



The impact of surgical downgrading on prostate cancer recurrence: systematic review and analysis of a multiethnic population

Denzel Zhu¹ · William Shyr¹ · Michelle Toker¹ · Ethan Fram² · Jinrong Cheng^{1,3} · Evan Z. Kovac^{1,2} · Ilir Agalliu^{2,4} · Ahmed Aboumohamed^{1,2} · Kara L. Watts^{1,2}

Received: 24 August 2021 / Accepted: 13 November 2021 / Published online: 30 November 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

Objective To perform a systematic review and a retrospective cohort analysis evaluating the rates of surgical downgrading of prostate cancer (PCa) from biopsy (PBx) to radical prostatectomy (RP), and their association with biochemical recurrence (BCR) in a multiethnic population.

Methods A systematic review of PubMed and other databases was performed. We included retrospective studies evaluating the relationship between surgical downgrading and BCR-free survival. Data regarding Gleason score (GL) downgrading were abstracted from the articles and categorized as follows: GL8-10 to GL7, GL7 to GL6, and GL 7(4+3) to GL7(3+4). We also performed a retrospective cohort review of patients who underwent RP at our institution from 2005 through 2020. Kaplan–Meier survival analysis and Cox proportional hazards models were used to compare BCR among downgraded versus non-downgraded men.

Results Systematic review yielded 137 abstracts; of these, 36 full-texts were reviewed, 8 of which were included in our systematic review. Despite substantial variability, all showed that GL at RP is one of the most important factors of BCR-free survival. A total of 1,484 men with PCa were analyzed from our institution. On multivariate analysis, GL7 to GL6 downgrading (HR = 0.50, $p = 0.022$) and GL8-10 to GL7 downgrading (HR = 0.42, $p = 0.011$) were associated with reduced risk of BCR when compared to men with GL7 and GL8-10 concordance, respectively. However, GL7(4+3) to GL7(3+4) downgrading was not significantly associated with reduced BCR (HR = 0.56, $p = 0.12$), when compared to GL7(4+3) concordance, although HR was similar.

Conclusion Surgical downgrading at RP was associated with a reduced risk of BCR compared to GL concordant disease, and these findings have been validated within our multiethnic population. Pathologic downgrading at the time of RP may be a more useful predictor of subsequent BCR in comparison to that associated with GL concordant pathology.

Keywords Prostate cancer · Prostate biopsy · Surgical downgrading · Radical prostatectomy · Non-Hispanic black · Non-Hispanic white · Hispanic · Multiethnic · Systematic review · Biochemical recurrence

Introduction

Prostate cancer (PCa) is the most common non-skin malignancy among US men [1]. The incidence and outcomes of PCa vary by race/ethnicity: non-Hispanic Black (NHB) men and Hispanic men have both higher risk for the development of PCa and lower 5-year survival rates for PCa, when

compared to non-Hispanic White (NHW) men [2]. Additionally, NHB men are at a higher risk of having biochemical recurrence (BCR) following radical prostatectomy (RP) compared to NHW men [3]. Efforts are needed to better characterize the prognosis of PCa within different racial and ethnic groups.

While surgical upgrading from prostate biopsy (PBx) to RP surgical pathology (SPx) has been well studied, to our knowledge, data on the outcomes of patients that are downgraded from PBx to RP are limited, particularly amongst high risk cohorts [4]. Furthermore, surgical downgrading can lead to suboptimal care for patients as these patients may

Denzel Zhu and William Shyr contributed equally as first authors.

✉ Kara L. Watts
kwatts@montefiore.org

Extended author information available on the last page of the article

have been reasonable candidates for less invasive therapy, such as active surveillance.

The goal of this study was to perform a systematic review of peer-reviewed published papers describing surgical downgrading of PCa at the time of RP and its impact on biochemical recurrence (BCR), with particular attention to the representation of racial/ethnic minority groups. A secondary aim of this study was to examine the impact of surgical downgrading on BCR within our own diverse, multiethnic patient population and compare our findings to those of the systematic review.

Materials and methods

Systematic review

Search strategy

The PubMed, EMBASE, MEDLINE, and Cochrane library databases were used to identify full text studies published in English language comparing patients who underwent surgical downgrading at the time of RP to their non-downgraded peers. The following search terms were used: (“prostate cancer”) AND (“downgrade”) AND (“surgical”). These terms were submitted in an identical manner to each database. Studies published from January 1984 through November 2020 were included.

Inclusion and exclusion criteria

All studies in which men older than 18 years underwent a standard or fusion prostate biopsy (PBx) with subsequent RP for PCa were reviewed. We included all studies which stratified surgical downgrading, defined as moving from a higher GL on PBx to a lower GL on RP, and had reported post-operative outcomes, such as recurrence-free survival or BCR (Supplementary Figure S1). We excluded studies that did not report on surgical downgrading by distinct GL groups, studies that pooled multiple GL groups into one combined group, studies that did not report post-operative oncologic outcomes, and those with fewer than 20 subjects in total. Articles published in a non-English language were also excluded. If it was unclear whether an article met inclusion criteria based on review of abstract, the full text was reviewed. The literature search was completed by one investigator (DZ).

Of the included studies, data were extracted by one investigator (DZ), including patient demographics, downgrading definition, clinical variables, and clinical outcome rates (e.g., BCR) for the downgraded group vs the non-downgraded group.

