) a lowest grade of 1, as opposed to 6, with the potential to reduce overtreatment of indolent PCa. As a result of modifications in 2005 and 2014, the current application of the Gleason grading system deviates sufficiently from Gleason's original system to justify a new grading system. The new grading system is not only supported by the current study but also by other studies published after 2005 demonstrating the excellent prognosis of Gleason score 6, the worse

prognosis of 4 + 3 = 7 versus 3 + 4 = 7, and the poor prognosis of Gleason scores 9–10. To avoid confusion, it will be prudent to report the new grading system, in conjunction with the Gleason system, until it becomes widely accepted and practiced (eg, Gleason score 3 + 3 = 6 [grade group 1]).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding/Support and role of the sponsor: None

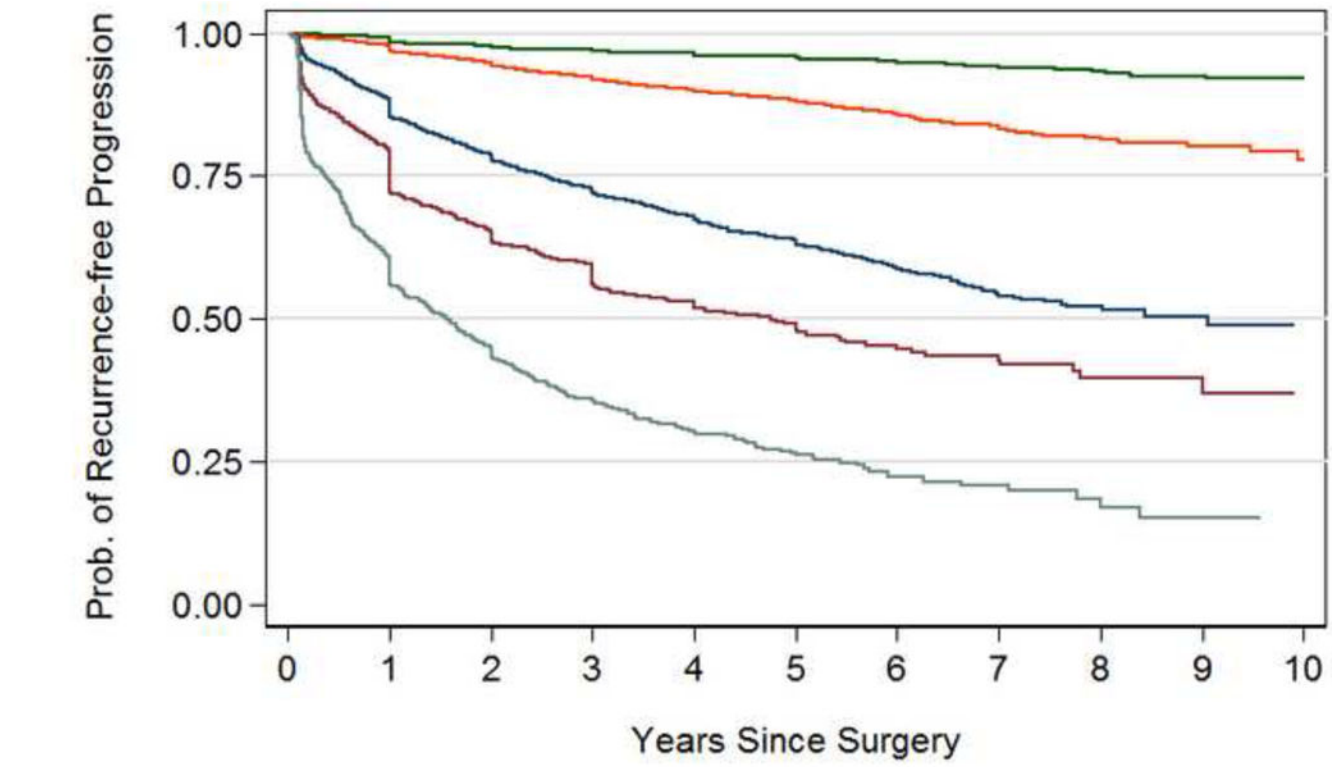
References

1. Gleason DF, Mellinger GT. Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging. J Urol. 1974; 111:58–64. [PubMed: 4813554]
2. Epstein JI, Allsbrook WC Jr, Amin MB, Egevad LL, ISUP Grading Committee. The 2005 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma. Am J Surg Pathol. 2005; 29:1228–42. [PubMed: 16096414]
3. Epstein JI, Egevad L, Amin MB, Delahunt B, Srigley JR, Humphrey PA. The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma: definition of grading patterns and proposal for a new grading system. Am J Surg Pathol. In press.
4. Pierorazio PM, Walsh PC, Partin AW, Epstein JI. Prognostic Gleason grade grouping: data based on the modified Gleason scoring system. BJU Int. 2013; 111:753–60. [PubMed: 23464824]
