Prostate Int 4 (2016) 145–151

Contents lists available at ScienceDirect

Prostate International

journal homepage: http://p-international.com

Original Article

Gleason group concordance between biopsy and radical prostatectomy specimens: A cohort study from Prostate Cancer Outcome Registry – Victoria

Sue M. Evans [1, *], Varuni Patabendi Bandarage [1], Caroline Kronborg [2], Arul Earnest [1], Jeremy Millar [1], David Clouston [3]

[1] School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia
[2] Department of Medicine, The Alfred Hospital, Alfred Health, Melbourne, Australia
[3] Tissupath Pathology Services, Mount Waverley, Melbourne, Australia

ARTICLE INFO

Article history:
Received 10 July 2016
Accepted 27 July 2016
Available online 3 August 2016

Keywords:
Biopsy
Gleason score
Prostate cancer
Prostatectomy
Registry

ABSTRACT

Background: A new prostate cancer (PCa) prognostic grading system [Gleason groups (GGs)] has been proposed based on the contemporary Gleason scores (GSs), which has five simplified prognostic categories. The objective of this study was to evaluate the agreement between the GGs of prostate biopsy and radical prostatectomy specimens and to identify predictive factors for upgrading GGs.

Methods: A total of 5339 cases of RP notified to the Prostate Cancer Outcomes Registry, Victoria, Australia over 6 years (2009–2014) from 46 hospitals, were included. The upgrading was evaluated using the new PCa prognostic grading system, the International Society of Urologic Pathology grade groups, which has five prognostic categories. GG 1 is GS ≤ 6, GG 2 is GS $3+4=7$, GG 3 is GS $4+3=7$, GG 4 is GS 8, and GG 5 is GS 9 and 10. Predictors of upgrading were assessed using univariate and multivariate models.

Results: The GG of prostate biopsies and RP specimens were concordant in 54.5% of cases, while 31.1% were upgraded and 14.3% were downgraded. Longer time interval between biopsy and RP [44–99 days: odds ratio (OR) = 1.3, 95% confidence interval (CI) = 1.1–1.6; > 99 days: OR = 3.0, 95% CI = 2.4–3.8), and RP performed in a metropolitan hospital (biopsy in a regional hospital: OR = 2.2, 95% CI = 1.6–3.2, biopsy in a metropolitan hospital: OR = 1.7, 95% CI = 1.2–2.2) were significant predictors of GG upgrading. Patients who were diagnosed by transperineal biopsy compared to transrectal ultrasound (OR = 0.6, 95% CI = 0.5–0.8) and higher percentage of positive biopsy cassettes (25–62.5%: OR = 0.7, 95% CI = 0.6–0.8, > 62.5: OR = 0.6, 95% CI = 0.5–0.8) were significantly associated with less likelihood of upgrade.

Conclusion: The lack of concordance among hospitals may be attributable to the specialist expertise of the pathologist. Expert review of specimens may help to overcome this discordance. Clinicians should consider clinical parameters and potential limitations of the GG at biopsy when making treatment decisions with regard to PCa.

1. Introduction

Histopathological assessment of biopsy tissue is the mainstay of diagnosing prostate cancer (PCa).[1] The Gleason score (GS) of biopsy material is a key parameter and plays a vital role in diagnostic evaluation, risk stratification, prognostication, and management decisions regarding PCa.[2]

GS upgrading refers to finding a higher grade in the operative specimen at radical prostatectomy (RP) than was seen in the biopsy, and is associated with poorer outcomes.[3] Studies have demonstrated significant histopathological discordance rates up to 62.8%,[4] with inaccurate biopsy specimens more typically undergraded than overgraded when compared with RP.

Many studies have examined variables that may help to predict pathological upgrading of GS from biopsy to RP, including high

* Corresponding author. Monash University, Department of Epidemiology and Preventive Medicine, Level 6 Alfred Centre, Commercial Road, Melbourne 3004, Australia.
E-mail address: sue.evans@monash.edu (S.M. Evans).

http://dx.doi.org/10.1016/j.prnil.2016.07.004

prostate-specific antigen (PSA) level,[5] advanced patient age,[5] the level of pathologist expertise,[6] time from biopsy to surgery,[7] serum testosterone level,[8] treatment with brachytherapy,[9] percentage tumor involvement,[10] prostate size or volume,[1] and number of core biopsies.[11]

A new PCa prognostic grading system has been proposed based on the contemporary GS, which is known as Gleason groups (GGs). It has five simplified prognostic categories that use a scale of 1–5.[12] The new PCa grading system is more accurate in grade stratification than previous systems, and the lowest grade is 1, as compared to 6 in the previous system, with the potential to reduce overtreatment of PCa.[13] There are limited data evaluating this new proposed GG system. One study has been conducted to investigate pathological outcomes using the new GGs,[14] another to verify whether the new GGs yield significant prognostic differences,[13] and one to examine the performance of the new GGs in men with PCa from a nationwide population-based cohort.[15]

In this study, we evaluated the agreement between the GGs of prostate biopsy and RP specimens and identify predictive factors for upgrading the GG in a cohort of men in Victoria, Australia.

