available at www.sciencedirect.com
journal homepage: www.europeanurology.com

European Association of Urology

Platinum Priority – Prostate Cancer

Editorial by Rodolfo Montironi, Liang Cheng, Alessia Cimadamore and Antonio Lopez-Beltran on pp. 32–34 of this issue

How Are Gleason Scores Categorized in the Current Literature: An Analysis and Comparison of Articles Published in 2016–2017

Amy G. Zhou [a], *Daniela C. Salles* [a], *Iryna V. Samarska* [a], *Jonathan I. Epstein* [a,b,c,*]

[a] *Department of Pathology, The Johns Hopkins Medical Institutions, Baltimore, MD, USA;* [b] *Department of Urology, The Johns Hopkins Medical Institutions, Baltimore, MD, USA;* [c] *Department of Oncology, The Johns Hopkins Medical Institutions, Baltimore, MD, USA*

Article info

Article history:
Accepted July 11, 2018

Associate Editor:
Giacomo Novara

Keywords:
Prostate Adenocarcinoma
Grade Groups
Gleason Score

Abstract

Background: A new prostate cancer grading system was proposed in 2013 and endorsed by major journals and societies in 2014, in part because of anecdotal evidence that Gleason scores (GSs) were incorrectly combined in the literature.
Objective: To examine how published studies categorized GSs in current practice.
Design, setting, and participants: A PubMed search was conducted on articles published in 2016–2017 using the search terms "Gleason" and "prostate". This literature review included 1576 articles after exclusions.
Results: (1) Separating GS 7: pathology journals were more likely than non-pathology journals to grade GS 7 separately (56.9% vs 40.0%, $p < 0.05$). Articles co-authored by a pathologist separated GS 7 more than those without a pathologist (53.2% vs 32.9%, $p < 0.001$). North American and European studies separated GS 7 more than Asian studies (47.6% and 44.1% vs 24.1%, $p < 0.001$). Clinical articles separated GS 7 more than research articles (43.7% vs 32.9%, $p < 0.001$). (2) Separating GS 8 from GS 9–10: pathology journals separated GS 8 from GS 9–10 more than non-pathology journals (55.2% vs 34.4%, $p = 0.001$). Articles co-authored by a pathologist separated GS 8 from GS 9–10 more often than those without a pathologist (44.9% vs 29.5%, $p < 0.001$). (3) Using grade groups as "ideal" with all other groupings "non-ideal": pathology journals used ideal more than non-pathology journals (32.2% vs 15.9%, $p < 0.001$). Ideal grouping is more likely in articles co-authored by a pathologist than in those without a pathologist (20.6% vs 11.0%, $p < 0.001$). North American and European studies used ideal grouping more than Asian studies (17.6% and 14.0% vs 9.1%, $p < 0.05$). (4) Arranging groupings in decreasing order from ideal to non-ideal: pathology journals were closer to ideal than non-pathology journals ($p = 0.002$). Articles co-authored by a pathologist were classified closer to ideal than those without a pathologist ($p < 0.001$). North American ($p < 0.001$) and European ($p = 0.02$) studies were closer to ideal than Asian studies.
Conclusions: There is still wide variation in how GSs are grouped world-wide. Only a minority of published articles group GSs accurately.
Patient summary: In this report, we looked at how GSs were grouped world-wide. We found that only a minority of published articles on prostate cancer were grouping GSs accurately, which could lead to inaccurate results and affect patient care with different prostate cancer grades. Our study calls for more widespread adoption of the new prostate cancer grading system composed of five grade groups to minimize incorrect grouping for future studies.

* Corresponding author. Department of Pathology, The Johns Hopkins Hospital, 401 N Broadway, Room 2242, Baltimore, MD 21231, USA. Tel. +1 410 955 5043; Fax: +1 443 287 3818.
E-mail address: jepstein@jhmi.edu (J.I. Epstein).

https://doi.org/10.1016/j.eururo.2018.07.021
0302-2838/

1. Introduction

A new classification system for prostate cancer was proposed in 2013 based on data from the Johns Hopkins Hospital [1] and validated in a multi-institutional study on 20 845 radical prostatectomy cases [2]. Compared with the commonly used D'Amico/National Comprehensive Cancer Network (NCCN) Risk Classification System which categorizes Gleason scores (GSs) into ≤6, 7 (not discriminating between 3 + 4 and 4 + 3), and 8–10, the new grade groups (GGs) are a more accurate predictor of progression [2]. The new system is composed of five prognostic GGs and is based on previously modified GSs [3]: GG 1 = GS ≤ 6; GG 2 = GS 3 + 4 = 7; GG 3 = GS 4 + 3 = 7; GG 4 = GS 4 + 4; and GG 5 = GS 9 and 10.