Retrospective cohort study

Patient population

Following institutional review board approval, we identified all patients who underwent an RP for biopsy-proven PCa from our institution from 2005 to 2020. Biochemical recurrence (BCR) was defined as the first post-operative PSA ≥ 0.2 ng/mL. Follow-up time was defined as the time from radical prostatectomy to the last encounter with our health system. Demographic and clinical variables, such as age at surgery, race/ethnicity, pre-operative body mass index (BMI), and pre-operative PSA were extracted by manual review of electronic medical records (EMR), queries of the EMR using Clinical Looking Glass (Streamline Health, Atlanta, GA) [5], and our institutional Cancer Registry [6]. Adverse histologic features, tumor grade, and pathologic tumor stage were extracted by manual review of RP pathology reports.

We examined different downgrading types, from GL7 to GL6, GL8-10 [7] to GL7, and GL7(4+3) to GL7(3+4). We compared patients who were downgraded against patients who had concordant GL scores (i.e., both PBx and RP GL scores are the same) of both the original and lower GL score (e.g., for GL7 to GL6 downgrading, we compared GL7 to GL6 downgraded patients against patients with GL7 concordance and GL6 concordance). Note that patients with GL7(4+3) to GL7(3+4) downgrading and GL7(3+4) to GL7(4+3) upgrading were included within the GL7 concordance group.

Statistical analysis

Baseline demographics were compared between patients who had surgical downgrading vs those without downgrading. Patients who experienced upgrading were excluded from this analysis. Continuous, normally distributed variables were compared using one-way analysis of variance (ANOVA). Continuous, not-normally distributed variables were compared using the Kruskal–Wallis H test, and categorical variables were compared using the χ^2 -test. Rates of BCR were compared using the log-rank test. All statistical tests were two-sided, using a significance level $p \leq 0.05$. We followed up patients from date of PBx through date of BCR, death (if the patient died), or end of follow-up (last patient contact with our institution). Kaplan–Meier survival analysis was used to compare BCR-free survival between different downgrading strata [8, 9]. A Cox proportional hazards (PH) model was used to calculate and compare the cumulative BCR-free survival among downgraded and non-downgraded patients, and to

estimate the hazard ratio (HR), after adjusting for age at biopsy (years), race/ethnicity, pre-operative PSA, and the presence of obesity (defined as $\text{BMI} \geq 30$) [9]. Lastly, we performed a sensitivity analysis in which we replicated the analysis using Gleason Group (GG) definitions (GG4/5 to GG3 downgrading, or GG4/5 to GG2 downgrading, GG3 to GG2 downgrading, GG3 to GG1 downgrading, and GG2 to GG1 downgrading). Kaplan–Meier survival curves were plotted using Stata v15.1 (StataCorp, College Station, TX); all other analyses were done in SPSS v27 (IBM, Armonk, NY).

Results

Systematic review

A total of 137 unique manuscripts and abstracts were identified using our search criteria (Supplementary Fig. S1). Of these, 101 were excluded based on their title or abstract details not reflecting the scope of our analysis. Of the remaining studies, 28 were excluded based on the criteria listed in Supplemental Fig. S1. Eight retrospective studies examining the relationship of BCR among men who had surgically downgraded PCa were analyzed for this review (Table 1). A total of 9,059 men composed the study population, of which 1,608 (17.8%) experienced surgical downgrading. All studies found that downgraded patients had lower rates of BCR compared to patients with concordant pathology. However, only 3 out of 8 included studies provided data on race and ethnicity, and in those studies, the majority were non-Hispanic White patients.

Retrospective cohort study

Of the 1621 patients that underwent RP at our institution, 137 were excluded from the analysis due to the lack of pre-operative prostate biopsy data. Among our eligible 1,484 patients, 37.6% were NHB and 28.8% were Hispanic (Supplementary Table S1). Patients downgraded from GL7 to GL6 were compared to GL6 concordant and GL7 concordant patients (Table 2A). Only 2.6% ($n=3$) of downgraded patients had a 6-month delay between PBx and RP. Patients downgraded from GL7 to GL6 had similar actuarial rates of BCR (15.5%) compared to those with GL6 concordance (13.0%), which were both lower than men with GL7 concordance (22.6%; $p < 0.001$, Fig. 1A).

Patients who were downgraded from high risk PCa (GL8–10) to intermediate-risk disease (GL7) were compared to patients who remained with high-risk PCa pathology (GL8–10 concordance) and GL7 concordance (Table 2B). Only 3.4% ($n=3$) of downgraded patients had a 6-month delay between PBx and RP. Patients with downgrading from high

risk PCa had lower actuarial rates of BCR ($n=25$, 28.7%) compared to GL8–10 concordance patients ($n=19$, 44.2%) and similar rates of BCR compared to GL7 concordant patients ($n=102$, 22.6%, Fig. 1B). We also reviewed our data for downgrading from GL8–10 to GL6 disease, and found 21 cases (Supplementary Figure S2).

Patients who were downgraded from GL4+3 to GL3+4 PCa were compared against GL4+3 concordant and GL3+4 concordant patients (Table 2C). A 6-month delay between PBx and RP occurred in only 4.3% ($n=4$) of downgraded patients. Patients who downgraded from GL4+3 to GL3+4 had similar actuarial rates of BCR ($n=17$, 18.5%) compared to GL3+4 patients ($n=41$, 17.1%), which were lower than GL4+3 concordance patients ($n=22$, 36.1%, $p=0.008$, Fig. 1C).