2. Materials and methods

2.1. Study population

The Prostate Cancer Outcomes Registry, Victoria (PCOR-Vic, previously known as the Victorian Prostate Cancer Registry; VCR) is a rapid case-ascertainment population-based registry established in 2009 as a means of investigating variation in cancer presentation and care provided to PCa patients in Victoria. Methods for participant recruitment and data collection have previously been described.[16] Men with biopsy-confirmed PCa diagnosis, in participating Victorian hospitals were notified to the PCOR-Vic. Clinical data were collected by trained data collectors through medical records, and histopathological data were captured through hospital information systems and pathology reports and de-identified. The biopsy and RP specimens were performed at a large number of institutions throughout Victoria, including teaching hospitals and private pathology laboratories. Given the wide range of institutions involved, no standardized handling of the surgical specimens was possible, and the specimens were reported by numerous pathologists with no central review.

2.2. Statistical analysis

For our analysis, biopsy GS and RP GS were classified into GGs as described above. GS ≤ 6 was GG 1, GS $3+4=7$ was GG 2, GS $4+3=7$ was GG 3, GS 8 was GG 4, and GS 9 and 10 was GG 5. The grades of the prostate biopsy and RP specimens were considered to be concordant if the GG was the same for the highest grade tumor in the prostate biopsy and the index tumor in the RP. Upgrading was defined as an increase in the GG of the RP specimens compared to the prostate biopsy GG. Cases with GS 10 at diagnosis were excluded, as that category cannot be upgraded, before dividing the patients into concordant and upgrade groups.

Patients' age at diagnosis, preoperative serum PSA level, number of biopsy cassettes, number of positive biopsy cassettes, the time interval between initial biopsy and RP, and RP annual surgeon volume were analyzed as continuous variables. The year of diagnosis, method of diagnosis, clinical categories (cT1, cT2, and cT3/4), percentage of positive biopsy cassettes (< 25%, 25–62.5%, and > 62.5% based on quartiles where percentage of cores positive was operationally defined as the percentage of individually labeled pathological specimens containing PCa of any amount, divided by the total number of individually labeled specimens received), National Comprehensive Cancer Network classification, RP approaches (robot-assisted laparoscopic RP, laparoscopic RP, or open retropubic RP) and the hospitals where the biopsy/RP was performed (private vs. public and metropolitan vs. regional), pathological categories (pT2/pT3/pT4), positive surgical margin status, and extraprostatic extension were analyzed as categorical variables. The annual surgeon volume was calculated by dividing the total number of RP procedures performed by each surgeon over the number of years they have contributed data to the PCOR-Vic. As the PCOR-Vic collects data on ~75% of men diagnosed with PCa, the surgeon volume calculate by PCOR-Vic was compared with that collected by the VCR, which collects surgeon details on all RPs in Victoria. There was no significant difference in annual surgeon volume between PCOR-Vic and VCR.

The differences in these factors in patients with concordant versus upgraded GG were compared using the Mann–Whitney U test for continuous variables and Pearson's χ^2 test for categorical variables. The effect of each of these factors on the odds of GG upgrading was analyzed using the univariate ordered logistic regression model, with odds ratios (ORs) and their corresponding 95% confidence intervals (CIs) provided as measures of effect size. Factors that were significant in the univariate analyses were included in the multivariate logistic regression model using the stepwise method. All statistical analyses were performed using STATA version 13 (StataCorp LP, College Station, TX, USA) and the level of significance was set at 5%.

2.3. Ethical approval

The study was approved by the Health Research Ethics Committee of each participating hospital as well as the Monash University Health Research Ethics Committee (Melbourne, Victoria, Australia; CF09/0931 – 2009000436) and the Cancer Council Victoria (Project No. 0908).

3. Results

Between January 2009 and December 2014, 12,366 PCa patients were included in the PCOR-Vic. Of these, 5,693 patients proceeded to have RP as their primary treatment. Three hundred and sixteen patients were excluded because the GS was missing for the biopsy, RP, or both the biopsy and RP. A further 38 patients who received radiotherapy (3) or androgen therapy (35) prior to RP were excluded as these treatments may affect the histopathology of the RP specimens. A total of 5,339 patients with complete information on paired biopsy and RP specimen histopathology were included in the current study.

Table 1 shows the clinical and demographic information of the patients in the study. The median age at diagnosis was 63.0 years [interquartile range (IQR) = 57.7–67.0]. Most patients (91.3%) were diagnosed via transrectal ultrasound (TRUS)-guided biopsy. RP was performed at a median of 59 days (IQR = 40–99) after diagnosis and 2,821 (52.8%) patients had robot-assisted laparoscopic RP, while 2,156 (40.4%) had open retropubic RP. The majority of the patients (72.2%) had the biopsy and RP performed in a private hospital. The median annual surgeon volume was 33 (IQR = 19–53) and most patients were histologically categorized into pT2 or pT3 categories (93.2%).