At the November 2014 meeting of the International Society of Urologic Pathologists (ISUP), there was a broad consensus (90%) to adopt this new system [4], which has also been accepted by the College of American Pathologists (CAP) and the World Health Organization (WHO) for the 2016 edition of *Pathology and Genetics: Tumors of the Urinary System and Male Genital Organs*. More recent studies show that GGs predict biochemical recurrence, clinical stage, metastases, and cancer-specific mortality [5–9]. Although there has been confusion in the terminology used for the new system, including "Grading Groups," "Grade Groups," "ISUP Groups," "ISUP/WHO grade," and "Epstein's Grade/Score," the recommended nomenclature by the CAP, WHO, and 2018 AJCC Cancer Staging Manual is "Grade Groups (GGs)."

The purpose of this study is to examine how published studies in the literature are currently categorizing Gleason grades, after the 2014 ISUP recommendation to classify prostate cancer according to five GGs. Is there still a problem with how prostate cancer grading is classified, justifying the need for more widespread adoption of the five-tiered GG system? The secondary goal of this study is to evaluate differences in reporting Gleason grades by journal type, author nationality, pathologist co-author, and clinical or research-oriented article.

2. Material and methods

A PubMed search was performed to include all studies involving prostate cancer published from January 1, 2016, to December 31, 2017, using the search terms "prostate" and "Gleason." We excluded review articles, case reports, meta-analyses, non-English articles, articles for which full text was not available, and articles for which a breakdown of Gleason grade was not given. We recorded the way Gleason groups were categorized into journal type (urology, pathology, radiology, radiation oncology, oncology), author nationality by continent, whether a pathologist was included as co-author, and type of article (clinical vs research).

For part of the analysis, we analyzed cases where the complete spectrum of GSs (from ≤6 to 10) was described in the study. Publications that only considered a subset of GSs were excluded from this analysis. We defined the five GGs as the "ideal" classification, while all other grouping systems were considered "non-ideal." We then separated the groupings into five categories and arranged them from ideal to most non-ideal: (1) ≤6, 3 + 4, 4 + 3, 8, 9–10 (ideal); (2) ≤6, 3 + 4, 4 + 3, 8–10; (3) ≤6, 7, 8, 9, 10; (4) ≤6, 7, 8–10; (5) ≤7, 8–10 or ≤6, 7–10. Cases that did not fit into the above categories were not analyzed.

Statistical analysis was performed using Stata software and the chi-square test. A p value <0.05 was considered statistically significant.

3. Results

The initial PubMed search produced 1183 articles published in 2016 and 1084 articles published in 2017. After all the aforementioned exclusions were made, a total of 822 articles from 2016 and 754 articles from 2017 were included in the final analysis. The breakdown of data collected from all the articles is presented in Table 1. The most common groupings of Gleason grades are listed in Table 2.

3.1. GS 7

First, we examined how often GS 3 + 4 and GS 4 + 3 were considered separately or combined as GS 7. This information was available for 1516/1576 (96.2%) cases. In 2016, GS 3 + 4 and GS 4 + 3 were reported separately in 300/788 (38.1%) cases compared with 318/728 (43.7%) cases in 2017, $p = 0.03$ (Table 3).