The association of surgical downgrading with BCR is reported in Table 3. On univariate analysis, downgrading from GL7 to GL6 was associated with a 51% reduction in the risk of BCR (HR = 0.49, 95% CI 0.29–0.83, $p=0.008$), with a similar HR magnitude compared to GL6 concordance (HR = 0.43, 95% CI 0.30–0.62, $p < 0.001$), when compared to GL7 concordance patients as reference. Downgrading from high risk PCa (GL 8–10) to GL7 was associated with a significant 52% reduction in the risk of BCR (HR = 0.48, 95% CI 0.26–0.88, $p=0.018$). The magnitude of the HR was slightly higher than GL7 concordance (HR = 0.37, 95% CI 0.22–0.60, $p < 0.001$). GL7(4+3) downgrading was associated with a 49% reduction in the risk of BCR (HR = 0.51, 95% CI 0.27–0.97, $p=0.039$), with a similar magnitude compared to GL7(3+4) concordance (HR = 0.46, 95% CI 0.27–0.77, $p=0.003$). When the models were adjusted for age, race/ethnicity, pre-operative PSA, and the presence of obesity, the results were similar for GL7 to GL6 downgrading (HR = 0.50, 95% CI 0.28–0.90, $p=0.022$) and GL8–10 to GL7 downgrading (HR = 0.42, 95% CI 0.22–0.82, $p=0.011$), but GL7(4+3) to GL7(3+4) downgrading was no longer statistically significant but with similar HR of 0.56 (95% CI 0.27–1.15, $p=0.12$). When we examined the association of BCR using Gleason Grade (GG) downgrading definitions, we failed to detect statistically significant associations between surgical downgrading and BCR (Supplementary Table S2).

Discussion

Discrepancies in Gleason score between PBx and RP are well recognized, with many studies examining the prognostic impact of surgical upgrading on BCR [10, 11]. Compared to surgical upgrading, the prognostic impact of surgical downgrading on BCR has been less explored. To our knowledge, no study has attempted to systematically examine the literature on this topic and summarize the effects

Table 1 Summary data for studies included in review

SGL7 → GL6 study	Population size, <i>n</i>		BCR, <i>n</i> (%)		Follow-up time, median, IQR, mo		Age at diagnosis, median, IQR, y		Race/ethnicity, <i>n</i> (%)		PSA, ng/mL, median, IQR	
	No downgrade	Downgrade	No downgrade	Downgrade	No downgrade	Downgrade	No downgrade	Downgrade	No downgrade	Downgrade	No downgrade	Downgrade
Su (2020) [16]	2650	223	769 (29.0)	223 (11.7)	36, 12–156		61, 55–65	60, 54–65	W: 2129 (80.3), B: 353 (13.3), H: 40 (1.5), A: 33 (1.2), O: 63 (2.4)	W: 186 (83.4), B: 26 (11.7), H: 1 (0.4), A: 2 (0.9), O: 7 (3.1)	5.3, 4.2–7.2	4.5, 3.5–5.8
Ham (2017) [17]	2665	648	1015 (38.1)	96 (14.8)	48, 12–312		60, 36–76	60, 34–74	W: 2254 (84.6), B: 280 (10.5), H: 131 (4.9)	W: 549 (84.7), B: 75 (11.6), O: 24 (3.7)	5.3, 0.01–79.1	4.7, 0.02–23.9
Gondo (2015) [18] <i>GL4 + 3 → GL3 + 4</i>	1202	115	NR	NR	22, NR		61, 56–66	60, 55–64	NR	NR	5.3, 4.0–7.3	4.8, 3.1–6.2
Jang (2019) [11]	161	125	83 (51.6)	43 (34.4)	65, 45–83		67, 62–71	65, 62–71	NR	NR	9.3, 6.4–15.9	8.7, 5.6–13.3
Gonzalvo (2006) [12]	68	16	NR	NR	NR		NR	NR	NR	NR	NR	NR
<i>GL8-10 → GL7</i>												
Altok (2019) [13]	45	46	22 (48.9)	13 (28.3)	60, 41–75		63, 57–66	61, 56–67	W: 33 (73.3), B: 6 (13.3), H: 5 (11.1), A: 0 (0), O: 1 (2.2)	W: 37 (80.4), B: 3 (6.5), H: 2 (4.3), A: 2 (4.3), O: 2 (4.3)	7.0, 5.2–9.4	6.0, 4.6–8.5
Chung (2019) [14]	528	332	404 (76.5)	163 (49.0)	51, 26–73		67, 62–71	66, 62–70	NR	NR	12.0, 7.1–22.7	8.6, 5.9–14.7
Donohue (2006) [15]	132	103	81 (61.4)	50 (48.5)	49.2, 25.2–86.4		NR	NR	NR	NR	NR	NR

NR not reported, W White, B Black, H Hispanic, A Asian, O Other

Table 2 Association of demographic, pathologic, and clinical variables with (A) Gleason score (GL) 7 to GL 6 downgrading, (B) GL 8–10 to GL 7 downgrading, and (C) GL 7(4+3) to GL 7(3+4) downgrading