5. Spratt DE, Zumsteg ZS, Ghadjar P, et al. Comparison of high-dose (86.4 Gy) IMRT vs combined brachytherapy plus IMRT for intermediate-risk prostate cancer. BJU Int. 2014; 114:360–7. [PubMed: 24447404]
6. Videtic, GM.; Woody, NM. Handbook of treatment planning in radiation oncology. 2. New York, NY: Demos Medical; 2014.
7. Resnick MJ, Koyama T, Fan KH, et al. Long-term functional outcomes after treatment for localized prostate cancer. N Engl J Med. 2013; 368:436–45. [PubMed: 23363497]
8. Wilt TJ, Brawer MK, Jones KM, et al. Radical prostatectomy versus observation for localized prostate cancer. N Engl J Med. 2012; 367:203–13. [PubMed: 22808955]
9. Bill-Axelson A, Holmberg L, Garmo H, et al. Radical prostatectomy or watchful waiting in early prostate cancer. N Engl J Med. 2014; 370:932–42. [PubMed: 24597866]
10. D'Amico AV, Whittington R, Malkowicz SB, et al. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. JAMA. 1998; 280:969–74. [PubMed: 9749478]
11. Chan TY, Partin AW, Walsh PC, Epstein JI. Prognostic significance of Gleason score 3+4 versus Gleason score 4+3 tumor at radical prostatectomy. Urology. 2000; 56:823–7. [PubMed: 11068310]
12. Stark JR, Perner S, Stampfer MJ, et al. Gleason score and lethal prostate cancer: does 3 + 4 = 4 + 3? J Clin Oncol. 2009; 27:3459–64. [PubMed: 19433685]
13. Tsao CK, Gray KP, Nakabayashi M, et al. Patients with biopsy Gleason 9 and 10 prostate cancer have significantly worse outcomes compared with Gleason 8 disease. J Urol. 2015; 194:91–7. [PubMed: 25623747]
14. Mellinger GT. Prognosis of prostatic carcinoma. Recent Results Cancer Res. 1977; 60:61–72. [PubMed: 866797]
15. Danneman D, Drevin L, Robinson D, Stattin P, Egevad L. Gleason inflation 1998–2011: a registry study of 97 168 men. BJU Int. 2015; 115:248–55. [PubMed: 24552193]

