

Clinical impact of prostate biopsy undergrading in an academic and community setting

Ashkan Mortezaei¹ · Etienne Xavier Keller¹ · Cédric Poyet¹ · Thomas Hermanns¹ · Karim Saba¹ · Marco Randazzo¹ · Christian Daniel Fankhauser¹ · Peter J. Wild² · Holger Moch² · Tullio Sulser¹ · Daniel Eberli¹

Received: 16 December 2015 / Accepted: 8 February 2016 / Published online: 1 March 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract

Purpose To evaluate whether the rate of Gleason score (GS) upgrade on final pathology, the rate of positive surgical margins (PSM) and the rate of biochemical recurrence (BCR) after radical prostatectomy (RP) were different if prostate biopsy (PB) was graded by community pathologists (CP) as compared to specialized uro-pathologists (UP).

Methods A consecutive series of patients undergoing RP in our institution between 2005 and 2013 were retrospectively reviewed. Any GS higher or lower in RP specimen as compared to PB GS was defined as GS upgrade or downgrade, respectively. Additionally, stratification for the new ISUP 2014 grading system was performed. Predictors of GS upgrade and PSMs and prognostic parameters for BCR were assessed by stepwise logistic regression models and by multivariable Cox regression analyses, respectively.

Results A total of 786 patients were available for analysis, and median follow-up was 36 months (1–101 months). A GS upgrade was found in 345 patients (43.9 %) and a GS downgrade in 91 patients (11.6 %). Discordance between PB GS and RP GS was significantly more frequent when grading had been performed by a CP (50.5 %

upgrade, 9.0 % downgrade) than by a UP (33.1 % upgrade, 15.7 % downgrade, $p < 0.001$). CP evaluation was an independent predictor for GS upgrade (odds ratio [OR] 1.91, $p < 0.001$) and for PSMs (OR 1.69, $p = 0.003$), as well as an independent predictor of BCR (hazard ratio [HR] 1.65, $p = 0.028$).

Conclusions Pathologic evaluation of PBs by a dedicated UP should be recommended to reduce the rate of biopsy undergrading, PSM and BCR after RP.

Keywords Prostate cancer · Transrectal biopsy · Gleason score · Undergrading · Prognosis · Positive surgical margin

Introduction

Since its first description in the 1960s, the Gleason grading system has been accepted as a grading standard and as the most accurate histopathological factor for the prognosis of prostate cancer (PC) patients [1]. Grading of prostate biopsy (PB) specimens is an important factor for counseling and decision making in men diagnosed with PC. The choice to perform active surveillance, nerve-sparing radical prostatectomy (RP), a pelvic lymphadenectomy or androgen deprivation therapy in addition to external beam radiation is all based on preoperative risk parameters such as PB Gleason score (GS) and pre-treatment prostate-specific antigen values (PSA) [2–4]. Thus, incorrect grading of PB specimens can result in inappropriate management of patients. Undergrading can lead to treatment delays or undertreatment (positive surgical margins, no lymph node resection) requiring secondary therapeutic options and compromising quality of life and survival.

Several studies have shown that discordance between PB GS and RP GS appears in almost 50 % of all cases [5].

Ashkan Mortezaei and Etienne Xavier Keller have contributed equally to this work.

✉ Ashkan Mortezaei
Ashkan.Mortezaei@usz.ch

✉ Daniel Eberli
Daniel.Eberli@usz.ch

¹ Department of Urology, University Hospital Zurich, University of Zurich, Frauenklinikstr.10, 8091 Zurich, Switzerland

² Institute of Surgical Pathology, University Hospital Zurich, University of Zurich, Zurich, Switzerland

Histopathological Gleason grade evaluation can significantly affect the accuracy of tumor classification [6]. Different levels of experience and skills of pathologists have been reported to be associated with GS discordance between PB and RP specimens [7–10]. However, the clinical significance of incorrect PB GS grading remains to be elucidated. In this study, we aimed to evaluate the accuracy of the PB GS grading depending on the diagnosing pathologist (community vs. uro-pathologist) and the prognostic impact on the oncological outcome in a contemporary RP series.

Materials and methods

Patient selection and data collection

All men who underwent robotic-assisted RP (RARP) between May 2005 and December 2013 in our tertiary-care academic center were retrospectively identified. Patients diagnosed with PC after ultrasound-guided transrectal PB performed in our center or externally by community urologists were eligible for this study. Men with PC diagnosed in specimens of transurethral resection of the prostate (TURP), or diagnosed after magnet resonance imaging (MRI)-guided biopsy and men who received neoadjuvant androgen deprivation treatment were excluded from further analysis. Electronic hospital charts were reviewed to collect peri- and postoperative data. Additionally, data were retrieved from referring urologist or patient's general practitioners if follow-up was not performed in our center.

Patients were divided into two groups, depending on whether their PB specimen had been evaluated in our institution by a minimum of two pathologists (uro-pathologist [UP] group)—of whom at least one expert in urologic pathology—or by pathologists in the community (community pathologist [CP] group). PB GS, core numbers and numbers of positive cores were retrieved from the respective pathology reports. PB slides of patients who were referred to our center by community urologists were not reviewed before RARP at our institution.

Surgical procedure and pathological analysis

RARP was performed by five experienced prostate surgeons using the three- and four-arm daVinci® Surgical System (Intuitive Surgical Inc., Sunnyvale, CA). Bilateral extended pelvic lymph node dissection (EPLND) was performed as described earlier in patients with either a PSA level of ≥ 10 ng/ml or a preoperative GS of ≥ 7 [11]. A nerve-sparing procedure was usually performed bilaterally in patients with cT1/2 PC and PB GS ≤ 7 and unilaterally in selected patients with GS 8 and small tumor volume identified on PB on the contralateral side.