A	GL7 → GL6 N=116	GL6 → GL6 N=330	GL7 → GL7 N=452	p
Age, yr, mean, SD	58.5, 7.8	57.9, 6.7	60.7, 7.0	<0.001
Follow-up, mo, median (IQR)	79 (53–123)	102 (67–137)	65 (32–114)	
Time (days) from PBx to RP, median (IQR)	79 (59–109)	75 (57.3–107.8)	71 (58–108)	
>6-Month delay, n (%)	3 (5.1)	21 (10.8)	20 (7.1)	
Race/ethnicity, n (%)				0.118
Non-Hispanic White	14 (12.1)	47 (14.2)	32 (7.1)	
Non-Hispanic Black	39 (33.6)	116 (35.2)	179 (39.6)	
Hispanic	35 (30.2)	97 (29.4)	135 (29.9)	
Other/unknown	28 (24.1)	70 (21.2)	106 (23.5)	
Fusion biopsy (vs Std), n (%)	3 (2.6)	0 (0.0)	33 (7.3)	<0.001
Pre-operative PSA, n (%)				<0.001
<4	20 (17.2)	62 (18.8)	38 (8.4)	
4–9.9	57 (49.1)	155 (47.0)	222 (49.1)	
≥10	22 (19.0)	36 (10.9)	122 (27.0)	
Missing	17 (14.7)	77 (23.3)	70 (15.5)	
Pathological tumor stage, n (%)				0.097
T1 or T2a/b	93 (80.2)	269 (81.5)	375 (83.0)	
T2c	2 (1.7)	8 (2.4)	14 (3.1)	
T3a	0 (0.0)	0 (0.0)	7 (1.5)	
T3b/c or T4	0 (0.0)	1 (0.3)	2 (0.4)	
Missing	21 (18.1)	52 (15.8)	54 (11.9)	
Cancer in prostate, %, median (IQR)	5.0 (3.3–15.0)	5.0 (3.0–10.0)	11.0 (6.0–20.0)	<0.001
Prostate weight, g, median (IQR)	43.6 (35.3–64.8)	48.0 (36.6–63.0)	44.8 (35.0–55.0)	0.12
Adverse pathologic features				
Lymphovascular invasion, n (%)	5 (4.3)	3 (0.9)	30 (6.6)	<0.001
Perineural invasion, n (%)	77 (66.4)	167 (50.6)	394 (87.2)	<0.001
Seminal vesical invasion, n (%)	0 (0.0)	1 (0.3)	51 (11.3)	<0.001
Positive surgical margin, n (%)	26 (22.4)	71 (21.5)	201 (44.5)	<0.001
B	GL8-10 → GL7 N=87	GL7 → GL7 N=452	GL8-10 → GL8-10 N=43	p
Age, yr, mean, SD	61.6, 6.7	60.7, 7.0	61.4, 6.8	0.495
Follow-up, mo, median (IQR)	75 (43–97)	65 (32–114)	49 (31–73)	
Time from PBx to RP, d, median (IQR)	69 (56–88)	79 (59–109)	58 (41–88)	
>6-Month delay, n (%)	3 (5.9)	20 (7.1)	0 (0)	
Race/ethnicity, n (%)				0.867
Non-Hispanic White	7 (8.0)	32 (7.1)	4 (9.3)	
Non-Hispanic Black	34 (39.1)	179 (39.6)	12 (27.9)	
Hispanic	26 (29.9)	135 (29.9)	16 (37.2)	
Other/unknown	20 (23.0)	106 (23.5)	11 (25.6)	
Fusion biopsy (VS STD), N (%)	4 (4.6)	33 (7.3)	2 (4.7)	0.559
Pre-operative PSA, n (%)				0.093
<4	4 (4.6)	38 (8.4)	0 (0.0)	
4–9.9	45 (51.7)	222 (49.1)	20 (46.5)	
≥10	31 (35.6)	122 (27.0)	16 (37.2)	
Missing	7 (8.0)	70 (15.5)	7 (16.3)	
Pathological tumor stage, n (%)				0.350
T1 or T2a/b	69 (79.3)	375 (83.0)	32 (74.4)	

Table 2 (continued)