16. McNeal JE, Yemoto CE. Spread of adenocarcinoma within prostatic ducts and acini. Morphologic and clinical correlations. Am J Surg Pathol. 1996; 20:802–14. [PubMed: 8669528]
17. Ross HM, Kryvenko ON, Cowan JE, Simko JP, Wheeler TM, Epstein JI. Do adenocarcinomas of the prostate with Gleason score (GS) </=6 have the potential to metastasize to lymph nodes? Am J Surg Pathol. 2012; 36:1346–52. [PubMed: 22531173]
18. Kweldam CF, Wildhagen MF, Steyerberg EW, Bangma CH, van der Kwast TH, van Leenders GJ. Cribriform growth is highly predictive for postoperative metastasis and disease-specific death in Gleason score 7 prostate cancer. Mod Pathol. 2015; 28:457–64. [PubMed: 25189638]
19. Nakabayashi M, Hayes J, Taplin ME, et al. Clinical predictors of survival in men with castration-resistant prostate cancer: evidence that Gleason score 6 cancer can evolve to lethal disease. Cancer. 2013; 119:2990–8. [PubMed: 23719969]
20. Esserman LJ, Thompson IM, Reid B, et al. Addressing overdiagnosis and overtreatment in cancer: a prescription for change. Lancet Oncol. 2014; 15:e234–42. [PubMed: 24807866]
21. Berman DM, Epstein JI. When is prostate cancer really cancer? Urol Clin North Am. 2014; 41:339–46. [PubMed: 24725494]
22. Carter HB, Partin AW, Walsh PC, et al. Gleason score 6 adenocarcinoma: should it be labeled as cancer? J Clin Oncol. 2012; 30:4294–6. [PubMed: 23032616]
23. Nelson JB. Observation for clinically localized prostate cancer. J Clin Oncol. 2014; 32:1295–8. [PubMed: 24637994]
24. Epstein JI, Feng Z, Trock BJ, Pierorazio PM. Upgrading and downgrading of prostate cancer from biopsy to radical prostatectomy: incidence and predictive factors using the modified Gleason grading system and factoring in tertiary grades. Eur Urol. 2012; 61:1019–24. [PubMed: 22336380]
25. Brockman JA, Alanee S, Vickers AJ, et al. Nomogram predicting prostate cancer-specific mortality for men with biochemical recurrence after radical prostatectomy. Eur Urol. 2015; 67:1160–7. [PubMed: 25301759]
26. Freedland SJ, Humphreys EB, Mangold LA, et al. Risk of prostate cancer-specific mortality following biochemical recurrence after radical prostatectomy. JAMA. 2005; 294:433–9. [PubMed: 16046649]
27. Trock BJ, Guo CC, Gonzalgo ML, Magheli A, Loeb S, Epstein JI. Tertiary Gleason patterns and biochemical recurrence after prostatectomy: proposal for a modified Gleason scoring system. J Urol. 2009; 182:1364–70. [PubMed: 19683280]
28. Cronin AM, Godoy G, Vickers AJ. Definition of biochemical recurrence after radical prostatectomy does not substantially impact prognostic factor estimates. J Urol. 2010; 183:984–9. [PubMed: 20083281]
29. Nielsen ME, Makarov DV, Humphreys E, Mangold L, Partin AW, Walsh PC. Is it possible to compare PSA recurrence-free survival after surgery and radiotherapy using revised ASTRO criterion–“nadir + 2”? Urology. 2008; 72:389–93. [PubMed: 18279937]
30. Lee BH, Kibel AS, Ciezki JP, et al. Are biochemical recurrence outcomes similar after radical prostatectomy and radiation therapy? Analysis of prostate cancer-specific mortality by nomogram-predicted risks of biochemical recurrence. Eur Urol. 2015; 67:204–9. [PubMed: 25294696]

Take-home message

A new prostate cancer (PCa) grading system was validated with more accurate grade stratification, including a simplified grading system of five grades and a lowest grade of 1, as opposed to 6, with the potential to reduce overtreatment of PCa.