Detailed analysis is presented in Table 4. By journal type, radiology (82/150, 54.7%), pathology (37/65, 56.9%), and urology (313/685, 45.7%) journals were significantly more likely to consider these groups separately than oncology journals (103/338, 30.5%, $p < 0.001$). Pathology journals (37/65, 56.9%) were more likely to separate GS 3 + 4 from GS 4 + 3 than non-pathology journals (581/1451, 40.0%, $p = 0.007$). Radiology journals (82/150, 54.7%) also outperformed urology (313/685, 45.7%, $p = 0.05$) and radiation oncology journals (19/55, 34.5%, $p = 0.01$). When a pathologist was included as co-author, GS 4 + 3 and GS 3 + 4 were considered separately in 53.2% (306/575) of cases versus 32.9% (301/916) when no pathologist was included ($p < 0.001$). North American (243/511, 47.6%) and European studies (199/451, 44.1%)

Table 1 – Data collected from articles by year

	Number of articles, *n* (%)	
	2016 (*N* = 822)	2017 (*N* = 754)
Type of journal		
Urology	393 (47.8)	325 (43.1)
Oncology	177 (21.5)	171 (22.6)
Pathology	29 (3.5)	38 (5.0)
Radiology	81 (9.9)	72 (9.6)
Radiation oncology	32 (3.9)	29 (3.9)
Other medical	42 (5.1)	50 (6.6)
Other surgical	5 (0.6)	8 (1.1)
Miscellaneous	63 (7.7)	61 (8.1)
Nationality		
North American	292 (35.5)	250 (33.2)
South/Central American	13 (1.6)	9 (1.2)
European	248 (30.2)	212 (28.1)
Asian	151 (18.4)	175 (23.2)
African	4 (0.5)	1 (0.1)
Multinational	99 (12.0)	96 (12.7)
Australian	15 (1.8)	11 (1.46)
Article type		
Clinical	576 (70.1)	579 (76.8)
Research	246 (29.9)	175 (23.2)
Pathologist co-author	2016 (*N* = 807)	2017 (*N* = 743)
No	477 (59.1)	477 (64.2)
Yes	330 (40.9)	266 (35.8)

Table 2 – Common Gleason score reporting formats

Gleason score reporting format	Cases, *n* (%)	
	2016 (*N* = 822)	2017 (*N* = 754)
≤6, 7, 8–10	219 (26.6)	161 (21.3)
≤6, 3 + 4 = 7, 4 + 3 = 7, 8–10	108 (13.1)	79 (10.5)
≤6, 7–10	101 (12.3)	85 (11.3)
≤6, 3 + 4 = 7, 4 + 3 = 7, 8, 9–10	71 (8.6)	132 (17.5)
≤7, 8–10	58 (7.1)	46 (6.1)
≤6, 7, 8, 9, 10	44 (5.3%)	38 (5.4%)
≤6, 3 + 4, 4 + 3, 4 + 4, 4 + 5, 5 + 4, 5 + 5	25 (3.0)	6 (0.8)
Other	196 (23.8)	207 (27.5)

Table 3 – Gleason score classification from all cases by year (chi-square test)

Classification	Cases, *n* (%)	Cases, *n* (%)	*p* value
GS 7	2016 (*N* = 788)	2017 (*N* = 728)	
Separate	300 (38.1)	318 (43.7)	0.03
Combined	488 (61.9)	410 (56.3)	
GS 8 vs GS 9–10	2016 (*N* = 601)	2017 (*N* = 684)	
Separate	196 (32.6)	258 (37.7)	0.056
Combined	405 (67.4)	426 (62.3)	
Complete spectrum	2016 (*N* = 713)	2017 (*N* = 680)	
Ideal	71 (10.0)	132 (19.4)	<0.001
Non-ideal	642 (90.0)	548 (80.6)	
Order to ideal	2016 (*N* = 601)	2017 (*N* = 541)	
<6, 3 + 4 = 7, 4 + 3 = 7, 8, 9–10	71 (11.8)	132 (34.4)	<0.001
<6, 3 + 4 = 7, 4 + 3 = 7, 8–10	108 (18.0)	79 (14.6)	
≤6, 7, 8, 9, 10	44 (7.3)	38 (70)	
<6, 7, 8–10	219 (36.4)	161 (29.8)	
<7, 8–10 or <6, 7–10	159 (26.46)	131 (24.2)	

GS = Gleason score.

separated GS 3 + 4 and GS 4 + 3 more often than Asian studies (77/320, 24.1%, $p < 0.001$). Clinical articles were more likely to separate the groups of GS 7 (483/1106, 43.7%) than research articles (135/410, 32.9%, $p < 0.001$).