B	GL8-10 → GL7 N=87	GL7 → GL7 N=452	GL8-10 → GL8-10 N=43	<i>p</i>
T2c	3 (3.4)	14 (3.1)	3 (7.0)	
T3a	2 (2.3)	7 (1.5)	3 (7.0)	
T3b/c or T4	0 (0.0)	2 (0.4)	0 (0.0)	
Missing	13 (14.9)	54 (11.9)	5 (11.6)	
Cancer in prostate, %, median (IQR)	10.0 (5.0–20.0)	11.0 (6.0–20.0)	20.0 (9.0–35.0)	0.117
Prostate weight, g, median (IQR)	45.0 (37.6–60.0)	44.8 (35.0–55.0)	48.0 (40.0–61.0)	0.192
Adverse pathologic features				
Lymphovascular invasion, <i>n</i> (%)	10 (11.5)	30 (6.6)	16 (37.2)	<0.001
Perineural invasion, <i>n</i> (%)	79 (90.8)	394 (87.2)	38 (88.4)	0.513
Seminal vesical invasion, <i>n</i> (%)	14 (16.1)	51 (11.3)	15 (34.9)	<0.001
Positive surgical margin, <i>n</i> (%)	34 (39.1)	201 (44.5)	28 (65.1)	0.069
C	GL4+3 → GL3+4 N=92	GL3+4 → GL3+4 N=240	GL4+3 → GL4+3 N=61	<i>p</i>
Age, yr, mean, SD	61.1, 7.1	60.6, 7.0	60.9, 6.9	0.826
Follow-up, mo, median (IQR)	64 (32–113)	66 (29–116)	60 (33–116)	
Time from PBx to RP, d, median (IQR)	81 (58–107)	80 (60–114)	79 (59–110)	
>6-Month delay, <i>n</i> (%)	4 (6.8)	12 (7.7)	2 (5.4)	
Race/ethnicity, <i>n</i> (%)				0.578
Non-Hispanic White	5 (5.4)	21 (8.8)	2 (3.3)	
Non-Hispanic Black	32 (34.8)	98 (40.8)	27 (44.3)	
Hispanic	31 (33.7)	67 (27.9)	19 (31.1)	
Other/unknown	24 (26.1)	54 (22.5)	13 (21.3)	
Fusion biopsy (vs Std), <i>n</i> (%)	6 (6.5)	20 (8.3)	6 (9.8)	0.752
Pre-operative PSA, <i>n</i> (%)				0.581
<4	8 (8.7)	21 (8.8)	5 (8.2)	
4–9.9	50 (54.3)	116 (48.3)	26 (42.6)	
≥10	23 (25.0)	59 (24.6)	21 (34.4)	
Missing	11 (12.0)	44 (18.3)	9 (14.8)	
Pathological tumor stage, <i>n</i> (%)				0.011
T1 or T2a/b	81 (88.0)	198 (82.5)	52 (85.2)	
T2c	3 (3.3)	10 (4.2)	1 (1.6)	
T3a	0 (0.0)	1 (0.4)	4 (6.6)	
T3b/c or T4	0 (0.0)	1 (0.4)	0 (0.0)	
Missing	8 (8.7)	30 (12.5)	4 (6.6)	
Cancer in prostate, %, median (IQR)	13.5 (10.0–20.0)	10.0 (5.0–20.0)	20.0 (10.0–30.0)	0.002
Prostate weight, g, median (IQR)	45.0 (36.3–54.0)	45.8 (36.0–56.0)	41.1 (34.0–55.0)	0.520
Adverse pathologic features				
Lymphovascular invasion, <i>n</i> (%)	4 (4.3)	10 (4.2)	9 (14.8)	0.016
Perineural invasion, <i>n</i> (%)	77 (83.7)	206 (85.8)	58 (95.1)	0.274
Seminal vesical invasion, <i>n</i> (%)	13 (14.1)	21 (8.8)	10 (16.4)	0.418
Positive surgical margin, <i>n</i> (%)	41 (44.6)	96 (40.0)	30 (49.2)	0.653

Data refers to results of our retrospective cohort study

p value refers to one-way ANOVA for normally distributed continuous variables (e.g. age), Kruskal–Wallis *H* test for not-normally distributed continuous variables (e.g. cancer in prostate, prostate weight), and χ^2 -test for categorical variables

Bolded *p* values are considered significant ($p < 0.05$). Note that difference in GL 7 concordance to sum of GL 4+3 to GL 4+3, GL 4+3 to GL 3+4, and GL 3+4 to GL 3+4 is due to 59 GL 3+4 to 4+3 upgraded cases, which are included in the GL 7 concordance stratum

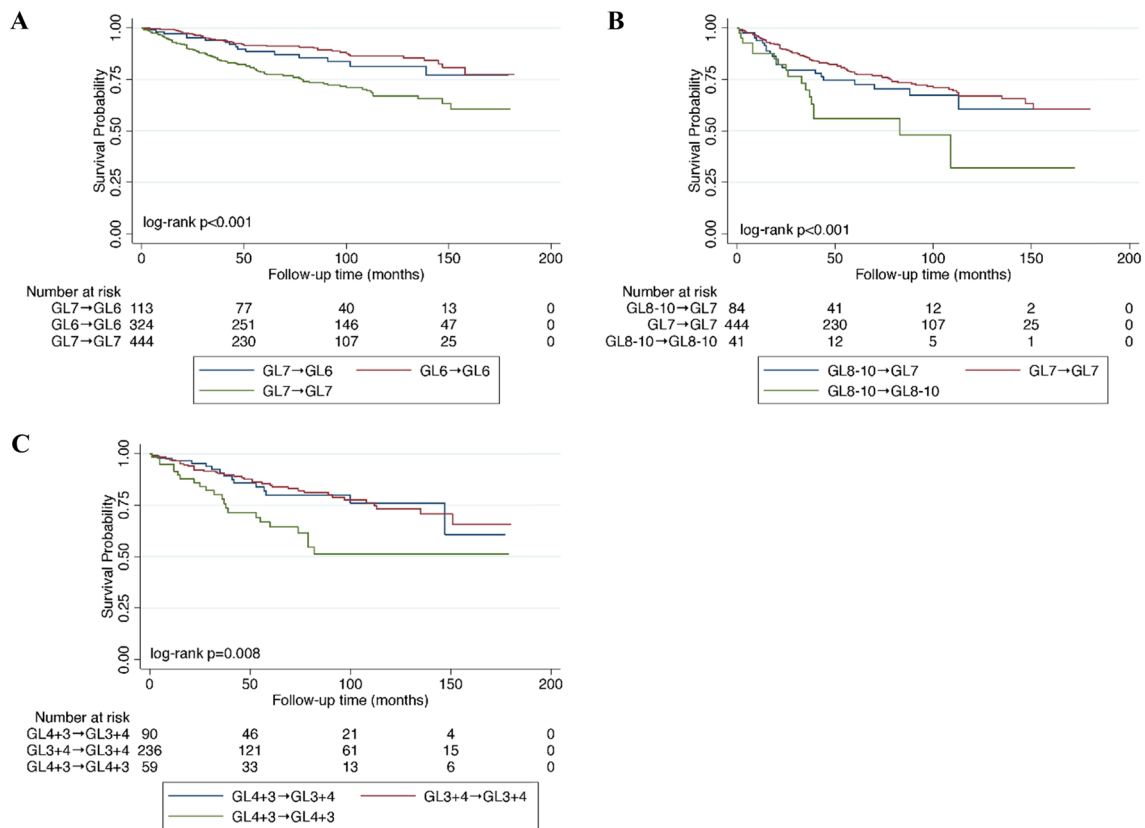


Fig. 1 Kaplan–Meier curves for biochemical recurrence-free survival following radical prostatectomy of prostate cancer **A** GL7→GL6 downgraded patients, **B** GL8-10→GL7 downgraded patients, **C** GL7(4+3)→GL7(3+4) downgraded patients