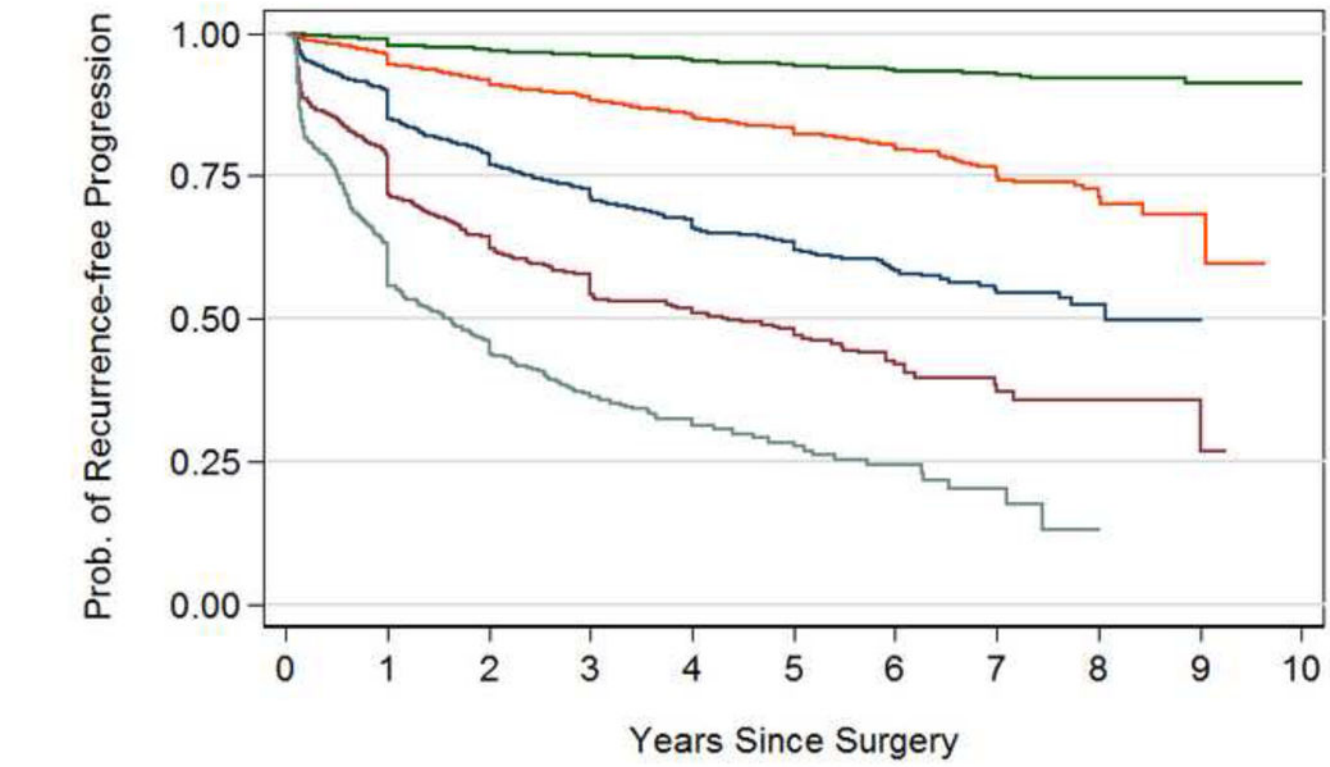
Number at risk

6	7397	6973	5104	4064	3226	2461	1768	1186	670	278	108
3+4	8353	7202	5298	3983	2955	2091	1299	778	393	135	45
4+3	3106	2452	1605	1152	839	569	350	199	90	38	15
8	917	678	412	280	191	129	86	59	35	14	7
>=9	1051	578	325	194	118	73	41	24	12	4	2

Fig. 1.

Recurrence-free progression following radical prostatectomy stratified by prostatectomy grade. Green line: Gleason score 6, grade group 1. Orange line: Gleason score 3 + 4, grade group 2. Dark blue line: Gleason score 4 + 3, grade group 3. Brown line: Gleason score 8, grade group 4. Gray line: Gleason score ≥9, grade group 5.

RFP = recurrence-free progression.