Table 2 shows the GG of the RP specimens stratified by biopsy GG. For the GG reported for the RP specimens, 2,911 (54.5%) were unchanged compared to the GG of the diagnostic biopsy (i.e., concordant), while 1,662 (31.1%) of the RP GGs were upgraded and 766 (14.3%) were downgraded (Table 3). In reviewing each grade group, the most frequent upgrading occurred in men with GG 1 on biopsy, who were upgraded in 69.7% of cases, mainly into GG 2,

Table 1
Preoperative and postoperative clinicopathological characteristics of the study population ($N = 5,339$).

Characteristic	*n* (%)
Age at diagnosis (y)	
Median (IQR)	63 (57.7–67)
Y of diagnosis (biopsy)	
2009	467 (8.7)
2010	567 (10.6)
2011	1162 (21.7)
2012	1186 (22.2)
2013	1060 (19.8)
2014	897 (16.8)
Method of diagnosis	
TRUS	4876 (91.3)
TURP	93 (1.7)
Transperineal biopsy	359 (6.7)
Other (TURBT, prostatectomy)	9 (0.2)
Not stated	2 (0.04)
Preoperative serum PSA level (ng/mL)	
Median (IQR)	6 (4.5–8.3)
Clinical categories	
cT1	2,407 (45.1)
cT2	1,606 (30.1)
cT3/cT4	217 (4.0)
Not stated	1,109 (20.7)
No. of biopsy cassettes	
Median (IQR)	8 (6–8)
No. of positive biopsy cassettes	
Median (IQR)	3 (2–5)
Positive biopsy cassettes (%)	
Median (IQR)	42.8 (25–62.5)
< 25	1025 (19.2)
25–62.5	2882 (53.9)
> 62.5	1273 (23.8)
Not stated	159 (2.9)
NCCN Classification	
Low risk	909 (17.0)
Intermediate risk	3007 (56.3)
Metastatic risk	1,134 (21.2)
Not stated	289 (5.4)
Interval between biopsy & surgery (d)	
Median (IQR)	59 (40–99)
< 40	1,327 (24.8)
40–99	2,675 (50.1)
> 99	1,330 (24.9)
Not stated	7 (0.1)
RP approach	
Robot-assisted laparoscopic	2,821 (52.8)
Laparoscopic	317 (5.9)
Open retropubic	2,156 (40.4)
Not stated	45 (0.8)
Hospital where biopsy/RP performed	
Private/private	3,853 (72.2)
Private/public	207 (3.8)
Public/private	189 (3.5)
Public/public	793 (14.8)
Not stated	297 (5.5)
Hospital where biopsy/RP performed	
Metropolitan/metropolitan	4,153 (77.8)
Metropolitan/regional	13 (0.2)
Regional/metropolitan	500 (9.3)
Regional/regional	454 (8.5)
Not stated	219 (4.1)
RP annual surgeon volume	
Median (IQR)	33 (19–53)
Pathological categories	
pT2	2,901 (54.3)
pT3	2,078 (38.9)
pT4	8 (0.1)
Not stated	352 (6.6)
Positive surgical margin	
Absent	3,798 (71.1)
Present	1,449 (27.1)
Not stated	92 (1.7)
Extraprostatic extension	
No	874 (16.3)
Yes	924 (17.3)
Data collected prior to database	3,296 (61.7)
Not stated	245 (4.1)

IQR, interquartile range; NCCN, National Comprehensive Cancer Network; PSA, prostate-specific antigen; RP, radical prostatectomy; TRUS, transrectal ultrasound; TURBT, transurethral resection of a bladder tumor; TURP, transurethral resection of the prostate.

which accounted for 78% of the upgraded cases in this group. Downgrading was more commonly seen in men with GG 3 and GG 4 on biopsy, with downgrading in 29.7% and 60.1%, respectively, in these two groups. Of the five grade groups, GG 1 and GG 4 were the least predictive of the final grade at RP.

Table 4 shows the characteristics of patients who have concordant and upgraded GG. Age at diagnosis, year of diagnosis, method of diagnosis, clinical categories at diagnosis, number and percentage of positive biopsy cassettes, time interval between biopsy and surgery, place where biopsy and surgery were performed, median annual surgeon volume, and status of positive surgical margin were significantly associated with GG upgrading.