All RARP specimens were processed at our institution. Comprehensive pathologic analysis was performed using standardized whole-mount sections [12].

Statistical analyses

SPSS version 22.0 (IBM Corp., Armonk, NY, USA) was used for statistical analyses. All two-sided p values ≤ 0.05 were considered statistically significant. The CP and the UP group were compared using Fisher's exact test for categorical variables and the Mann–Whitney U test for continuous variables.

PB GS was compared with the GS of their respective RP specimen individually. A RP GS higher or lower than the PB GS was defined as GS upgrade or downgrade, respectively. Concordance between PB GS and GS of the surgical specimen was calculated using the Cohen's kappa test (correction for agreement expected by chance) [13]. Additionally, stratification according to the recent ISUP 2014 prognostic grade groups (PGG) revision was performed [1, 14, 15].

Subgroup analyses evaluating clinically significant up- and downgrading for the following three treatment groups were performed: GS 5–6 = potential candidate for active surveillance [16], GS 7 = eligible for nerve-sparing [2] and GS 8–10 = high-risk tumors who should have undergone RP and EPLND without NS [2, 16] according to our institute's guidelines.

Logistic regression models adjusting for preoperative parameters were built to identify predictive factors for GS upgrades.

BCR was defined as PSA value ≥ 0.1 ng/ml with subsequent confirmation after reaching a PSA nadir ≤ 0.1 ng/ml postoperatively. The predictive impact of pre- and postoperative parameters on PSM and BCR rates was assessed by a stepwise logistic regression and a Cox regression model, respectively.

To assess differences in prognostic accuracy of the CP versus UP PB grading, Kaplan–Meier analyses of BCR-free survival (BCRFS) were performed and estimates were compared between the CP and UP group using the log-rank test. Subgroups were formed according to the D'Amico criteria [17].

Results

Of 826 patients undergoing RARP, 40 were excluded due to PC diagnosis by TURP ($n = 22$) or MRI-guided biopsy ($n = 11$) or due to neoadjuvant androgen deprivation therapy ($n = 7$) resulting in 786 patients eligible for the final analysis. Pre-, intra- and postoperative data of these patients are summarized in Table 1.

Table 1 Patient characteristics

	Total cohort (<i>n</i> = 786)	Community pathology (<i>n</i> = 487)	Uro-pathology (<i>n</i> = 299)	<i>p</i> *
<i>Preoperative data</i>				
Age, year				
Median (IQR)	64 (58–68)	63 (58–67)	64 (59–68)	0.008
Preoperative PSA, ng/ml				
Median (IQR)	7.1 (4.9–10.9)	7.5 (5.3–11.1)	6.8 (4.4–10.0)	<0.001
Clinical T stage				
cT1	528 (72.0 %)	317 (68.6 %)	211 (77.9 %)	0.007
cT2	205 (28.0 %)	145 (31.4 %)	60 (22.1 %)	
Number of PB cores				
Median (IQR)	10 (8–12)	10 (8–12)	12 (8–12)	<0.001
Number of positive PB cores				
Median (IQR)	3 (2–5)	3 (2–5)	3 (1–5)	0.430
PB Gleason score				
5–6	317 (40.3 %)	204 (41.9 %)	113 (37.8 %)	0.248
7a (3 + 4)	260 (33.1 %)	152 (31.2 %)	108 (36.1 %)	
7b (4 + 3)	104 (13.2 %)	73 (15.0 %)	31 (10.4 %)	
8–10	105 (13.4 %)	58 (11.9 %)	47 (15.7 %)	
<i>Intraoperative data</i>				
Lymph node dissection				
All	522/786 (66.4 %)	326/487 (66.9 %)	196/299 (65.6 %)	0.698
Gleason 5–6	32/149 (21.5 %)	18/81 (22.2 %)	14/68 (20.6 %)	0.844
Gleason 7	355/492 (72.2 %)	218/307 (71.0 %)	137/185 (74.1 %)	0.533
Gleason 8–10	135/145 (93.1 %)	90/99 (90.9 %)	45/46 (97.8 %)	0.170
Nerve-sparing				
All	510/785 (65 %)	331/487 (68.0 %)	179/298 (60.1 %)	0.026
Unilateral	268/785 (34.1 %)	162/487 (33.3 %)	106/298 (35.6 %)	
Bilateral	242/785 (30.8 %)	169/487 (34.7 %)	73/298 (24.5 %)	
PB Gleason 5–6	247/316 (78.2 %)	164/204 (80.4 %)	83/112 (74.1 %)	0.203
RP Gleason 5–6	124/148 (83.8 %)	70/81 (86.4 %)	54/67 (80.6 %)	0.376
PB Gleason 7	238/364 (65.4 %)	153/225 (68.0 %)	85/139 (61.2 %)	0.212
RP Gleason 7	328/492 (66.7 %)	218/307 (71.0 %)	110/175 (59.5 %)	0.010
PB Gleason 8–10	25/105 (23.8 %)	14/58 (24.1 %)	11/47 (23.4 %)	1.000
RP Gleason 8–10	58/145 (40 %)	43/99 (43.4 %)	15/46 (32.6 %)	0.275
<i>Postoperative data</i>				
pT stage				
pT2a–c	578 (73.5 %)	345 (70.5 %)	233 (78.0 %)	0.029
pT3ab	206 (26.5 %)	141 (28.9 %)	65 (22.7 %)	
pT4	2 (0.3 %)	1 (0.2 %)	3 (0.3 %)	
Gleason Score				
5–6	149 (19.0 %)	81 (16.6 %)	68 (22.7 %)	0.026
7a (3 + 4)	331 (42.1 %)	196 (40.2 %)	135 (45.2 %)	
7b (4 + 3)	161 (20.5 %)	111 (22.8 %)	50 (16.7 %)	
8–10	145 (18.4 %)	99 (20.3 %)	46 (15.4 %)	
Positive margin status				
All	237 (30.2 %)	169 (34.7 %)	68 (22.7 %)	<0.001
pT2	129 (22.3 %)	88 (25.5 %)	41 (17.6 %)	0.025
pT3	108 (51.9 %)	81 (57 %)	27 (40.9 %)	0.037
Gleason 5–6	22/149 (14.8 %)	16/81 (19.8 %)	6/68 (8.8 %)	0.068