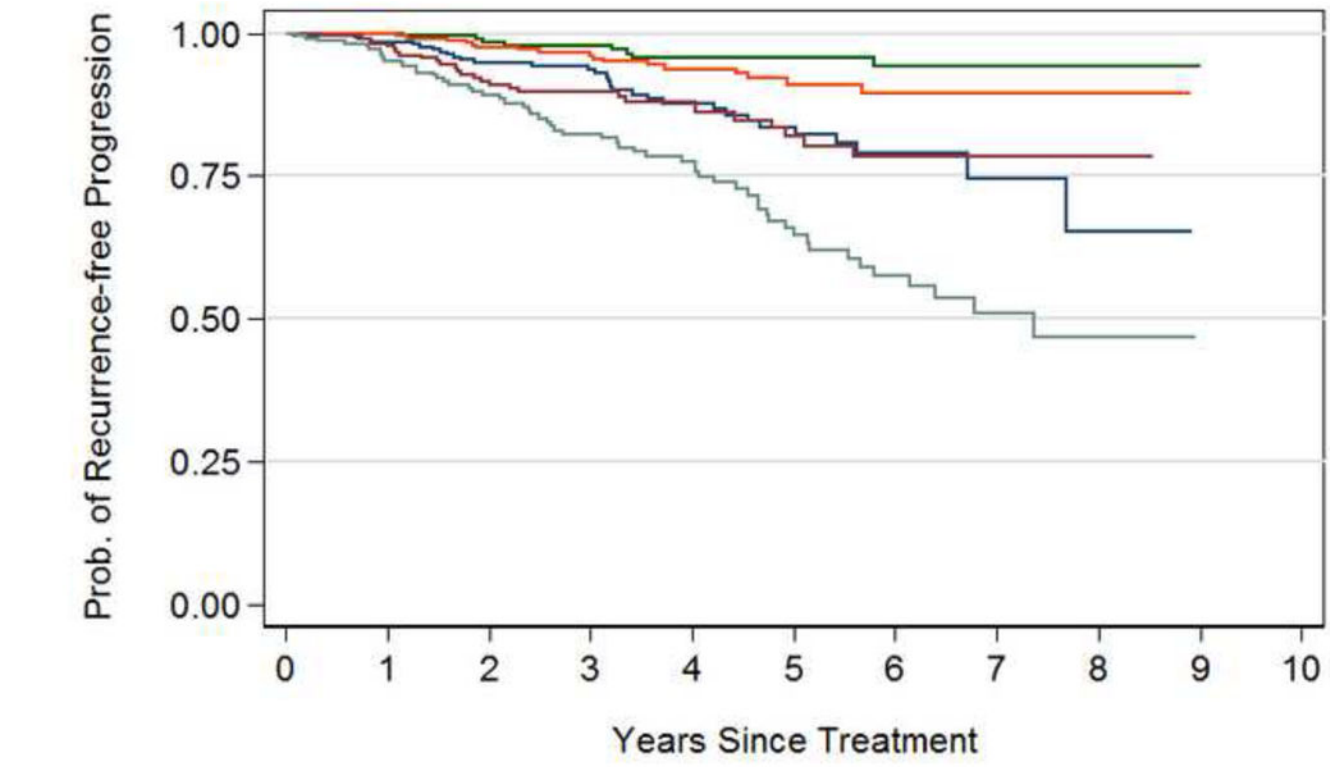
Number at risk

6	8039	7264	5154	3943	3018	2177	1383	818	371	80	1
3+4	4595	3875	2624	1845	1291	845	470	244	102	21	0
4+3	1872	1511	934	634	432	282	157	90	35	5	0
8	1005	710	413	279	185	120	68	33	13	4	0
>=9	661	365	199	118	75	41	25	11	2	0	0

Fig. 2.

Recurrence-free progression following radical prostatectomy stratified by pre-prostatectomy biopsy grade. Green line: Gleason score 6, grade group 1. Orange line: Gleason score 3 + 4, grade group 2. Dark blue line: Gleason score 4 + 3, grade group 3. Brown line: Gleason score 8, grade group 4. Gray line: Gleason score ≥9, grade group 5.

RFP = recurrence-free progression.



Number at risk

6	237	211	183	147	113	85	58	33	14	0	0
3+4	355	311	247	183	132	78	54	29	13	0	0
4+3	279	248	183	142	97	68	38	13	3	0	0
8	235	198	147	120	89	53	33	10	3	0	0
>=9	228	194	143	111	86	54	32	17	6	0	0

Fig. 3.

Recurrence-free progression following radiation stratified by pre–radiation therapy biopsy grade (entire cohort). Green line: Gleason score 6, grade group 1. Orange line: Gleason score 3 + 4, grade group 2. Dark blue line: Gleason score 4 + 3, grade group 3. Brown line: Gleason score 8, grade group 4. Gray line: Gleason score ≥9, grade group 5.
RFP = recurrence-free progression.