Table 3 Association of surgical downgrading with biochemical recurrence following robotic assisted radical prostatectomy, among Gleason 7 to 6 downgrading (A), Gleason 8–10 to Gleason 7 downgrading (B), and Gleason 4+3 to 3+4 downgrading (C)

Type	BCR, N (%)	Univariate		Multivariate*	
		HR (95% CI)	<i>p</i>	HR (95% CI)	<i>p</i>
A					
GL7→GL6 (N=116)	18 (15.5)	0.49 (0.29–0.83)	0.008	0.50 (0.28–0.90)	0.022
GL6→GL6 (N=330)	43 (13.0)	0.43 (0.30–0.62)	< 0.001	0.41 (0.27–0.64)	< 0.001
GL7→GL7 (N=452)	102 (22.6)	1	ref	1	ref
B					
GL8-10→GL7 (N=87)	25 (28.7)	0.48 (0.26–0.88)	0.018	0.42 (0.22–0.82)	0.011
GL7→GL7 (N=452)	102 (22.6)	0.37 (0.22–0.60)	< 0.001	0.33 (0.19–0.58)	< 0.001
GL8-10→GL8-10 (N=43)	19 (44.2)	1	ref	1	ref
C					
GL4+3→GL3+4 (N=92)	17 (18.5)	0.51 (0.27–0.97)	0.039	0.56 (0.27–1.15)	0.116
GL3+4→GL3+4 (N=240)	41 (17.1)	0.46 (0.27–0.77)	0.003	0.43 (0.24–0.79)	0.007
GL4+3→GL4+3 (N=61)	22 (36.1)	1	ref	1	ref

* The multivariate model was adjusted for age at surgery, race/ethnicity, pre-operative PSA, and obesity at RP (yes vs no)

GL Gleason (score), HR Hazard ratio, 95% CI 95% Confidence interval

Bolded *p* values are considered significant ($p < 0.05$). Log-rank *p* value for BCR in Table 3A is $p < 0.001$, for Table 3B is $p < 0.001$, and for Table 3C is $p = 0.008$

of our proposed groupings of surgical downgrading (e.g., GL8–10 to GL7, GL7 to GL6, GL7(4+3) to GL7(3+4)) on clinical outcomes.

When we examined the impact of surgical downgrading on BCR, we found that GL at RP is a stronger determinant of BCR compared to GL at PBx. In particular, there was significant reduction in the risk of BCR by at least 50% for men who were downgraded from GL8–10 to GL7 as well as those downgraded from GL7 to GL6. Our models of GL7(4+3) to GL7(3+4) downgrading had similar HR compared to GL7(3+4) concordance patients; however, it failed to reach significance. This may be limited due to a smaller cohort of men with GL7(4+3) downgrading ($n=42$). Lastly, when we examined our data using the GG downgrading definitions, we failed to detect statistically significant associations between surgical downgrading and BCR, but this is likely due to a small sample size of men with GG3 and GG2 on either biopsy or RP.

Our data are consistent with prior studies which found that GL at RP more accurately determines the risk of BCR compared to GL at PBx. In our systematic review, we found two studies that examined downgrading from GL7(4+3) to GL7(3+4). Jang et al. [12], in a review of 286 patients, of which 125 were downgraded, found that downgraded patients had lower rates of BCR (34.4% vs 51.6%, $p=0.001$). Furthermore, a smaller study of 84 patients (16 downgraded) also found a difference in BCR-rates between the two groups, although they did not report the numbers of patients who developed BCR [13]. We similarly detected a difference in the rates of BCR among our cohort (downgraded: 18.5%, concordance: 36.1%, $p=0.008$, Table 2C), although we did not measure a significant result within our multivariate models (Table 3).

Three studies examined downgrading from high-risk PCa (GL8–10) to GL7 [14–16]. A large study of 860 patients (332 downgraded at RP) found that rates of BCR differed when comparing downgraded patients to non-downgraded patients (49% vs 76.5%, respectively, $p<0.001$) [15]. These rates were similar to a medium-sized review of 235 patients (103 downgraded) which reported similar rates of BCR among downgraded men (48.5%) compared to GL8–10 concordant (61.4%, $p=0.0004$) patients. A smaller review of 91 patients (46 downgraded) found significantly lower rates of BCR among downgraded (28.3%) vs non-downgraded (48.9%) patients, although low sample size may have biased results among this cohort [14]. Our results were more similar to the smaller study (downgraded: 28.7% vs concordant: 44.2%, $p<0.001$, Table 2B), suggesting that rates of BCR vary significantly between sites.

Downgrading from GL7 to GL6 was evaluated by three studies included in the systematic analysis [17–19]. Large retrospective reviews from Su et al. [17] and Ham et al. [18] found that downgrading was associated with lower rates of

BCR (11.7% and 14.8%, respectively) compared to GL7 concordance (29% and 38.1%, respectively). Additionally, Ham et al. estimated HRs of BCR among GL7 downgraded patients and reported that GL7 downgrading was associated with a HR:1.87 (95% CI 1.40–2.51, $p<0.0001$) compared to GL6 concordance (reference) and GL7 concordance (HR:4.09, 95% CI 3.50–4.78, $p<0.0001$). Our adjusted models found that GL7 downgrading and GL6 concordance have similar risks of BCR, which represents a difference between our cohorts. Lastly, a smaller study of 1317 men (115 downgraded) found similar results, but did not report their rates of BCR [19].