3.2. GS 8

Next, we examined whether GS 8 was reported separately from or combined with GS 9 and 10. Of 1285 articles in which this distinction could be made, 196/601 (32.6%) cases in 2016 and 258/684 (37.7%) cases in 2017 separated GS 8 from GS 9–10 (Table 3). Further analysis (Table 5) showed that pathology journals (32/58, 55.2%) more often separated GS 8 from 9 and 10 than non-pathology journals (422/1227, 34.4%, $p = 0.001$). Radiology journals (62/150, 51.7%) also outperformed oncology (82/291, 28.2%, $p < 0.001$), urology (197/574, 34.3%, $p < 0.001$), and radiation oncology (16/48, 33.3%, $p = 0.03$) journals. Articles with a pathologist co-author more often separated GS 8 from GS 9 and 10 (216/481, 44.9%) than those without a pathologist (230/780, 29.5%, $p < 0.001$). Nationality did not show significant differences in this analysis. Clinical articles (342/919, 37.2%) versus research articles (112/366, 30.1%) also showed a significant difference ($p = 0.02$).

3.3. Ideal grouping

A complete spectrum of GSs was reported in 713 cases in 2016 and 680 cases in 2017 (Table 3). In 2016, 71/713 (10.0%) used the Gleason GGs, while in 2017, 132/680 (19.4%) used GGs ($p < 0.001$). A detailed analysis is presented in Table 6. By journal type, pathology journals (19/59, 32.2%) were more likely to use GGs than non-pathology journals (183/1150, 15.9%, $p < 0.001$). Additionally, radiology journals (25/132, 18.9%) more often utilized GGs than radiation oncology journals (3/50, 6.0%, $p = 0.03$). When a pathologist was included in the study, ideal grouping was used more often (107/519, 20.6%) than when no pathologist was included (93/847, 11.0%, $p < 0.001$). North American (80/454, 17.6%, $p = 0.001$) and European studies (58/413, 14.0%, $p = 0.04$) were more likely to use the ideal grouping than Asian studies (28/308, 9.1%). Clinical articles (153/999, 15.3%) versus research articles (49/393, 12.5%) showed no significant difference ($p = 0.2$).

Table 4 – Reporting GS 7 as separate 3 + 4/4 + 3 (chi-square test)

	n/*N* (%)		*n*/*N* (%)	*p* value
Journal type				
Oncology	103/338 (30.5)	Urology	313/685 (45.7)	<0.001
		Pathology	37/65 (56.9)	<0.001
		Radiology	82/150 (54.7)	<0.001
		Rad-Onc	19/55 (34.5)	0.5
Radiology	82/150 (54.7)	Urology	313/685 (45.7)	0.05
		Pathology	37/65 (56.9)	0.8
		Rad-Onc	19/55 (34.5)	0.01
Pathology	37/65 (56.9)	Urology	313/685 (45.7)	0.08
		Rad-Onc	19/55 (34.5)	0.01
		Non-pathology	581/1451 (40.0)	0.007
Urology	313/685 (45.7)	Rad-Onc	19/55 (34.5)	0.11
Nationality				
Asian	77/320 (24.1)	North American	243/511 (47.6)	<0.001
		European	199/451 (44.1)	<0.001
North American	243/511 (47.6)	European	199/451 (44.1)	0.3
Other				
Pathologist	306/575 (53.2)	No pathologist	301/916 (32.9)	<0.001
Clinical	483/1106 (43.7)	Research	135/410 (32.9)	<0.001

GS = Gleason score; Rad-Onc = radiation oncology.

Table 5 – Reporting GS 8–10 as separate GS 8 versus GS 9–10 (chi-square test)

	n/*N* (%)		*n*/*N* (%)	*p* value
Journal type				
Oncology	82/291 (28.2)	Urology	197/574 (34.3)	0.07
		Pathology	32/58 (55.2)	<0.001
		Radiology	62/120 (51.7)	<0.001
		Rad-Onc	16/48 (33.3)	0.5
Radiology	62/120 (51.7)	Urology	197/574 (34.3)	<0.001
		Pathology	32/58 (55.2)	0.7
		Rad-Onc	16/48 (33.3)	0.03
Pathology	32/58 (55.2)	Urology	197/574 (34.3)	0.002
		Rad-Onc	16/48 (33.3)	0.02
		Non-pathology	422/1227 (34.4)	0.001
Urology	197/574 (34.3)	Rad-Onc	16/48 (33.3)	0.9
Nationality				
Asian	86/290 (29.7)	North American	160/421 (38.0)	0.02
		European	132/374 (35.3)	0.1
North American	160/421 (38.0)	European	132/374 (35.3)	0.4
Other				
Pathologist	216/481 (44.9)	No pathologist	230/780 (29.5)	<0.001
Clinical	342/919 (37.2)	Research	112/366 (30.1)	0.02

GS = Gleason score; Rad-Onc = radiation oncology.