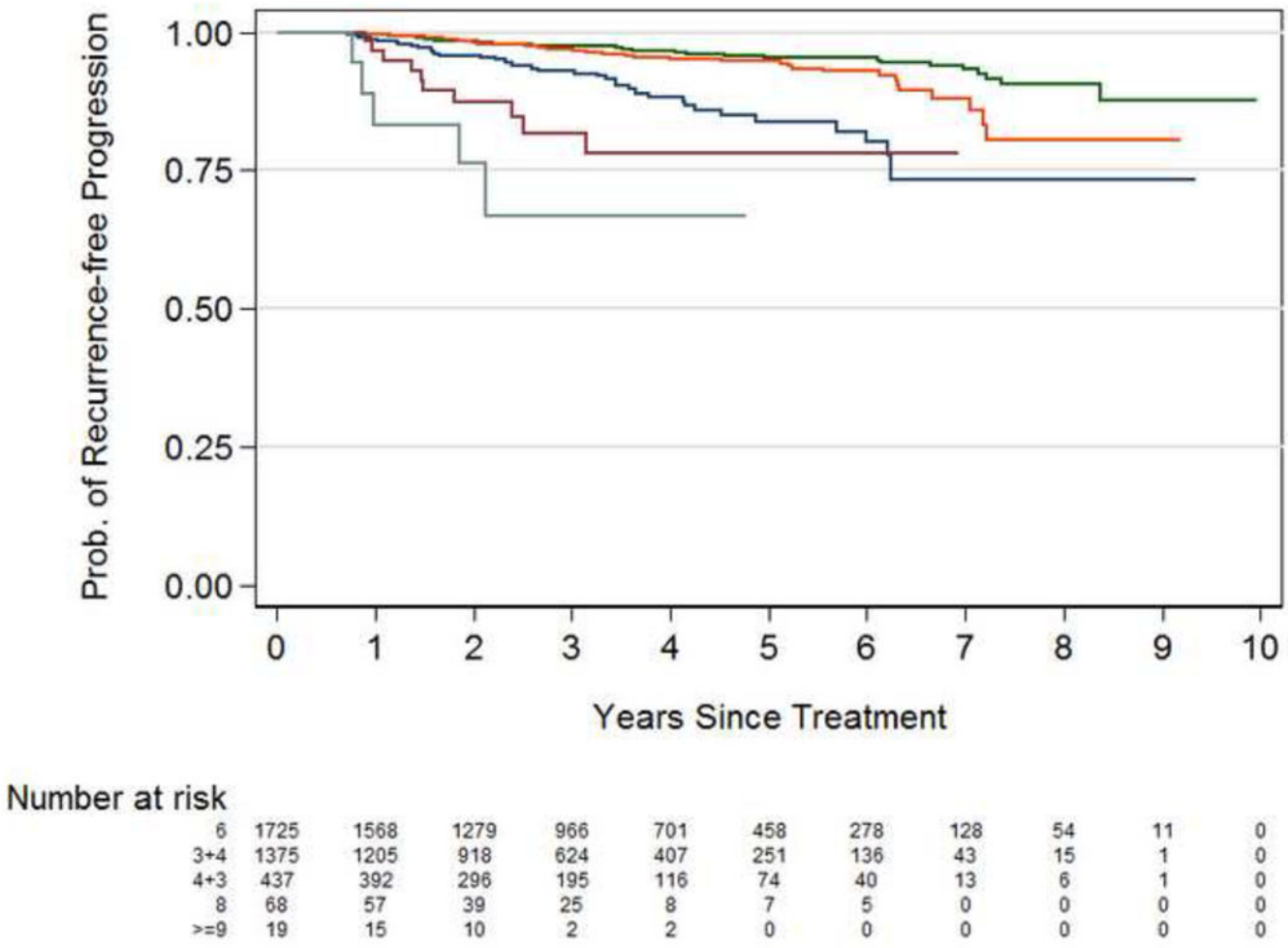


Fig. 4.

Recurrence-free progression following radiation stratified by pre–radiation therapy biopsy grade (no hormone therapy cohort). Green line: Gleason score 6, grade group 1. Orange line: Gleason score 3 + 4, grade group 2. Dark blue line: Gleason score 4 + 3, grade group 3. Brown line: Gleason score 8, grade group 4. Gray line: Gleason score ≥9, grade group 5. RFP = recurrence-free progression.