Table 1 continued

	Total cohort (n = 786)	Community pathology (n = 487)	Uro-pathology (n = 299)	p*
Gleason 7	159/492 (32.3 %)	104/307 (33.9 %)	55/185 (29.7 %)	0.371
Gleason 8–10	56/145 (38.6 %)	49/99 (49.5 %)	7/46 (15.2 %)	<0.001
Months follow-up, m				
Median (IQR)	36 (19–59)	37 (17–59)	36 (22–58)	0.842

IQR interquartile range, SD standard deviation, PB prostate biopsy, RP radical prostatectomy

* p values ≤ 0.05 are marked in bold

Table 2 Gleason score up- and downgrades stratified by origin of pathology report

	Total cohort (n = 786)	Community pathology (n = 487)	Uro-pathology (n = 299)	p*
<i>Gleason Score shift</i>				
Upgrade	345/786 (43.9 %)	246/487 (50.5 %)	99/299 (33.1 %)	<0.001
Downgrade	91/786 (11.6 %)	44/487 (9.0 %)	47/299 (15.7 %)	0.006
Concordance (GS)	0.324	0.273	0.411	<0.001^a
Concordance (GS 5–6, 7, 8–10)	0.396	0.355	0.467	<0.001^a
<i>Clinically significant Gleason Score shift for potential active surveillance candidates (GS 5–6 vs. 7–10)</i>				
Upgrade of PB GS 5–6	186/317 (58.7 %)	131/204 (64.2 %)	55/113 (48.7 %)	0.009
Downgrade of PB GS 7–10	18/469 (3.8 %)	8/283 (2.8 %)	10/186 (5.4 %)	0.218
<i>Clinically significant Gleason Score shift for nerve-sparing (GS 5–7 vs. 8–10)</i>				
Upgrade of				
PB GS 5–7	71/681 (10.4 %)	54/429 (12.6 %)	17/252 (6.7 %)	0.019
PB GS 5–6	17/317 (5.4 %)	13/204 (6.4 %)	4/113 (3.5 %)	
PB GS 7	54/364 (14.8 %)	41/225 (18.2 %)	13/139 (9.4 %)	
Downgrade of PB GS 8–10	31/105 (29.5 %)	13/58 (22.4 %)	18/47 (38.3 %)	0.088

GS Gleason score, PB prostate biopsy

* p values ≤ 0.05 are marked in bold

^a Kappa statistic

Patients in the CP group had higher preoperative PSA levels (mean 10.1 vs. 8.1 ng/ml) and had more often palpable tumors (31 vs. 22 %), and fewer cores were taken during PB (mean 9.7 vs. 10.7). The distribution of PB GS was not significantly different between two groups.

Final pathology revealed significantly higher rates of extraprostatic extension, more GS 8–10 tumors and more PSMs in patients initially diagnosed by a CP. Even when stratified for pathological tumor stage and final GS, persistently higher rates of PSMs could be observed in the CP group. The PSM rate for Gleason 8–10 tumors was 49.5 % in the CP group and 15.2 % in the UP group ($p < 0.001$).

Tables 2 and 3 show the number and percentage of upgrades and downgrades between biopsy and RARP specimen. The overall GS concordance measured by kappa was fair (0.273) in the CP group and moderate (0.411) in the UP group. Significantly higher rates of overall and clinically significant upgrades were found when biopsies were graded by a CP.

Table 4 displays the results of the uni- and multivariable logistic regression analyses. A higher preoperative PSA level, a lower number of biopsy cores as well as grading by a CP predicted GS upgrade in the univariable analysis (Table 4a). In the multivariable analysis, a higher PSA level and grading by a CP remained significant predictors of GS upgrade. The second model assessed the risk of PSMs (Table 4b). In the multivariable analysis, a RP GS of 7, pT3 tumor stage and CP grading remained independent predictive factors for PSMs.

Median follow-up time was 36 (range 1–101) months. A postoperative PSA nadir of <0.1 ng/ml was not reached by 50 (10.3 %) men in the group of patients graded by a CP and by 17 (7.5 %) men in the group of patients graded by a UP ($p = 0.025$). This difference remained significant in a subgroup analysis stratified for risk groups (data not shown). Table 5 displays the results of the Cox regression analysis. In the univariate model, all variables but nerve-sparing were significantly associated with a lower BCRFS rate. In

Table 3 Radical prostatectomy grades stratified by biopsy Gleason scores and tumor stage