Results of univariate and multivariate analysis are shown in Table 5. Higher age at diagnosis was associated with reduced risk of GG upgrading (OR = 0.98, 95% CI = 0.97–0.99). Men diagnosed in more recent years were less likely to be upgraded (2013: OR = 0.7, 95% CI = 0.6–0.9; 2014: OR = 0.5, 95% CI = 0.4–0.7) than those diagnosed in 2009. Patients who were diagnosed by transperineal biopsy were less likely to be upgraded (OR = 0.6, 95% CI = 0.5–0.8). When compared with the total number of cassettes used in the two methods of diagnosis, the median number of cassettes in TRUS was 7.9 and the mean number of cassettes used in transperineal biopsy was 9.7. This difference between the mean values was significant (2 sample Wilcoxon rank sum test = −2.2, $P = 0.02$). Men with a clinical T category of cT3/4 were less likely than those diagnosed with cT1 to have GG upgrading (OR = 0.6, 95% CI = 0.4–0.8). A higher number of positive biopsy cassettes (OR = 0.9, 95% CI = 0.8–0.9) and higher percentage of positive biopsy cassettes was associated with reduced likelihood of GG upgrading. Compared with those with < 25% positive cassettes, 25–62.5% positive cassettes had OR of 0.6 (95% CI = 0.5–0.7), and > 62.5% positive cassettes had OR of 0.5 (95% CI = 0.4–0.6). A longer interval between diagnosis and RP was associated with increased risk of GG upgrading. Compared with < 40 days, 40–99 days had a 1.4 times higher risk of upgrade (OR = 1.4, 95% CI = 1.1–1.6), and > 99 days had a three times higher risk of upgrade (OR = 3.1, 95% CI = 2.6–3.7). Additionally, patients who had both biopsy and RP performed in private hospitals were less likely to have GG upgrading, compared with those who had biopsy and surgery in public hospitals (OR = 0.6, 95% CI = 0.5–0.8). Lastly, patients who had RP in a metropolitan hospital irrespective of the place of biopsy were more likely to have GG upgrading (biopsy at a regional hospital: OR = 2.3, 95% CI = 1.7–3.1, biopsy at a metropolitan hospital: OR = 1.2, 95% CI = 1.0–1.6).

In multivariate analysis, a longer interval between biopsy and RP (44–99 days: OR = 1.3, 95% CI = 1.1–1.6; > 99 days: OR = 3.0, 95% CI = 2.4–3.8), and RP performed in a metropolitan hospital (biopsy in a regional hospital: OR = 2.2, 95% CI = 1.6–3.2; biopsy in a metropolitan hospital: OR = 1.7, 95% CI = 1.2–2.2) remained as significant predictors of GG upgrading. Patients who were

Table 2
Gleason groups of biopsy and RP specimens.

Biopsy Gleason groups	RP specimens Gleason groups, *n* (%)					
	1 (≤ 6)	2 (3 + 4)	3 (4 + 3)	4 (8)	5 (9 & 10)	Total
1 (≤ 6)	438 (30.3)	787 (54.5)	175 (12.1)	15 (1.0)	31 (2.1)	1,446 (100.0)
2 (3 + 4)	71 (3.4)	1545 (74.7)	382 (18.6)	32 (1.5)	38 (1.8)	2,068 (100.0)
3 (4 + 3)	19 (2.1)	255 (27.6)	551 (59.7)	35 (3.8)	63 (6.8)	923 (100.0)
4 (8)	4 (0.7)	83 (16.0)	229 (44.0)	100 (19.2)	104 (20.1)	520 (100.0)
5 (9 & 10)	3 (0.8)	18 (4.7)	63 (16.5)	21 (5.5)	277 (72.5)	382 (100.0)
Total	535 (10.0)	2,688 (50.3)	1400 (26.2)	203 (3.8)	513 (9.7)	5,339 (100.0)

RP, radical prostatectomy.

Table 3
Prostate biopsy and RP specimen GG concordance rate.

	RP specimen GG, *n* (x, y%)					
Biopsy GG	Downgraded *n* (x, y)	Concordant *n* (x, y)	Upgraded by any group *n* (x, y)	Total *n* (x, y)	Upgraded by 1 group *n* (x)	Upgraded by ≥ 2 groups *n* (x)
1 (≤ 6)	0 (0.0, 0.0)	438 (15.0, 30.3)	1008 (60.6, 69.7)	1,446 (27.0, 100.0)	787 (60.2)	221 (62.4)
2 (3 + 4)	71 (9.3, 3.4)	1545 (53.1, 74.7)	452 (27.2, 21.8)	2,068 (38.7, 100.0)	382 (29.2)	70 (19.8)
3 (4 + 3)	274 (35.8, 29.7)	551 (18.9, 59.7)	98 (5.9, 10.6)	923 (17.3, 100.0)	35 (2.7)	63 (17.8)
4 (8)	316 (41.2, 60.8)	100 (3.4, 19.2)	104 (6.3, 20.0)	520 (9.7, 100.0)	104 (7.9)	0
5 (9 & 10)	105 (13.7, 27.5)	277 (9.5, 72.5)	0 (0.0, 0.00)	382 (7.1, 100.0)	0	0
Total	766 (100.0, 14.3)	2911 (100.0, 54.5)	1662 (100.0, 31.1)	5339 (100.0, 100.0)	1308 (100.0)	354 (100.0)

GG, Gleason group; RP, radical prostatectomy.

diagnosed by transperineal biopsy compared with TRUS (OR = 0.6, 95% CI = 0.5–0.8) and higher percentage of positive biopsy cassettes (25–62.5%: OR = 0.7, 95% CI = 0.6–0.8; > 62.5%: OR = 0.6, 95% CI = 0.5–0.8) remained as significant factors for less likelihood of upgrade. Age at diagnosis, year of diagnosis, clinical T categories, and number of positive biopsy cassettes were no longer significantly associated with upgrading.