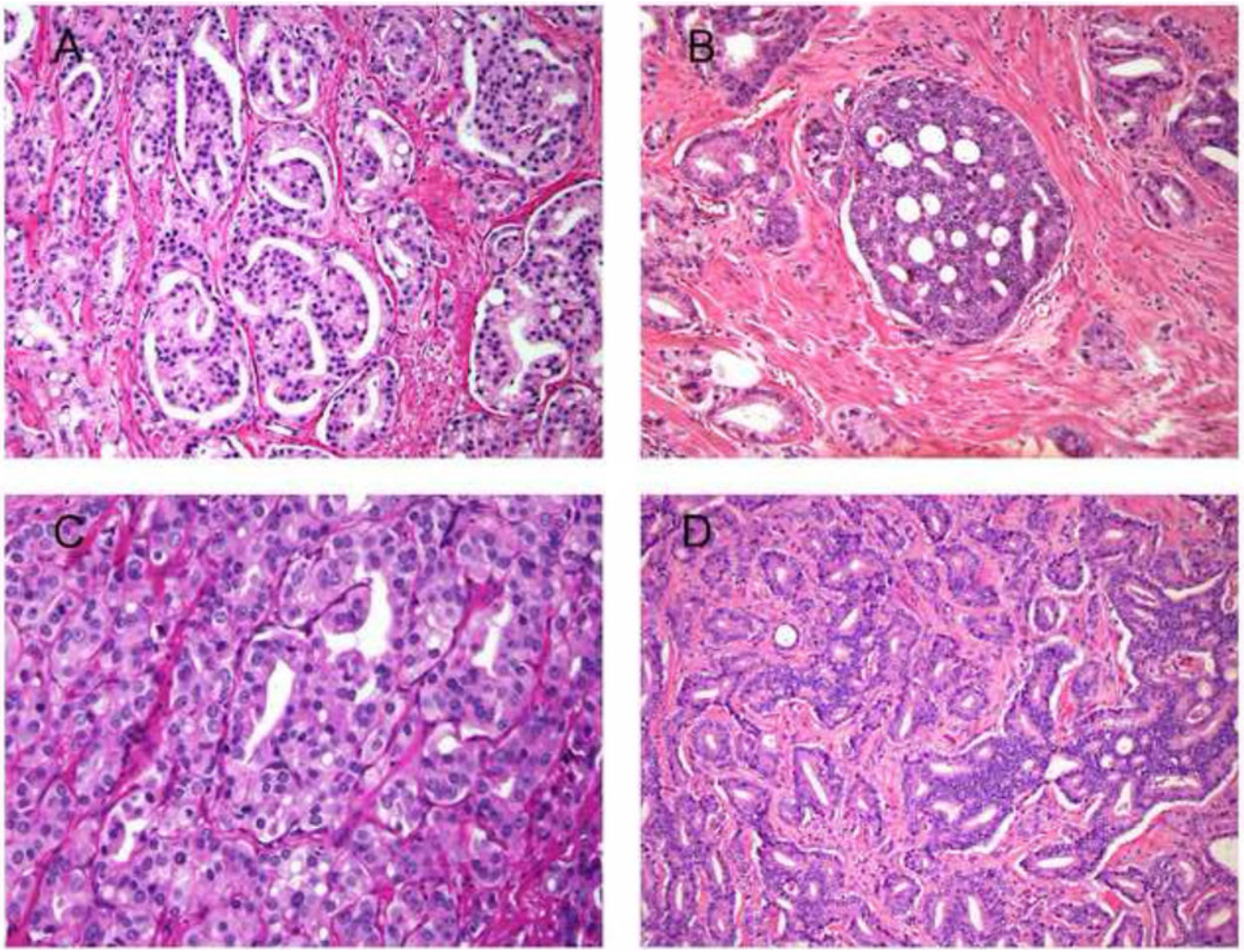


Fig. 5.

Images of cases diagnosed as Gleason score 6 prior to 2005 from Johns Hopkins Hospital, Henry Ford Hospital, University of California San Francisco, or Baylor College of Medicine (modified from Ross et al [17]). (a) Gleason pattern 4 with glomeruloid glands. (b) Gleason pattern 4 with medium-sized rounded cribriform gland. Associated small glands of pattern 3. (c) Gleason pattern 4 with poorly formed and fused glands. (d) Gleason pattern 4 with irregular cribriform gland with adjacent Gleason pattern 3.

Table 1

Patient characteristics

	Before radical prostatectomy (*n* = 16 172)	After radical prostatectomy (*n* = 20 824)	RT (*n* = 5501)
Age at treatment, yr	60 (55–65)	61 (56–65)	NA
Pretreatment PSA, ng/ml	5.0 (3.8–6.9)	5.3 (4.0–7.5)	6.1 (4.6–9.0)
Clinical stage			
T1	10 774 (67)	13 518 (65)	3825 (70)
T2	4066 (25)	5607 (27)	1405 (26)
T3/4	306 (1.9)	461 (2.2)	233 (4.2)
Unknown	1026 (6.3)	1238 (5.9)	38 (0.7)
Gleason			
≤6	8039 (50)	7397 (36)	2029 (37)
3 + 4	4595 (28)	8353 (40)	1883 (34)
4 + 3	1872 (12)	3106 (15)	805 (15)
8	1005 (6.2)	917 (4.4)	431 (7.8)
≥9	661 (4.1)	1051 (5.0)	353 (6.4)
Radical prostatectomy			
Positive surgical margin	NA	3808 (18)	NA
Unknown		116 (0.6)	
Pathology stage	NA		NA
pT0		6 (<0.1)	
pT2		13 806 (66)	
pT3a		5258 (25)	
pT3b		1307 (6.3)	
pT4		147 (0.7)	
Unknown		300 (1.4)	
RT			
Type	NA	NA	
Brachytherapy			3361 (61)
EBRT			2140 (39)
Peri-RT treatment hormone therapy	NA	NA	1845 (34)
Unknown			32 (0.6)

EBRT = external-beam radiation therapy; NA = not applicable; PSA = prostate-specific antigen; RT = radiation therapy.

Statistics presented are median (interquartile range) or number (percentage).