Table 6 – Reporting Gleason grades as ideal (chi-square test)

	n/*N* (%)		*n*/*N* (%)	*p* value
Journal type				
Oncology	29/314 (9.3)	Urology	103/626 (16.5)	0.003
		Pathology	19/59 (32.2)	<0.001
		Radiology	25/132 (18.9)	0.004
		Rad-Onc	3/50 (6.0)	0.5
Radiology	25/132 (18.9)	Urology	103/626 (16.5)	0.5
		Pathology	19/59 (32.2)	0.04
		Rad-Onc	3/50 (6.0)	0.03
Pathology	19/59 (32.2)	Urology	103/626 (16.5)	0.003
		Rad-Onc	3/50 (6.0)	0.001
		Non-pathology	183/1150 (15.9)	<0.001
Urology	103/626 (16.5)	Rad-Onc	3/50 (6.0)	0.05
Nationality				
Asian	28/308 (9.1)	North American	80/454 (17.6)	0.001
		European	58/413 (14.0)	0.04
North American	80/454 (17.6)	European	58/413 (14.0)	0.1
Other				
Pathologist	107/519 (20.6)	No pathologist	93/847 (11.0)	<0.001
Clinical	153/999 (15.3)	Research	49/393 (12.5)	0.2

Rad-Onc = radiation oncology.

3.4. *Range to ideal grouping*

A total of 601 cases in 2016 and 541 cases in 2017 fit into one of the five categories defined in the methods section as range to ideal grouping (Table 3). The detailed analysis by journal type, pathologist co-author, author nationality, and clinical versus research article is shown in Table 7. Pathology journals were closer to ideal than non-pathology journals ($p = 0.002$). Urology ($p = 0.003$), pathology ($p < 0.001$), and radiology journals ($p = 0.001$) were closer to ideal than oncology journals. Radiology journals also outperformed urology ($p = 0.01$) and radiation oncology journals ($p = 0.02$). When a pathologist was included, cases were much more likely to be classified closer to ideal than when there was no pathologist ($p < 0.001$). When comparing nationality, North American ($p < 0.001$) and European studies ($p < 0.001$) were more likely to classify cases closer to the ideal than Asian studies. Clinical articles versus research articles showed no significant difference ($p = 0.3$).

4. Discussion

Donald Gleason, MD, first proposed grading prostate cancer based on its architectural pattern in 1966. Since then, there have been many modifications to the original Gleason system based on new pathologic and clinical developments [3,4]. Since the 2005 modified Gleason grading system [3] was set in place, many cases originally considered Gleason pattern 3 in the original Gleason system are now considered Gleason pattern 4, leading to contemporary GS 6 cancers

Table 7 – Range to ideal grading (chi-square test)

		p value
Journal type		
Oncology	Urology	0.003
	Pathology	<0.001
	Radiology	0.001
	Rad-Onc	0.3
Radiology	Urology	0.006
	Pathology	0.2
	Rad-Onc	0.015
Pathology	Urology	0.03
	Rad-Onc	0.003
	Non-pathology	0.002
Urology	Rad-Onc	0.02
Nationality		
Asian	North American	<0.001
	European	<0.001
North American	European	0.7
Other		
Pathologist	No pathologist	<0.001
Clinical	Research	0.3

Rad-Onc = radiation oncology.

having a better prognosis than before [10]. Since GS 2–5 are no longer assigned, the lowest possible GS is 6. However, patients often believe that they have an intermediate-grade cancer based on a scale of GS 2–10 [4]. GGs arose out of the need for a more accurate and contemporary method of reporting GS.