Etiologically, the discordance between GL on PBx and RP is likely in part due to sampling error during PBx. The small amount of tissue obtained during PBx likely does not reflect the totality of PCa when the prostate is examined following RP. Discordance between PBx and SPx is worsened due to different reporting guidelines for PBx and SPx specimens: for PBx, pathologists report the most prevalent Gleason pattern and the highest grade pattern, while for SPx specimens, pathologists report the first and second most prevalent Gleason patterns. This is supported by studies which have been able to use percentages of cores positive for PCa to predict surgical downgrading [20–22]. Interobserver heterogeneity in grading PBx and surgical pathology has been suggested as well, although a well-validated study on this subject has failed to demonstrate significant heterogeneity occurring between pathologists [23]. Lastly, it now appears that targeted biopsy of MRI lesions may lead to more downgrading compared to standard PBx, as fusion biopsy has been shown to be associated with downgrading [24]. This may be due to the fact that cores from prostatic lesions are likely to be higher grade compared to cores taken systematically, resulting in discordance when the prostate is evaluated following surgery.

Our study has several strengths to consider. Our sample size is relatively large and with longer follow-up times compared to previously reported studies. In addition, our population is racially/ethnically diverse, contributing to the body of literature in this domain. Of the studies reviewed, only 3 reported race/ethnicity data, with NHW men composing $\geq 73\%$ of the study population (Table 1) [14, 17, 18]. We found this to be a significant limitation during our systematic review of the literature on this topic, given that PCa incidence and mortality varies by race/ethnicity [2] and that NHB men are at higher risk for BCR following RP compared to NHW men [3]. In our study, we could not perform stratified analyses by race/ethnicity to better examine the impact of race/ethnicity on surgical downgrading. We can only conclude that we were able to replicate findings of other studies in a majority NHB (37.6%) and Hispanic (28.8%) population (Supplementary Table S1). Analysis of larger datasets of men with PCa treated with RP is needed

to better understand the relationship between race/ethnicity and surgical downgrading.

Limitations of our study include its retrospective nature, including some incomplete biopsy data. We also could not perform a detailed analysis on the number or percentage of cores involved in PCa, which may provide data on the risk of downgrading. Additionally, there have been changes in the definition of the Gleason score from 2005 to 2015, with some of the older Gleason pattern 3 morphologies being shifted to Gleason pattern 4. Given that this period takes place during our systematic review, our results should be interpreted with this finding in mind. Lastly, we could not replicate our analysis when we used GG definitions for surgical downgrading, due to low sample sizes. Future work is needed to re-assess these findings using the GG definitions for surgical downgrading. Our findings further support the need for more advanced methods of PBx in order to minimize discordance between PBx and surgical pathology, as we have demonstrated that surgical pathology is more likely to determine the clinical outcome of patients compared to their PBx results.

Conclusions

Our systematic review and cohort study support prior findings from other authors that surgical downgrading of prostate cancer is associated with a lower risk of BCR compared to concordant Gleason score on PBx pathology. We have validated these findings in our multiethnic population. Further research is needed in the development of advanced PBx techniques and biomarkers to limit discordance between PBx and RP pathology, and to better optimize care.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00345-021-03892-2>.

Acknowledgements We thank Julie Chung of the Montefiore Medical Center Cancer Registry for assistance with data collection.

Funding This study received no funding.

Data availability The data to support this study are not publicly available since the data contain potentially identifiable patient information. Deidentified data are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest None of the named authors have any conflicts of interests, financial or otherwise, to declare.

Ethics standards This study was approved by the Institutional Review Board of the Albert Einstein College of Medicine under protocol #2013–2712.