Table 2

Univariate and multivariable results of Cox proportional hazards regression using varying Gleason grade categorizations

	RP biopsy Gleason grade						Post-RP Gleason grade						RT Gleason grade without hormone therapy					
	Univariate			Multivariable			Univariate			Multivariable			Univariate			Multivariable		
	HR	95% CI	*p*	HR	95% CI	*p*	HR	95% CI	*P*	HR	95% CI	*p*	HR	95% CI	*p*	HR	95% CI	*p*
≤6	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
3 + 4	3.23	2.82–3.71	<0.0001	2.54	2.18–2.95	<0.0001	2.66	2.32–3.06	<0.0001	1.94	1.67–2.24	<0.0001	1.47	1.08–2.00	0.014	1.32	0.97–1.81	0.076
4 + 3	8.14	7.08–9.36	<0.0001	5.70	4.88–6.67	<0.0001	9.94	8.67–11.40	<0.0001	5.14	4.43–5.97	<0.0001	3.65	2.69–4.95	<0.0001	2.83	2.06–3.88	<0.0001
8	14.56	12.59–16.84	<0.0001	9.14	7.73–10.80	<0.0001	16.76	14.33–19.59	<0.0001	7.99	6.73–9.48	<0.0001	4.26	3.03–6.00	<0.0001	2.87	2.00–4.12	<0.0001
≥9	26.07	22.47–30.24	<0.0001	13.78	11.53–16.47	<0.0001	33.16	28.73–38.28	<0.0001	11.68	9.92–13.76	<0.0001	7.58	5.58–10.30	<0.0001	4.47	3.17–6.31	<0.0001
≤6	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
7	4.54	4.01–5.14	<0.0001	3.38	2.95–3.89	<0.0001	4.41	3.87–5.02	<0.0001	2.73	2.38–3.13	<0.0001	2.11	1.62–2.76	<0.0001	1.77	1.35–2.33	<0.0001
≥8	18.36	16.14–20.89	<0.0001	10.25	8.81–11.94	<0.0001	24.06	21.02–27.53	<0.0001	8.50	7.31–9.90	<0.0001	5.78	4.39–7.63	<0.0001	3.43	2.52–4.67	<0.0001

CI = confidence interval; HR = hazard ratio; Ref. = reference; RP = radical prostatectomy; RT = radiation therapy.

Multivariable biopsy Gleason Cox model includes preoperative prostate-specific antigen (PSA) and clinical stage (T1 vs T2 vs T3/4), and post-RP Cox model includes preoperative PSA, surgical margin status, and pathology stage (pT2 vs pT3a vs pT3b vs pT4).

Table 3

Discrimination of varying Gleason grade categorizations

	RP biopsy Gleason grade		Post-RP Gleason grade		RT Gleason grade	
	Univariate	Multivariable	Univariate	Multivariable	Univariate	Multivariable
≤6 vs 7 vs ≥8	0.760	0.805	0.744	0.83	0.662	0.729
≤6 vs 3 + 4 vs 4 + 3 vs ≥8	0.781	0.811	0.791	0.842	0.684	0.736
≤6 vs 7 vs 8 vs ≥9	0.762	0.806	0.747	0.831	0.666	0.729
≤6 vs 3 + 4 vs 4 + 3 vs 8 vs ≥9	0.783	0.813	0.793	0.842	0.687	0.737

RP = radical prostatectomy; RT = radiation therapy.

Multivariable biopsy Gleason Cox model includes preoperative prostate-specific antigen (PSA) and clinical stage (T1 vs T2 vs T3/4), and post-RP Cox model includes preoperative PSA, surgical margin status, and pathology stage (pT2 vs pT3a vs pT3b vs pT4). The C-index has been corrected for optimism using 10-fold cross-validation.

Table 4

Histologic definition of new grading system

Grade group 1 (Gleason score 3 + 3 = 6): Only individual discrete well-formed glands
Grade group 2 (Gleason score 3 + 4 = 7): Predominantly well-formed glands with lesser component of poorly formed/fused/cribriform glands
Grade group 3 (Gleason score 4 + 3 = 7): Predominantly poorly formed/fused/cribriform glands with lesser component of well-formed glands †
Grade group 4 (Gleason score 8)
- Only poorly formed/fused/cribriform glands *or*
- Predominantly well-formed glands and lesser component lacking glands ††
- Predominantly lacking glands and lesser component of well-formed glands ††
Grade group 5 (Gleason scores 9–10): Lack of gland formation (or with necrosis) with or without poorly formed/fused/cribriform glands †

† For cases with >95% poorly formed/fused/cribriform glands or lack of glands on a core or at radical prostatectomy, the component of <5% well-formed glands is not factored into the grade.

†† Poorly formed/fused/cribriform glands can be a more minor component.