For decades of both research and patient care (therapy and prognostication), various combinations of Gleason grades have been used to categorize patients into different groups based on risk. The most commonly accepted and widely used system is the D'Amico (NCCN) classification, which divides GSs into three tiers: ≤6, 7, and 8–10 [11]. In our study, the D'Amico classification was used in 380/1393 (27.3%) of articles that included a full spectrum of GSs. However, the biggest problem with the D'Amico classification is oversimplification, combining GS 3 + 4 and 4 + 3 as one group and considering GS 8 and GS 9–10 together.

There is ample evidence that GS 3 + 4 behaves better than GS 4 + 3 [1,2,12–20]. In the early 2000s, publications have shown that the prognosis of GS 3 + 4 is better than that of GS 4 + 3 with a significant difference in recurrence-free survival [14], pathological stage [20], and preoperative prostate-specific antigen [20]. More recent studies report that, compared with GS 3 + 4, GS 4 + 3 has an increased risk of recurrence or progression [12,16], higher prostate cancer–specific mortality [19], higher pathological stage [12,17], and lower recurrence-free survival [1,13]. In our study, only 618/1516 (40.8%) published articles considered GS 3 + 4 and GS 4 + 3 separately. However, an improvement was seen from 2016 (300/788, 38.1%) to 2017 (318/728, 43.7%, $p = 0.03$). The articles that tended to separate GS 3 + 4 and GS 4 + 3 included pathology (37/65, 56.9%) and radiology (82/150, 54.7%) journals, those including a pathologist co-author (306/575, 53.2%), North American (243/511, 47.6%) and European (199/451, 44.1%) studies, and articles with a clinical focus (483/1106, 43.7%).

Similarly, there has been a growing literature supporting the need to categorize GS 8 separately from GS 9–10, as the latter includes Gleason pattern 5. Multiple studies have shown a considerably worse prognosis for GS 9–10 than GS 8 [1,2,21–25]. Compared with GS 8, the risk of progression of GS 9–10 nearly doubles [1] and has increased risk for metastasis [26], higher pathological stage [27], and lower cancer-specific survival [27]. In our study, only 454/1285 cases (35.3%) separated GS 8 from GS 9–10. A borderline significant improvement was seen from 2016 (196/601, 32.6%) to 2017 (258/684, 37.7%, $p = 0.056$). Pathology (32/58, 55.2%) and radiology journals (62/120, 51.7%), articles with a pathologist co-author (216/481, 44.9%), North American studies (160/421, 38.0%), and articles with a clinical focus (342/919, 37.2%) were more likely to separate GS 8 from GS 9–10.

Factoring in the above, GGs clearly offer a more accurate stratification of prostate cancer than other categorizations. However, in our study in 2016, only 71/713 (10%) of published articles used GGs or GS equivalents when describing the full spectrum of GSs. This nearly doubled to 132/680 (19.4%) in 2017. Comparing journal types, pathology journals (19/59, 32.2%) were significantly more likely to use GGs than non-pathology journals (183/150, 15.9%). Articles with a pathologist co-author (107/519, 20.6%) were more likely to use GG categorization than those without (93/847, 11.0%). North American (80/454, 17.6%) and European studies (58/413, 14.0%) were more likely to use GGs than Asian studies (28/308, 9.1%).

When we compared GS classification types along a spectrum of degree away from the ideal (ie, GGs), the results were very similar to those of reporting GS 7. Pathology journals were significantly closer to the ideal groupings than all other journal types, except radiology. Radiology journals were closer to the ideal groupings than oncology, urology, and radiation oncology journals. Studies including a pathologist and North American and European studies used ideal or closer to ideal groupings.

5. Conclusion

Our study shows that there is still wide variation in how GSs are grouped world-wide. Based on articles published in 2016–2017, the most common method in current use is still the D'Amico risk stratification groups (≤6, 7, 8–10) at 380/1393 (27.3%). Only a minority of published articles (203/1393, 14.6%) group GSs accurately. Studies including a pathologist co-author were more likely to use GGs or their GS equivalents for separating GSs, although numbers are still far from expected. Studies based in North America and Europe, and those published in radiology, pathology, and/or urology journals were also more likely to adhere to GGs or their equivalents.