RP GS	PB GS 4–6 (PGG 1)		PB GS 3 + 4 = 7 (PGG 2)		PB GS 4 + 3 = 7 (PGG 3)		PB GS 8 (PGG 4)		PB GS 9–10 (PGG 5)		
	All	CP	All	UP	All	CP	All	CP	All	CP	
5–6	131 (41.3)	73 (35.8)	58 (51.3)	15 (5.8)	7 (4.6)	8 (7.4)	3 (2.9)	1 (1.4)	2 (6.5)	0 (0.0)	0 (0.0)
3 + 4 = 7	127 (40.1)	86 (42.2)	41 (36.3)	159 (61.2)	89 (58.6)	70 (64.8)	31 (29.8)	17 (23.3)	14 (45.2)	12 (18.5)	3 (7.9)
4 + 3 = 7	42 (13.2)	32 (15.7)	10 (8.8)	60 (23.1)	39 (25.7)	21 (19.4)	42 (40.4)	31 (42.5)	11 (35.5)	14 (21.5)	9 (23.7)
8	11 (3.5)	8 (3.9)	3 (2.7)	15 (5.8)	8 (5.3)	7 (6.5)	13 (12.5)	10 (13.7)	3 (9.7)	12 (18.5)	5 (13.2)
9–10	6 (1.9)	5 (2.5)	1 (0.9)	11 (4.2)	9 (5.9)	2 (1.9)	15 (14.4)	14 (19.2)	1 (3.2)	27 (41.5)	21 (55.3)
Total	317 (100)	204 (100)	113 (100)	260 (100)	152 (100)	108 (100)	104 (100)	73 (100)	31 (100)	65 (100)	38 (100)
pT2	277 (87.4)	174 (85.3)	103 (91.2)	190 (73.1)	105 (69.1)	85 (78.7)	68 (65.4)	48 (65.8)	20 (64.5)	29 (44.6)	12 (31.6)
pT3a	27 (8.5)	19 (9.3)	8 (7.1)	57 (21.9)	37 (24.3)	20 (15.5)	21 (20.2)	17 (23.3)	4 (12.9)	22 (33.8)	17 (44.7)
pT3b	11 (3.5)	10 (4.9)	1 (0.9)	13 (5)	10 (6.6)	3 (2.8)	15 (14.4)	8 (11.0)	7 (22.6)	14 (21.5)	9 (23.7)
pT4	2 (0.6)	1 (0.5)	1 (0.9)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0.0)	0 (0.0)
Total	317 (100)	204 (100)	113 (100)	260 (100)	150 (100)	108 (100)	104 (100)	73 (100)	31 (100)	65 (100)	38 (100)

RP radical prostatectomy, GS Gleason score, PGG prognostic grade group [1, 14, 15], PB prostate biopsy, CP community pathologist, UP uro-pathologist

the multivariable model, grading by a CP remained an independent predictor of BCR with a hazard ratio (HR) of 1.65 ($p = 0.028$).

The Kaplan–Meier analyses of BCRFS comparing the CP and UP group are shown in Fig. 1a–d. For low-risk patients ($n = 177$), a higher BCR rate was detected in the CP group (Fig. 1b). Comparison of these estimates with survival analyses based on final pathology showed that the UP PB GS 6 curve resembled more the RP GS 6 curve than the CP PB GS 6 curve did (Fig. 1b).

In the intermediate-risk group ($n = 319$), comparison of PB GS 7 BCRFS with estimates based on final pathology showed that the UP PB GS 7 curve resembled more the RP GS 7 curve than the CP PB GS 7 curve did (Fig. 1c).

A significantly lower BCRFS rate could be observed for patients with a RP GS of 8–10 and preoperative PSA level of <20 ng/ml ($n = 92$) when biopsies had been graded by a CP (Fig. 1d).

Discussion

The present study is the first to evaluate the association between pathology report origin and concordance between PB and RP GS, as well as its impact on oncological outcome in a contemporary series of patients treated by RARP for PC. We were able to show that PB GS undergrading was significantly more frequent if PB grading had been performed by a CP compared to a UP. In addition, PB GS grading by a CP was an independent predictor of worse oncological outcome. Men in the CP group more often had PSM, postoperative PSA persistence and BCR compared to the UP group.

In 1992, DF Gleason himself raised the problem of non-dedicated pathologists having the tendency to not recognize small amounts of higher tumor grade [18]. This might explain why in the present study more than half (54.5 %) of the tumors with a Gleason 8–10 on final pathology were not recognized as such by the CP compared to 37 % missed by the UP. Steinberg and colleagues backed Gleason’s statement in 1997 with reporting a higher rate of GS upgrade between the biopsy and the prostatectomy specimen when the biopsy material was analyzed in non-academic settings (37 vs. 28 %) [7]. Kuroiwa et al. [10] reported in 2011 a 16 % higher rate of undergrading by CP. Notably, their studies included data from the pre-ISUP 2005 era and did not assess the prognostic impact of the pathologist on oncological outcome.

An improved concordance rate is not only a theoretical advantage, but is of clinical relevance. Undergrading of PB samples leads to an underestimation of the actual disease burden and can have considerable consequences for therapeutic decision making after diagnosis of PC. Active surveillance, brachytherapy and new therapeutic approaches such as focal therapy are currently considered inappropriate

Table 4 Uni- and multivariable logistic regression models to predict (a) Gleason score upgrade from biopsy to radical prostatectomy and (b) positive surgical margin status