4. Discussion

This study demonstrates that despite efforts of standardization, inaccuracy of PCa biopsy persists in Victoria with discrepancies in GG between biopsy and RP. This has significant clinical implications for management of PCa; both on an individual and public health level. Concordance between biopsy and RP specimens was seen 54.5% of the time, and if not concordant, GG was more likely to be upgraded. In our cohort, lower GG (GG 1) on biopsy was more likely to be upgraded on RP (60.6%). After multivariate analysis, the predictors of GG upgrading were the time interval between biopsy and RP, the hospital where the RP was performed, method of diagnosis, and percentage of positive biopsy cassettes.

An Australian population-based series of RP undertaken between 1995 and 2000 demonstrated low concordance between prostate biopsy and RP (31%), with 42% of biopsy specimens undergraded.[17] The difference between their average level of concordance and ours (31% vs. 54%) may reflect an improvement of accuracy in Australia over time. However, lack of uniformity in the various grading systems, especially the allocation of histology to grades changed after International Society of Urologic Pathology 2005,[18] and that precludes meaningful comparison between studies.

There are several possible explanations for upgrading of PCa. First, it may reflect interobserver variation among the pathologists. A means of overcoming discrepancy between biopsy and RP GGs would be to have biopsy specimens examined by higher-volume, centrally located pathologists prior to making treatment decisions. Indeed, it has been suggested that a second opinion on prostate biopsy specimens should be mandatory as expert review may lead to a significant difference in score and hence recommended therapy.[19] Second, upgrading may reflect sampling error, where a small sample of PCa obtained on biopsy may not be representative of the cancer as a whole. As such, the grade on biopsy frequently differs from the final GG of the RP specimen; potentially with significant clinical repercussions.[20] It follows that the amount of tumor in the biopsy should be considered when gauging the likely true pathology. In our study, having > 25% of the biopsy specimens positive for cancer was associated with a higher level of concordance between biopsy and RP GG, compared with having < 25% of the specimens positive for cancer. While percentage of positive biopsy cores is a proxy measure of tumor volume, limited literature on the correlation between percentage tumor volume and accuracy of tumor grade exists. Our results are consistent with previous studies that demonstrated an inverse relationship between percentage of positive cores and GG upgrading.[6] The literature suggests this sampling error can be overcome in part by taking more cores, with reports that increased number of cores taken at the time of biopsy achieve improved concordance.[21] This improvement in sampling is possible with transperineal biopsies, and although limited, the data presented here show better concordance of tumor grade between biopsy and RP for transperineal biopsy than TRUS biopsy.

Several studies have investigated predictors of upgrading in the literature with conflicting results; our aim was to strengthen the evidence base. Predictors from the literature include low prostate volume/weight,[3] higher PSA level,[5,22,23] higher PSA density,[24] older age,[5] clinical stage T2,[25] time interval between diagnosis to RP,[7,23] percentage positive biopsy cores,[7,22] Cancer of the Prostate Risk Assessment (CAPRA) score,[22] higher body mass index,[22,24] and low serum testosterone.[8]

Our results demonstrate that a longer interval between diagnostic biopsy and surgery is predictive of upgrading. This is consistent with previous findings.[7,23] Furthermore, we have shown that a delay of > 9 months or even 6 months not only predicted upgrading, but was also associated with greater biochemical recurrence and positive surgical margins,[26] suggesting that a delay may compromise patient outcomes. An alternative explanation would be that the association is confounded by subsequent biopsies after the initial biopsy, which showed more advanced disease, in turn prompting RP.

We found that the method of diagnostic biopsy is a predictor for Gleason upgrading. Patients diagnosed using transperineal biopsy, introduced recently into clinical practice in Victoria, were less likely to be upgraded compared to those undergoing TRUS biopsies.

Table 4
Characteristics of patients with concordant versus upgraded Gleason groups

Characteristic	Concordant ([a])$n = 2,891$)	Upgraded ($n = 1,662$)	P[b)]
Age at diagnosis (y)	62.1	61.5	0.003
Year of diagnosis (biopsy)			
2009	239 (8.3)	164 (9.8)	< 0.001
2010	268 (9.3)	194 (11.7)	
2011	601 (20.8)	364 (21.9)	
2012	633 (21.9)	397 (23.9)	
2013	603 (20.8)	318 (19.1)	
2014	547 (18.9)	225 (13.5)	
Diagnostic method			
TRUS	2620 (90.9)	1535 (92.4)	0.004
TURP	46 (1.6)	39 (2.3)	
Transperineal biopsy	215 (7.5)	87 (5.2)	
Preoperative PSA level (ng/mL)	7.6	7.4	0.5
Clinical categories,			
cT1	1280 (56.6)	814 (60.4)	0.01
cT2	857 (37.9)	485 (35.9)	
cT3/4	125 (5.5)	49 (3.6)	
No. of biopsy cassettes	8.01	8.02	0.8
No. of positive biopsy cassettes	3.5	3.2	< 0.001
Positive biopsy cassettes (%)			
< 25	502 (17.9)	419 (26.1)	< 0.001
25–64.2	1,592 (56.8)	870 (54.1)	
> 64.2	709 (25.3)	318 (19.8)	
Interval between biopsy & surgery (d)			
< 40	796 (27.5)	280 (16.8)	< 0.001
40–99	1,509 (52.3)	744 (44.8)	
> 99	582 (20.2)	638 (38.4)	
Surgical approach			
robot-assisted laparoscopic RP	1,538 (53.8)	916 (55.3)	0.2
Laparoscopic prostatectomy	191 (6.7)	91 (5.5)	
Open prostatectomy	1,127 (39.5)	650 (39.2)	
Hospital where RP performed			
Private/private	2192 (79.8)	1119 (72.3)	< 0.001
Private/public	105 (3.8)	70 (4.5)	
Public/private	80 (2.9)	82 (5.3)	
Public/public	368 (13.4)	276 (17.8)	
Hospital where biopsy/RP performed			
Metropolitan/Metropolitan	2,293 (82.3)	1244 (78.8)	< 0.001
Metropolitan/regional	6 (0.2)	6 (0.3)	
Regional/metropolitan	214 (7.7)	213 (13.5)	
Regional/regional	272 (9.8)	115 (7.3)	
RP median annual surgeon volume	34.5	36.7	0.0009
Pathological T categories			
pT2	1,598 (59.4)	904 (57.8)	0.3
pT3	1,092 (40.6)	659 (42.2)	
Positive surgical margin			
Absent	2,108 (74.2)	1,136 (69.3)	< 0.001
Present	732 (25.7)	502 (30.6)	
Extraprostatic extension			
No	493 (49.5)	280 (48.6)	0.7
Yes	503 (50.5)	296 (51.4)	