A limitation of this study is that the results may not reflect the proportion of pathology practices using GGs in their clinical/diagnostic reporting. Our data is from publications, which may be driven to use less specific groupings such as the D'Amico classification to summarize data based on small numbers of cases in certain cohorts (ie, GS 9–10). In addition, sometimes depending on the purpose of the study, it may be adequate to combine GGs. For example, a study that examines drug response in patients with metastatic prostate cancer may find combining GS 8–10 to be

sufficient. In 2017, 19.4% of articles used GGs, which is much improved from 10% in 2016. However, it is unlikely that GGs are being used in the vast majority of clinical settings, and considerable room for improvement remains.

GGs or their GS equivalents have not been incorporated into most of the literature. Some of the most cited clinical trials published in high-impact journals have used inappropriate groupings. For example, the following three trials were published in the *New England Journal of Medicine*. The PLCO prostate cancer screening trial (2009) categorized GS into 2–4, 5–6, 7, and 8–10. While meaningless to subdivide GS 2–4 from GS 5–6, GS 3 + 4 and GS 4 + 3 were combined together [28]. The Prostate Cancer Intervention versus Observation trial (2012) classified GS into 2–6 and 7–10, hence combining GS 3 + 4 = 7 with an 88% cure rate with GS 9–10 with a 25% cure rate [2,29]. The ProtectT study group (2016) compared mortality for active monitoring, radical prostatectomy, and external-beam radiotherapy. This study classified GS into 6, 7, and 8–10, lumping together GS 3 + 4 and GS 4 + 3, as well as GS 8 and GS 9–10, despite very different prognoses between these GGs [30].

Current recommendations are to report GGs in conjunction with GS in a case, such as "GS 3 + 3 = 6 (GG 1)" [4]. Our study demonstrated the need for more widespread adoption of GGs in research and to minimize incorrect grouping for future studies.

Author contributions: Jonathan I. Epstein had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Epstein.
Acquisition of data: Zhou, Salles, Samarska.
Analysis and interpretation of data: Zhou, Epstein.
Drafting of the manuscript: Zhou, Epstein.
Critical revision of the manuscript for important intellectual content: Epstein.
Statistical analysis: Zhou, Epstein.
Obtaining funding: None.
Administrative, technical, or material support: Zhou, Epstein.
Supervision: Epstein.
Other (specify): None.

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

Funding/Support and role of the sponsor: None.

References

[1] 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.
[2] Epstein JI, Zelefsky MJ, Sjoberg DD, et al. A contemporary prostate cancer grading system: a validated alternative to the Gleason score. Eur Urol 2016;69:428–35.
[3] Epstein JI, Allsbrook Jr WC, Amin MB, Egevad LL, Committee IG. The 2005 International Society of Urological Pathology (ISUP) consensus conference on Gleason grading of prostatic carcinoma. Am J Surg Pathol 2005;29:1228–42.
[4] Epstein JI, Egevad L, Amin MB, et al. 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 2016;40:244–52.
[5] Schulman AA, Howard LE, Tay KJ, et al. Validation of the 2015 prostate cancer grade groups for predicting long-term oncologic outcomes in a shared equal-access health system. Cancer 2017;123:4122–9.
[6] Leapman MS, Cowan JE, Simko J, et al. Application of a prognostic Gleason grade grouping system to assess distant prostate cancer outcomes. Eur Urol 2017;71:750–9.
[7] 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.
[8] He J, Albertsen PC, Moore D, Rotter D, Demissie K, Lu-Yao G. Validation of a contemporary five-tiered Gleason grade grouping using population-based data. Eur Urol 2017;71:760–3.
[9] Berney DM, Beltran L, Fisher G, et al. Validation of a contemporary prostate cancer grading system using prostate cancer death as outcome. Br J Cancer 2016;114:1078–83.
[10] Kryvenko ON, Epstein JI. Prostate cancer grading: A decade After the 2005 modified Gleason grading system. Arch Pathol Lab Med 2016;140:1140–52.
[11] D'Amico AV, Whittington R, Malkowicz SB, et al. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. JAMA 1998;280:969–74.
[12] Amin A, Partin A, Epstein JI. Gleason score 7 prostate cancer on needle biopsy: relation of primary pattern 3 or 4 to pathological stage and progression after radical prostatectomy. J Urol 2011;186:1286–90.
[13] Burdick MJ, Reddy CA, Ulchaker J, et al. Comparison of biochemical relapse-free survival between primary Gleason score 3 and primary Gleason score 4 for biopsy Gleason score 7 prostate cancer. Int J Radiat Oncol Biol Phys 2009;73:1439–45.
[14] Chan TY, Partin AW, Walsh PC, Epstein JI. Prognostic significance of Gleason score 3 + 4 versus Gleason score 4 + 3 tumor at radical prostatectomy. Urology 2000;56:823–7.
[15] Gonzalgo ML, Bastian PJ, Mangold LA, et al. Relationship between primary Gleason pattern on needle biopsy and clinicopathologic outcomes among men with Gleason score 7 adenocarcinoma of the prostate. Urology 2006;67:115–9.
[16] Kang DE, Fitzsimons NJ, Presti Jr JC, et al. Risk stratification of men with Gleason score 7 to 10 tumors by primary and secondary Gleason score: results from the SEARCH database. Urology 2007;70:277–82.
[17] Koontz BF, Tsivian M, Mouraviev V, et al. Impact of primary Gleason grade on risk stratification for Gleason score 7 prostate cancers. Int J Radiat Oncol Biol Phys 2012;82:200–3.
[18] Sakr WA, Tefilli MV, Grignon DJ, et al. Gleason score 7 prostate cancer: a heterogeneous entity? Correlation with pathologic parameters and disease-free survival. Urology 2000;56:730–4.
[19] Wright JL, Salinas CA, Lin DW, et al. Prostate cancer specific mortality and Gleason 7 disease differences in prostate cancer outcomes between cases with Gleason 4 + 3 and Gleason 3 + 4 tumors in a population based cohort. J Urol 2009;182:2702–7.
[20] Makarov DV, Sanderson H, Partin AW, Epstein JI. Gleason score 7 prostate cancer on needle biopsy: is the prognostic difference in Gleason scores 4 + 3 and 3 + 4 independent of the number of involved cores? J Urol 2002;167:2440–2.
[21] Sabolch A, Feng FY, Daignault-Newton S, et al. Gleason pattern 5 is the greatest risk factor for clinical failure and death from prostate