Parameter	Univariable			Multivariable		
	OR	95 % CI	<i>p</i> *	OR	95 % CI	<i>p</i> *
<i>(a)</i>						
Grading pathologist						
Uro-pathologist	1.00	–	–	1.00	–	–
Community pathologist	2.06	1.53–2.8	<0.001	1.91	1.40–2.61	<0.001
Age						
Age ≤64 year	1.00	–	–	–	–	–
Age > 64 year	0.99	0.75–1.31	0.943	–	–	–
PSA value						
PSA <10 ng/ml	1.00	–	–	1.00	–	–
PSA ≥10 ng/ml	1.84	1.35–2.5	<0.001	1.81	1.31–2.5	<0.001
Number of PB cores						
PB cores ≥10	1.00	–	–	1.00	–	–
PB cores <10	1.47	1.10–1.96	0.01	1.32	0.98–1.78	0.07
No. of positive PB cores	1.01	0.96–1.07	0.627	–	–	–
Clinical stage						
cT1	1.00	–	–	–	–	–
cT2	0.90	0.65–1.25	0.537	–	–	–
<i>(b)</i>						
Grading pathologist						
Uro-pathologist	1.00	–	–	1.00	–	–
Community pathologist	1.81	1.3–2.5	<0.001	1.69	1.20–2.38	0.003
PSA value						
PSA <10 ng/ml	1.00	–	–	1.00	–	–
PSA ≥10 ng/ml	1.78	1.28–2.46	0.001	1.21	0.84–1.74	0.302
Number of PB cores						
PB cores ≥10	1.00	–	–	–	–	–
PB cores <10	1.08	0.79–1.48	0.65	–	–	–
Final pathology GS						
RP Gleason score <7	1.00	–	–	1.00	–	–
RP Gleason score = 7	2.76	1.69–4.5	<0.001	2.03	1.21–3.4	0.007
RP Gleason score >7	3.63	2.07–6.38	<0.001	1.57	0.82–3.0	0.173
Clinical stage						
cT2	1.00	–	–	–	–	–
cT3	1.38	0.98–1.95	0.065	–	–	–
Pathological stage						
pT2	1.00	–	–	1.00	–	–
pT3	3.76	2.69–5.26	<0.001	3.28	2.23–4.81	<0.001
Nerve-sparing						
No nerve-sparing	1.00	–	–	–	–	–
Unilateral nerve-sparing	0.68	0.47–0.97	0.036	–	–	–
Bilateral nerve-sparing	0.68	0.47–0.99	0.042	–	–	–

OR odds ratio, 95 % CI 95 % confidence interval, PB prostate biopsy, RP radical prostatectomy

**p* values ≤0.05 are marked in bold

treatment options for most patients with intermediate- or high-risk PC [16, 19]. Furthermore, the extent of surgical resection (LND, nerve-sparing) but also the indication for concomitant androgen deprivation with radiotherapy is

based on preoperative risk stratification [2–4] and can have significant impact on patient's quality of life. Additional treatments may become necessary or opportunities for cure may be missed in patients with misclassified tumors. Our

Table 5 Uni- and multivariable Cox regression analysis of predictors for biochemical recurrence-free survival

Parameter	Univariable			Multivariable		
	HR	95 % CI	<i>p</i> *	HR	95 % CI	<i>p</i> *
Grading pathologist						
Uro-pathologist	1.00	–	–	1.00	–	–
Community pathologist	2.21	1.42–3.43	<0.001	1.65	1.06–2.6	0.028
PSA value						
PSA <10 ng/ml	1.00	–	–	1.00	–	–
PSA ≥10 ng/ml	1.52	1.02–2.27	0.041	0.85	0.55–1.30	0.439
Final pathology GS						
RP Gleason score <7	1.00	–	–	1.00	–	–
RP Gleason score = 7	4.00	1.82–8.70	0.001	2.66	1.20–5.87	0.016
RP Gleason score >7	10.70	4.78–24.2	<0.001	5.18	2.18–12.3	<0.001
Clinical stage						
cT1	1.00	–	–	1.00	–	–
cT2	1.72	1.16–2.56	0.007	1.63	1.09–2.45	0.018
Pathological stage						
pT2	1.00	–	–	1.00	–	–
pT3	6.17	4.22–9.03	<0.001	3.65	2.40–5.55	<0.001
Nerve-sparing						
Negative surgical margin	1.00	–	–	1.00	–	–
Positive surgical margin	2.99	2.06–4.35	<0.001	2.16	1.46–3.20	<0.001
No nerve-sparing	1.00	–	–	–	–	–
Unilateral nerve-sparing	0.80	0.51–1.27	0.345	–	–	–
Bilateral nerve-sparing	0.84	0.53–1.32	0.449	–	–	–

HR hazard ratio, 95 % CI 95 % confidence interval, RP radical prostatectomy

* *p* values ≤0.05 are marked in bold

subgroup analysis for clinically significant undergrading revealed higher rates of misclassification of patients when biopsies were graded by a CP compared to a UP based on inclusion criteria for active surveillance (64.2 vs. 48.7 %) and recommendations for nerve-sparing (12.6 vs. 6.7 %).

Accordingly, based on our hypothesis that an inaccurate PB GS compromises an adequate surgical approach and therefore the oncological outcome, we could show that grading by a CP is an independent predictor of PSMs. A surgeon underestimating the tumor might select inadequate patients for nerve-sparing and perform a more extensive preservation of surrounding structures. The impact of GS underestimation on PSM rates has been described before; Corocoran et al. [20] reported in a retrospective analysis significantly higher PSM rates in upgraded tumors than in corresponding concordant tumors. The PSM rate in undergraded Gleason 3 + 4 tumors was significantly higher than in accurately diagnosed counterparts. In the present investigation, nearly half of the patients with a RP GS of 8–10 had PSMs when PBs were graded by a CP, compared to only 15 % in patients graded by UP. This observation gains importance considering that more than every second patient with a RP GS 8–10 had been assigned a lower GS

preoperatively by a CP. Consequently, the observed rate of nerve-sparing was also significantly higher in this group compared to the group of patients whose PB cores had been graded by a UP. This further indicates the important role of surgeon's decisions in the high-risk situation for a favorable surgical resection of the tumor and possible consequences of undergrading for the oncological outcome (Fig. 1d).