References

1. Siegel RL, Miller KD, Jemal A (2020) Cancer statistics, 2020. *CA Cancer J Clin* 70(1):7–30
2. Siegel DA, O’Neil ME, Richards TB, Dowling NF, Weir HK (2020) Prostate cancer incidence and survival, by stage and race/ethnicity—United States, 2001–2017. *MMWR Morb Mortal Wkly Rep* 69(41):1473–1480
3. Faisal FA, Sundi D, Cooper JL et al (2014) Racial disparities in oncologic outcomes after radical prostatectomy: long-term follow-up. *Urology* 84(6):1434–1441
4. Bullock N, Simpkin A, Fowler S, Varma M, Kynaston H, Narahari K (2019) Pathological upgrading in prostate cancer treated with surgery in the United Kingdom: trends and risk factors from the British Association of Urological Surgeons Radical Prostatectomy Registry. *BMC Urol* 19(1):94
5. Bellin E, Fletcher DD, Geberer N, Islam S, Srivastava N (2010) Democratizing information creation from health care data for quality improvement, research, and education—the Montefiore Medical Center Experience. *Acad Med* 85(8):1362–1368
6. Suresh T, Sharma J, Goel S, Kovac E, Gao Q, Gartrell B (2019) Patient characteristics at prostate cancer diagnosis in different races at an academic center serving a diverse population. *Clin Genitourin Cancer* 17(2):139–144
7. Hurwitz LM, Agalliu I, Albanes D et al (2020) Recommended definitions of aggressive prostate cancer for etiologic epidemiologic research. *J Natl Cancer Inst*. <https://doi.org/10.1093/jnci/djaa154>
8. Hosmer DW, Lemeshow S, May S (2008) Applied survival analysis: regression modeling of time-to-event data, 2nd edn. Wiley-Interscience, Hoboken
9. Cox DR (1972) Regression models and life-tables. *J Roy Stat Soc B* 34(2):187–220
10. Cohen MS, Hanley RS, Kurteva T et al (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
11. Muntener M, Epstein JI, Hernandez DJ et al (2008) Prognostic significance of Gleason score discrepancies between needle biopsy and radical prostatectomy. *Eur Urol* 53(4):767–775 (**discussion 775–766**)
12. Jang WS, Koh DH, Kim J et al (2019) The prognostic impact of downgrading and upgrading from biopsy to radical prostatectomy among men with Gleason score 7 prostate cancer. *Prostate* 79(16):1805–1810
13. Gonzalgo ML, Bastian PJ, Mangold LA et al (2006) Relationship between primary Gleason pattern on needle biopsy and clinicopathologic outcomes among men with Gleason score 7 adenocarcinoma of the prostate. *Urology* 67(1):115–119
14. Altok M, Troncoso P, Achim MF, Matin SF, Gonzalez GN, Davis JW (2019) Prostate cancer upgrading or downgrading of biopsy Gleason scores at radical prostatectomy: prediction of “regression to the mean” using routine clinical features with correlating biochemical relapse rates. *Asian J Androl* 21(6):598–604
15. Chung DY, Lee JS, Goh HJ et al (2019) Predictive factors and oncologic outcome of downgrade to pathologic Gleason score 6–7 after radical prostatectomy in patients with biopsy Gleason score 8–10. *J Clin Med*. <https://doi.org/10.3390/jcm8040438>
16. Donohue JF, Bianco FJ Jr, Kuroiwa K et al (2006) Poorly differentiated prostate cancer treated with radical prostatectomy: long-term outcome and incidence of pathological downgrading. *J Urol* 176(3):991–995
17. Su ZT, Patel HD, Epstein JI, Pavlovich CP, Allaf ME (2020) Downgrading of grade group 2 intermediate-risk prostate cancer from biopsy to radical prostatectomy: comparison of outcomes

- and predictors to identify potential candidates for active surveillance. *Cancer* 126(8):1632–1639
18. Ham WS, Chalfin HJ, Feng Z et al (2017) The impact of downgrading from biopsy Gleason 7 to prostatectomy Gleason 6 on biochemical recurrence and prostate cancer specific mortality. *J Urol* 197(4):1060–1067
 19. Gondo T, Poon BY, Matsumoto K, Bernstein M, Sjoberg DD, Eastham JA (2015) Clinical role of pathological downgrading after radical prostatectomy in patients with biopsy confirmed Gleason score 3 + 4 prostate cancer. *BJU Int* 115(1):81–86
 20. Epstein JI, Feng Z, Trock BJ, Pierorazio PM (2012) 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* 61(5):1019–1024
 21. Chen CC, Huang CP, Tsai YT, Hsieh TF, Shyr CR (2017) The genomic alterations of 5 α -reductases and their inhibitor Finasteride's effect in bladder cancer. *Anticancer Res* 37(12):6893–6898
 22. Whitson JM, Porten SP, Cowan JE, Simko JP, Cooperberg MR, Carroll PR (2013) Factors associated with downgrading in patients with high grade prostate cancer. *Urol Oncol* 31(4):442–447
 23. Treurniet KM, Trudel D, Sykes J, Evans AJ, Finelli A, Van der Kwast TH (2014) Downgrading of biopsy based Gleason score in prostatectomy specimens. *J Clin Pathol* 67(4):313–318
 24. Beksac AT, Sobotka S, Xu P et al (2019) Downgrading of grade group after radical prostatectomy: comparison of multiparametric magnetic resonance imaging guided fusion biopsy and standard 12-core biopsy. *Urology* 127:80–85
 25. Epstein JI (2018) Prostate cancer grading: a decade after the 2005 modified system. *Modern Pathol* 31(1):47–63
 26. Epstein JI, Allsbrook WC Jr, Amin MB, Egevad LL (2005) The 2005 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma. *Am J Surg Pathol* 29(9):1228–1242

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Denzel Zhu¹ · William Shyr¹ · Michelle Toker¹ · Ethan Fram² · Jinrong Cheng^{1,3} · Evan Z. Kovac^{1,2} · Ilir Agalliu^{2,4} · Ahmed Aboumohamed^{1,2} · Kara L. Watts^{1,2}

Denzel Zhu
denzel.zhu@einsteinmed.org

William Shyr
william.shyr@einsteinmed.org

Michelle Toker
michelle.toker@einsteinmed.org

Ethan Fram
efram@montefiore.org

Jinrong Cheng
jcheng@montefiore.org

Evan Z. Kovac
ekovac@montefiore.org

Ilir Agalliu
ilir.agalliu@einsteinmed.org

Ahmed Aboumohamed
aaboumoh@montefiore.org

¹ Albert Einstein College of Medicine, Bronx, NY, USA

² Department of Urology, Montefiore Medical Center, Montefiore Medical Center, Albert Einstein College of Medicine, 1250 Waters Place, Tower 1; Penthouse, Bronx, NY 10461, USA

³ Department of Pathology, Montefiore Medical Center, Bronx, NY, USA

⁴ Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY, USA