cancer after dose-escalated radiation therapy and hormonal ablation. Int J Radiat Oncol Biol Phys 2011;81:e351–60.
[22] Stenmark MH, Blas K, Halverson S, Sandler HM, Feng FY, Hamstra DA. Continued benefit to androgen deprivation therapy for prostate cancer patients treated with dose-escalated radiation therapy across multiple definitions of high-risk disease. Int J Radiat Oncol Biol Phys 2011;81:e335–44.
[23] Stone NN, Stone MM, Rosenstein BS, Unger P, Stock RG. Influence of pretreatment and treatment factors on intermediate to long-term outcome after prostate brachytherapy. J Urol 2011;185:495–500.
[24] Sylvester JE, Grimm PD, Wong J, Galbreath RW, Merrick G, Blasko JC. Fifteen-year biochemical relapse-free survival, cause-specific survival, and overall survival following I(125) prostate brachytherapy in clinically localized prostate cancer: Seattle experience. Int J Radiat Oncol Biol Phys 2011;81:376–81.
[25] Stock RG, Cesaretti JA, Stone NN. Disease-specific survival following the brachytherapy management of prostate cancer. Int J Radiat Oncol Biol Phys 2006;64:810–6.
[26] Tsao CK, Gray KP, Nakabayashi M, et al. Patients with biopsy Gleason 9 and 10 prostate cancer have significantly worse outcomes compared to patients with Gleason 8 disease. J Urol 2015;194:91–7.
[27] Ham WS, Chalfin HJ, Feng Z, et al. New prostate cancer grading system predicts long-term survival following surgery for Gleason score 8-10 prostate cancer. Eur Urol 2017;71:907–12.
[28] Andriole GL, Crawford ED, Grubb RL, et al. Mortality results from a randomized prostate-cancer screening trial. N Engl J Med 2009;360:1310–9.
[29] Wilt TJ, Brawer MK, Jones KM, et al. Radical prostatectomy versus observation for localized prostate cancer. N Engl J Med 2012;367:203–13.
[30] Donovan JL, Hamdy FC, Lane JA, et al. Patient-reported outcomes after monitoring, surgery, or radiotherapy for prostate cancer. N Engl J Med 2016;375:1425–37.