Data are presented as n (%).
For continuous variables, median, and for categorical variables, n (%) is shown.
[a)] Patients who had Gleason sum score at diagnosis =10 were excluded in the analysis.
[b)] Numerical-Mann–Whitney U test, categorical – chi square test/fishers exact depending on the number in each cell.
PSA, prostate-specific antigen; RP, radical prostatectomy; TRUS, transrectal ultrasound; TURP, transurethral resection of the prostate.

However, the mean number of biopsy cassettes was higher for the transperineal biopsy method, hence this should be interpreted carefully. Few studies have demonstrated this association. In one recent study,[27] the likelihood of upgrade with transperineal biopsies (30.41%) was less than with TRUS biopsies (33.22%), but the reduction was not statistically significant ($P = 0.55$).

The factor most significantly associated with upgrading related to where biopsies and surgery was undertaken. GG upgrading was significantly more likely to occur if the surgery was carried out in a metropolitan hospital compared with a regional hospital, irrespective of the place of biopsy. Although there is variability in the level and type of services available in regional areas compared to metropolitan (urban and more developed) areas,[28] the reasons for this are uncertain and require further investigation. It may reflect that the skill of the pathologist affected the agreement between the two sets of pathologists. This interobserver variation was confirmed in a previous study in which all biopsies were reviewed by a central panel and a consensus opinion was achieved. It demonstrated that there was a high degree of concordance between the original GS and the consensus scores derived from the central review (28%).[29]

Our data did not support the predictive power of preoperative PSA level in GG upgrading. Variable evidence exists for an association between preoperative PSA and upgrading, with some literature in support of higher PSA levels correlating with upgrading in some,[23] while other studies show that PSA is not helpful in predicting Gleason upgrading.[4]

Table 5
Factors predictive of RP specimen Gleason group upgrading in univariate and multivariate logistic regression[b)]

	Univariate analysis			Multivariate analysis		
	OR	95% CI	*P*	OR	95% CI	*P*
Age at diagnosis	0.98	0.97–0.99	0.008	–		
Y of diagnosis (biopsy)						
2009	Reference					
2010	1.0	0.8–1.3	0.6	–		
2011	0.8	0.6-1.1	0.3			
2012	0.9	0.7–1.1	0.4			
2013	0.7	0.6–0.9	0.03			
2014	0.5	0.4–0.7	< 0.001			
Diagnostic method						
TRUS	Reference					
TURP	1.4	0.9–2.2	0.09			
Transperineal biopsy	0.6	0.5–0.8	0.005	0.6	0.5–0.8	< 0.001
Preoperative serum PSA level (ng/mL)	0.9	0.9–1.0	0.5	a)	–	–
Clinical categories at diagnosis						
cT1	Reference			–	–	–
cT2	0.8	0.7–1.0	0.1			
cT3/4	0.6	0.4–0.8	0.006			
No. of biopsy cassettes	1.0	0.9–1.0	0.8	a)	–	–
No. of positive biopsy cassettes	0.9	0.8–0.9	< 0.001			
Positive biopsy cassettes (%)						
< 25	Reference			Reference		
25–62.5	0.6	0.5–0.7	< 0.001	0.7	0.6–0.8	0.002
> 62.5	0.5	0.4–0.6	< 0.001	0.6	0.5–0.8	< 0.001
Interval between biopsy & surgery (d)						
< 40	Reference					
40–99	1.4	1.1–1.6	< 0.001	1.3	1.1–1.6	0.002
> 99	3.1	2.6–3.7	< 0.001	3.0	2.4–3.8	< 0.001
Surgical approach						
Robot-assisted LRP	Reference			a)		
LRP	0.7	0.6–1.0	0.09	–	–	–
Open retropubic RP	0.9	0.8–1.0	0.6	–	–	–
Hospital where biopsy/RP performed						
Public/public	Reference			Reference		
Public/private	1.3	0.9–1.9	0.07	1.6	1.0–2.3	0.01
Private/public	0.8	0.6–1.2	0.4	0.7	0.5–1.1	0.1
Private/private	0.6	0.5–0.8	< 0.001	0.9	0.7–1.1	0.3
Hospital where biopsy/RP performed						
Regional/regional	Reference					
Regional/metropolitan	2.3	1.7–3.1	< 0.001	2.2	1.6–3.2	< 0.001
Metropolitan/regional	2.3	0.7–7.4	0.1	4.8	0.8–29.7	0.08
Metropolitan/metropolitan	1.2	1.0–1.6	0.03	1.7	1.2–2.2	< 0.001
RP median annual surgeon volume	1.005	1.002–1.008	< 0.001			
Pathological T categories						
pT2	Reference				a)	
pT3	1.0	0.9–1.2	0.3			
Surgical margins						
Absent	Reference				a)	
Present	1.2	1.1–1.4	< 0.001			
Extraperitoneal invasion						
No	Reference				a)	
Yes	1.03	0.8–1.2	0.7			