The assessment of BCRFS was performed in this study for two reasons: One motive was to evaluate the impact of preoperative parameters including the grading pathologist on a more objective criterion of midterm oncological outcome. The second reason was to address the issue of an innate bias: Our pathologic department was grading both of the specimens (PB and RP). This might have led to a better concordance between the UP PB GS and RP GS. Therefore, we investigated the prognostic accuracy of the PB and RP GS for BCRFS and discriminated which of the PB GS (CP or UP) was a more precise reflection of tumor behavior.

Kaplan–Meier analysis indicated RP GS 5–6 as an excellent predictor of tumor aggressiveness with a 5-year estimated BCRFS of 96 %. This is in line with large

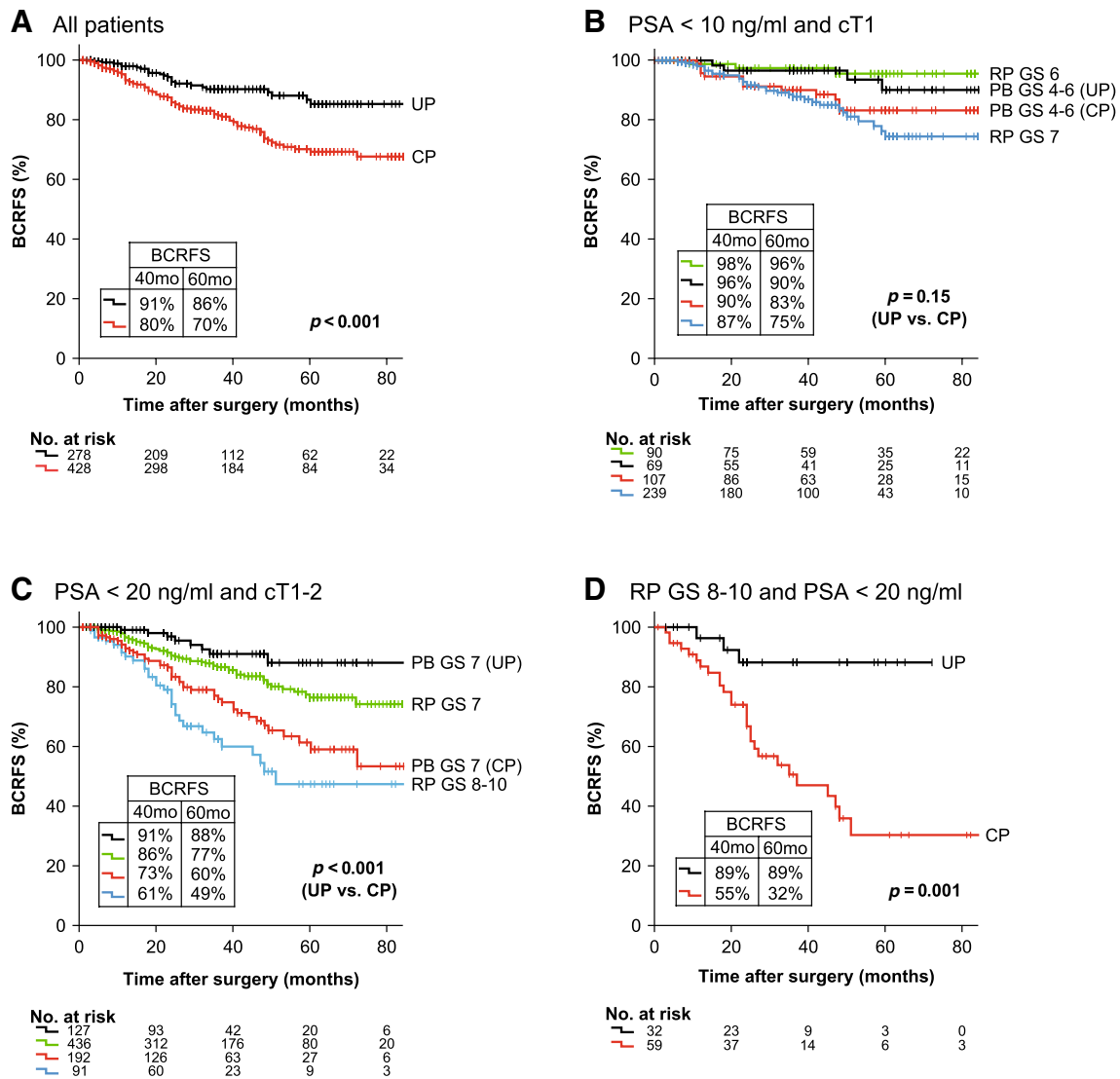


Fig. 1 Kaplan–Meier analyses for biochemical recurrence-free survival. **a** All Patients stratified for origin of pathology report. **b** Patients with a PSA of <10 ng/ml, clinical T1 stage and PB GS of 4–6 stratified for origin of pathology report and compared to concordant and higher RP GS. **c** Patients with a PSA of <20 ng/ml, clinical

T1–2 stage and PB GS of 7 stratified for origin of pathology report and compared to concordant and higher RP GS. **d** Patients with a PSA <20 ng/ml and a RP GS of 8–10 (final pathology) stratified for origin of pathology report

reported series and the expected behavior of these tumors [21]. With an estimated 5-year BCRFS of 90 % in the group of low-risk tumors (based on PB GS), grading by UP achieved a higher predictive accuracy than grading by a CP (5y-BCRFS of 83 %).

A similar constellation was observed in the intermediate-risk group. While the better outcome of patients with a PB GS 7 graded by a UP is probably due to the higher rate of downgrading at RP, the Kaplan–Meier curve of CP-graded patients with PB GS 7 estimates a BCRFS rate more similar to patients with a RP GS 8–10 than RP GS 7, hinting at the larger proportion of actual higher GSs in this group.