a) Variables not entered in to the final model.
b) Dependent variable was upgraded 0–4 in an ordered logistic model.
CI, confidence interval; LRP, laparoscopic radical prostatectomy; OR, odds ratio; PSA, prostate-specific antigen; RP, radical prostatectomy; TRUS, transrectal ultrasound; TURP, transurethral resection of the prostate.

The strengths of our study were that the PCOR-Vic is a registry that collects data systematically on PCa; the data are collected by trained staff; and it captures a complete summary of patient history, diagnosis, treatment, and quality of life outcomes of patients diagnosed with PCa in Victoria. Recruitment occurs from hospitals concurrently with cancer notifications to the VCR. All Victorian hospitals, pathology services, and prescribed registers (public or private) are mandated to report cancer diagnosis information to the VCR.[16]

There were several limitations to our study. The new GGs were not entirely a straight transfer of GS to GG but also considered tertiary patterns as well. However, PCOR-Vic does not collect data about the tertiary pattern. In addition, the PCOR-Vic does not collect details of the pathologist reviewing the specimen or how the RP specimens are submitted, and other potential contributing factors such as prostate volume, patient symptoms, body mass index, and other past medical or social history. As such, we cannot determine the extent to which these variables explain the variation or perhaps moderate our existing findings. The PCOR-Vic currently collects data from ~75% of the Victorian population. As such, it is difficult to generalize our findings to regions not represented in the registry.

In conclusion, much international data exists, but a large cohort of Australian men has not been evaluated with regard to concordance and predictive factors. Our data suggest the need for measures to enhance consistency of pathological assessment of prostate

biopsies due to the high rate of discordance and the importance of GG as a parameter in risk stratification and management decision-making. Measures might include central review, or enhancing or supporting pathological reporting in regional Victoria. It is important for the clinician and patient to consider potential limitations of the GG at biopsy and how well it represents the true pathology when making treatment decisions.

Conflicts of interest

All authors have no conflict of interest to declare.

Acknowledgments

We would like to thank the participating clinicians and data collectors for their valuable contribution to the PCOR-Vic.

This project has been funded by Movember, the Australian Prostate Cancer Research Centre and Cancer Australia (Priority-driven Collaborative Cancer Research Scheme APP 1010384).

References

1. Moon SJ, Park SY, Lee TY. Predictive factors of Gleason score upgrading in localized and locally advanced prostate cancer diagnosed by prostate biopsy. Kor J Urol 2010;51:677–82.
2. Amin M, Boccon-Gibod L, Egevad L, Epstein JI, Humphrey PA, Mikuz G, et al. Prognostic and predictive factors and reporting of prostate carcinoma in prostate needle biopsy specimens. Scand J Urol Nephrol Suppl 2005;216:20–33.
3. Sarici H, Telli O, Yigitbasi O, Ekici M, Ozgur BC, Yuceturk CN, et al. Predictors of Gleason score upgrading in patients with prostate biopsy Gleason score </=6. Can Urol Assoc J 2014;8:E342–6.
4. Djavan B, Kadesky K, Klopukh B, Marberger M, Roehrborn CG. Gleason scores from prostate biopsies obtained with 18-gauge biopsy needles poorly predict Gleason scores of radical prostatectomy specimens. Eur Urol 1998;33:261–70.
5. Caster JM, Falchook AD, Hendrix LH, Chen RC. Risk of pathologic upgrading or locally advanced disease in early prostate cancer patients based on biopsy Gleason score and PSA: a population-based study of modern patients. Int J Radiat Oncol Biol Phys 2015;92:244–51.
6. Dong F, Jones JS, Stephenson AJ, Magi-Galluzzi C, Reuther AM, Klein EA. Prostate cancer volume at biopsy predicts clinically significant upgrading. J Urol 2008;179:896–900.
7. Eroglu M, Doluoglu OG, Sarici H, Telli O, Ozgur BC, Bozkurt S. Does the time from biopsy to radical prostatectomy affect Gleason score upgrading in patients with clinical t1c prostate cancer? Kor J Urol 2014;55:395–9.
8. Pichon A, Neuzillet Y, Botto H, Raynaud JP, Radulescu C, Molinie V, et al. Preoperative low serum testosterone is associated with high-grade prostate cancer and an increased Gleason score upgrading. Prostate Cancer Prostatic Dis 2015;18:382–7.
9. Bowes D, Crook JM, Wallace K, Evans A, Toi A, Finelli A, et al. Results of a surgically derived nomogram to predict Gleason score upgrading applied to a cohort of patients with “favorable-risk” prostate cancer treated with permanent seed brachytherapy. Urology 2012;80:649–55.
10. Fu Q, Moul JW, Banez LL, Sun L, Mouraviev V, Xie D, et al. Association between percentage of tumor involvement and Gleason score upgrading in low-risk prostate cancer. Med Oncol 2012;29:3339–44.
11. Turley RS, Terris MK, Kane CJ, Aronson WJ, Presti Jr JC, Amling CL, et al. The association between prostate size and Gleason score upgrading depends on the number of biopsy cores obtained: results from the Shared Equal Access Regional Cancer Hospital Database. BJU Int 2008;102:1074–9.
12. Carter HB, Partin AW, Walsh PC, Trock BJ, Veltri RW, Nelson WG, et al. Gleason score 6 adenocarcinoma: should it be labeled as cancer? J Clin Oncol 2012;30:4294–6.
13. Epstein JI, Zelefsky MJ, Sjoberg DD, Nelson JB, Egevad L, Magi-Galluzzi C, et al. A contemporary prostate cancer grading system: a validated alternative to the Gleason Score. Eur Urol 2016;69:428–35.
14. 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.
15. Loeb S, Folkvaljon Y, Robinson D, Lissbrant IF, Egevad L, Stattin P. Evaluation of the 2015 Gleason Grade Groups in a nationwide population-based cohort. Eur Urol 2016;69:1135–41.
16. Evans SM, Millar JL, Wood JM, Davis ID, Bolton D, Giles GG, et al. The Prostate Cancer Registry: monitoring patterns and quality of care for men diagnosed with prostate cancer. BJU Int 2013;111:E158–66.
17. Bolton D, Severi G, Millar JL, Kelsall H, Davidson AJ, Smith C, et al. A whole of population-based series of radical prostatectomy in Victoria, 1995 to 2000. Aust N Z J Public Health 2009;33:527–33.
18. Egevad L, Mazzucchelli R, Montironi R. Implications of the International Society of Urological Pathology modified Gleason grading system. Arch Pathol Lab Med 2012;136:426–34.
19. Brimo F, Schultz L, Epstein JI. The value of mandatory second opinion pathology review of prostate needle biopsy interpretation before radical prostatectomy. J Urol 2010;184:126–30.
20. Boorjian SA, Karnes RJ, Crispen PL, Carlson RE, Rangel LJ, Bergstralh EJ, et al. The impact of positive surgical margins on mortality following radical prostatectomy during the prostate specific antigen era. J Urol 2010;183:1003–9.
21. Yang CW, Lin TP, Huang YH, Chung HJ, Kuo JY, Huang WJ, et al. Does extended prostate needle biopsy improve the concordance of Gleason scores between biopsy and prostatectomy in the Taiwanese population? J Chin Med Assoc 2012;75:97–101.
22. Vora A, Large T, Aronica J, Haynes S, Harbin A, Marchalik D, et al. Predictors of Gleason score upgrading in a large African-American population. Int Urol Nephrol 2013;45:1257–62.
23. Kvale R, Moller B, Wahlqvist R, Fossa SD, Berner A, Busch C, et al. Concordance between Gleason scores of needle biopsies and radical prostatectomy specimens: a population-based study. BJU Int 2009;103:1647–54.
24. Truong M, Slezak JA, Lin CP, Iremashvili V, Sado M, Razmaria AA, et al. Development and multi-institutional validation of an upgrading risk tool for Gleason 6 prostate cancer. Cancer 2013;119:3992–4002.
25. Jain S, Loblaw A, Vesprini D, Zhang L, Kattan MW, Mamedov A, et al. Gleason upgrading with time in a large prostate cancer active surveillance cohort. J Urol 2015;194:79–84.
26. Abern MR, Aronson WJ, Terris MK, Kane CJ, Presti Jr JC, Amling CL, et al. Delayed radical prostatectomy for intermediate-risk prostate cancer is associated with biochemical recurrence: possible implications for active surveillance from the SEARCH database. Prostate 2013;73:409–17.
27. Scott S, Samaratunga H, Chabert C, Breckenridge M, Gianduzzo T. Is transperineal prostate biopsy more accurate than transrectal biopsy in determining final Gleason score and clinical risk category? A comparative analysis. BJU Int 2015;116:26–30.
28. Department of Health Victoria. *Rural and regional health plan technical paper update*. Melbourne, Victoria: Department of Health; 2014.
29. Salmo EN. An audit of inter-observer variability in Gleason grading of prostate cancer biopsies: The experience of central pathology review in the North West of England. Integr Cancer Sci Ther 2015;2:104–6.