This study has limitations. It was not possible to distinguish whether the grading in 2005 and 2006 had been performed according to the classic system or according to the ISUP 2005 consensus. It can be reasonably assumed that the modernized Gleason grading scheme was more rapidly adopted by dedicated UPs than among CPs. However, the high number of patients and a period of almost 9 years strengthen the hypothesis of permanent and significant difference in the accuracy of Gleason grading in the two groups. Furthermore, there were statistically significant differences in the preoperative parameters of the two groups. One potential explanation of these differences is that community urologists do refer patients with more advanced

tumors to academic centers for surgery. To overcome this problem, we performed multivariable logistic and Cox regression analyses and subclassification of patients into risk groups. Finally, a retrospective evaluation of an accurate biopsy technique is not feasible. We analyzed the number of biopsy cores, which was slightly lower for patients diagnosed by a CP. We further assume that a median of 10 cores in the CP and 12 cores in the UP-graded group exceeded the critical cutoff of 6 cores for a significantly increased risk of GS upgrade [6] and had therefore limited impact on the accuracy of the grading in this study. We did not perform a re-evaluation of CP graded Gleason scores by a UP because of legal (necessity of signed declarations of agreement by every single patient) and resource issues (collection of slides from multiple CP archives and re-evaluation of 487 biopsy sets by an experienced UP). We believe that these results demonstrate real life data and sensitize to the actual oncological impact of grading by non dedicated uro-pathologists on patient outcome in daily practice.

Conclusion

Dedication of pathologists not only affects the rate of discordance but also the oncological outcome of patients treated for localized prostate cancer. We strongly recommend prostate biopsy specimen to be reviewed by a dedicated uro-pathologist to improve the concordance between prostate biopsy and radical prostatectomy Gleason score, which in turn allows to choose the most appropriate treatment option and eventually results in a better oncological outcome.

Acknowledgments We kindly thank Dr. Niels Hagenbuch for assistance with methodology and statistical support.

Author contributions A Mortezaei and EX Keller were involved in data collection, data analysis, data interpretation, literature search, generation of figures and writing of the manuscript. C Poyet and T Hermanns were involved in data analysis, data interpretation, writing and final review of the manuscript. K Saba, M Randazzo and C Fankhauser were involved in data collection and data analysis. PJ Wild, H Moch and T Sulser had the idea for this investigation, designed the study and reviewed the results and the manuscript. D Eberli is the senior author of this study. He designed the investigation and was involved in data analysis, data interpretation and writing of the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

Ethical standards The study was approved by local ethics committee (Study Number: StV 25-2008-rev) and complied with the Declaration of Helsinki principles (including revisions), as well as with the guidelines of the International Conference on Harmonisation of Good Clinical Practice (ICH-GCP).

References

- Pierorazio PM, Walsh PC, Partin AW, Epstein JI (2013) Prognostic Gleason grade grouping: data based on the modified Gleason scoring system. *BJU Int* 111(5):753–760. doi:[10.1111/j.1464-410X.2012.11611.x](https://doi.org/10.1111/j.1464-410X.2012.11611.x)
- Montorsi F, Wilson TG, Rosen RC, Ahlering TE, Artibani W, Carroll PR, Costello A, Eastham JA, Ficarra V, Guazzoni G, Menon M, Novara G, Patel VR, Stolzenburg JU, Van der Poel H, Van Poppel H, Mottrie A, Pasadena Consensus P (2012) Best practices in robot-assisted radical prostatectomy: recommendations of the Pasadena Consensus Panel. *Eur Urol* 62(3):368–381. doi:[10.1016/j.eururo.2012.05.057](https://doi.org/10.1016/j.eururo.2012.05.057)
- Horwitz EM, Bae K, Hanks GE, Porter A, Grignon DJ, Brereton HD, Venkatesan V, Lawton CA, Rosenthal SA, Sandler HM, Shipley WU (2008) Ten-year follow-up of radiation therapy oncology group protocol 92-02: a phase III trial of the duration of elective androgen deprivation in locally advanced prostate cancer. *J Clin Oncol* 26(15):2497–2504. doi:[10.1200/JCO.2007.14.9021](https://doi.org/10.1200/JCO.2007.14.9021)
- Pilepich MV, Winter K, Lawton CA, Krisch RE, Wolkov HB, Movsas B, Hug EB, Asbell SO, Grignon D (2005) Androgen suppression adjuvant to definitive radiotherapy in prostate carcinoma—long-term results of phase III RTOG 85-31. *Int J Radiat Oncol Biol Phys* 61(5):1285–1290. doi:[10.1016/j.ijrobp.2004.08.047](https://doi.org/10.1016/j.ijrobp.2004.08.047)
- Cohen MS, Hanley RS, Kurteva T, Ruthazer R, Silverman ML, Sorcini A, Hamawy K, Roth RA, Tuerk I, Libertino JA (2008) Comparing the Gleason prostate biopsy and Gleason prostatectomy grading system: the Lahey Clinic Medical Center experience and an international meta-analysis. *Eur Urol* 54(2):371–381. doi:[10.1016/j.eururo.2008.03.049](https://doi.org/10.1016/j.eururo.2008.03.049)
- Van Praet C, Libbrecht L, D'Hondt F, Decaestecker K, Fonteyne V, Verschuere S, Rottey S, Praet M, De Visschere P, Lumen N, Uro-Oncology Group G (2014) Agreement of Gleason score on prostate biopsy and radical prostatectomy specimen: is there improvement with increased number of biopsy cylinders and the 2005 revised Gleason scoring? *Clin Genitourin Cancer* 12(3):160–166. doi:[10.1016/j.clgc.2013.11.008](https://doi.org/10.1016/j.clgc.2013.11.008)
- Steinberg DM, Sauvageot J, Piantadosi S, Epstein JI (1997) Correlation of prostate needle biopsy and radical prostatectomy Gleason grade in academic and community settings. *Am J Surg Pathol* 21(5):566–576
- Karakiewicz PI, Chun FKH, Gallina A, Suardi N, Briganti A, Erbersdobler A, Schlomm T, Walz J, Currilin E, Michl U, Haese A, Arjane P, Heinzer H, Graefen M, Huland H (2008) Biopsies performed at tertiary care centers are superior to referral biopsies in predicting pathologic Gleason sum. *J Endourol* 22(3):533–538. doi:[10.1089/end.2007.0219](https://doi.org/10.1089/end.2007.0219)
- Fine SW, Epstein JI (2008) A contemporary study correlating prostate needle biopsy and radical prostatectomy Gleason score. *J Urol* 179(4):1335–1339. doi:[10.1016/j.juro.2007.11.057](https://doi.org/10.1016/j.juro.2007.11.057)
- Kuroiwa K, Shiraiishi T, Naito S, Clinicopathological Research Group for Localized Prostate Cancer I (2011) Gleason score correlation between biopsy and prostatectomy specimens and prediction of high-grade Gleason patterns: significance of central pathologic review. *Urology* 77(2):407–411. doi:[10.1016/j.urology.2010.05.030](https://doi.org/10.1016/j.urology.2010.05.030)
- Feicke A, Baumgartner M, Talimi S, Schmid DM, Seifert HH, Muntener M, Fatzer M, Sulser T, Strebel RT (2009) Robotic-assisted laparoscopic extended pelvic lymph node dissection for prostate cancer: surgical technique and experience with the first 99 cases. *Eur Urol* 55(4):876–883. doi:[10.1016/j.eururo.2008.12.006](https://doi.org/10.1016/j.eururo.2008.12.006)
- Umbuhr M, Kessler TM, Sulser T, Kristiansen G, Probst N, Steurer J, Bachmann LM (2008) ProCOC: the

- prostate cancer outcomes cohort study. *BMC urology* 8:9. doi:[10.1186/1471-2490-8-9](https://doi.org/10.1186/1471-2490-8-9)
13. Kramer MS, Feinstein AR (1981) Clinical biostatistics. LIV. The biostatistics of concordance. *Clin Pharmacol Ther* 29(1):111–123
 14. Epstein JI, Egevad L, Amin MB, Delahunt B, Srigley JR, Humphrey PA, and the Grading C (2015) 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*. doi:[10.1097/PAS.0000000000000530](https://doi.org/10.1097/PAS.0000000000000530)
 15. Epstein JI, Zelefsky MJ, Sjoberg DD, Nelson JB, Egevad L, Magi-Galluzzi C, Vickers AJ, Parwani AV, Reuter VE, Fine SW, Eastham JA, Wiklund P, Han M, Reddy CA, Ciezki JP, Nyberg T, Klein EA (2015) A contemporary prostate cancer grading system: a validated alternative to the Gleason score. *Eur Urol*. doi:[10.1016/j.eururo.2015.06.046](https://doi.org/10.1016/j.eururo.2015.06.046)
 16. Heidenreich A, Bastian PJ, Bellmunt J, Bolla M, Joniau S, van der Kwast T, Mason M, Matveev V, Wiegel T, Zattoni F, Mottet N, European Association of U (2014) EAU guidelines on prostate cancer. part 1: screening, diagnosis, and local treatment with curative intent-update 2013. *Eur Urol* 65(1):124–137. doi:[10.1016/j.eururo.2013.09.046](https://doi.org/10.1016/j.eururo.2013.09.046)
 17. D'Amico AV, Whittington R, Malkowicz SB, Schultz D, Blank K, Broderick GA, Tomaszewski JE, Renshaw AA, Kaplan I, Beard CJ, Wein A (1998) Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. *JAMA* 280(11):969–974
 18. Gleason DF (1992) Histologic grading of prostate cancer: a perspective. *Hum Pathol* 23(3):273–279
 19. Donaldson IA, Alonzi R, Barratt D, Barret E, Berge V, Bott S, Bottomley D, Eggener S, Ehdaie B, Emberton M, Hindley R, Leslie T, Miners A, McCartan N, Moore CM, Pinto P, Polascik TJ, Simmons L, van der Meulen J, Villers A, Willis S, Ahmed HU (2015) Focal therapy: patients, interventions, and outcomes—a report from a consensus meeting. *Eur Urol* 67(4):771–777. doi:[10.1016/j.eururo.2014.09.018](https://doi.org/10.1016/j.eururo.2014.09.018)
 20. Corcoran NM, Hong MK, Casey RG, Hurtado-Coll A, Peters J, Harewood L, Goldenberg SL, Hovens CM, Costello AJ, Gleave ME (2011) Upgrade in Gleason score between prostate biopsies and pathology following radical prostatectomy significantly impacts upon the risk of biochemical recurrence. *BJU Int* 108(8 Pt 2):E202–E210. doi:[10.1111/j.1464-410X.2011.10119.x](https://doi.org/10.1111/j.1464-410X.2011.10119.x)
 21. Han M, Partin AW, Pound CR, Epstein JI, Walsh PC (2001) Long-term biochemical disease-free and cancer-specific survival following anatomic radical retropubic prostatectomy. The 15-year Johns Hopkins experience. *The Urologic clinics of North America* 28(3):555–565