Pathology 57 (2025) 293–296

Contents lists available at ScienceDirect

Pathology

journal homepage: www.pathologyjournal.rcpa.edu.au

CONTROVERSIES IN PATHOLOGY

Gleason scores provide more accurate prognostic information than grade groups

Lars Egevad [1,*], Chiara Micoli [2], Brett Delahunt [1,3], Hemamali Samaratunga [4], Hans Garmo [5], Pär Stattin [5,6], Martin Eklund [2]

[1] Department of Oncology-Pathology, Karolinska Institutet, Stockholm, Sweden
[2] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
[3] Malaghan Institute of Medical Research, Wellington, New Zealand
[4] Aquesta Uropathology and University of Queensland, Brisbane, Qld, Australia
[5] Regional Cancer Centre, Uppsala University Hospital, Uppsala, Sweden
[6] Department of Surgical Sciences Uppsala University, Uppsala, Sweden

ARTICLE INFO

Key words:
Gleason grade
mortality
needle biopsy
prostate cancer

SUMMARY

Prostate cancer grade is currently often reported both by Gleason scores and by grouping of the scores into five so-called International Society of Urological Pathology (ISUP) grades (also known as grade groups). Using population-based registry data from 172,112 men diagnosed with prostate cancer on needle biopsy, we recently investigated the outcome of Gleason score 8–10 prostate cancer with death due to prostate cancer and death from any cause as endpoints. There was a prognostic heterogeneity between Gleason scores 3+5, 4+4 and 5+3 (ISUP grade 4) and between Gleason scores 4+5, 5+4 and 5+5 (ISUP grade 5). This heterogeneity was lost when the grades collapsed into ISUP grades 4 and 5, respectively. On the other hand, there was also a prognostic overlap between these ISUP grades. The outcome of Gleason score 5+3 and 4+5 cancers was very similar. The prostate-specific mortality of Gleason scores 5+3 and 4+5 was 0.32 (95% confidence interval 0.27–0.36) and 0.30 (0.29–0.31), respectively, after 5 years and 0.44 (0.39–0.49) and 0.45 (0.44–0.46), respectively, after 10 years. The findings emphasise the importance of reporting the Gleason grades and scores for more accurate prognostic information of highly heterogeneous high-grade prostate cancers. It also questions the clinical value of the current recommendations of grouping of Gleason scores into ISUP grades or grade groups.

The last decade has seen a lively debate as to whether or not reporting of prostate cancer grade should be reduced to five so-called International Society of Urological Pathology (ISUP) grades (also known as grade groups). This grouping was endorsed at an ISUP consensus conference in 2014 and has subsequently been variably adopted.[1] In the fifth edition of the WHO Classification of Tumours, ISUP grade, grade groups and WHO grade are all recommended and accepted nomenclatures for this grouping.[2] It has been argued that each of the five grades is associated with a unique prognosis and that this 5-tier system is easier to understand than the original Gleason system.[3] However, the validity of the grouping has been questioned both in the pathology and urology communities.[4] Specifically, it has been noted that by reducing the Gleason scores into a limited number of categories, prognostic granularity may be lost, and no additional prognostic information is provided.

In two recent studies we have employed population-based registry data from the National Prostate Cancer Register (NPCR) in Sweden, consisting of a large cohort of men with prostate cancer diagnosed on needle biopsy (n=172,112) between the inclusive years 2000–2020, to investigate the outcome of Gleason score 8–10 prostate cancer. In these studies, death due to prostate cancer and death from any cause were taken as the endpoints.[5,6] A problem in many previous studies on this topic has been the small samples of some grade combinations (i.e., Gleason score 3+5, 5+3 and 5+5), causing difficulties in conclusive assessment of their prognostic impact.[7–10] For example, only a total of 6–76 cases in these studies were Gleason scores 3+5 and 5+3, while our study included 2,505 of such tumours (2,085 and 420 with 3+5 or 5+3, respectively).[5] By using a national database with a coverage as high as 96% of all newly diagnosed cancers, we have been able to gather

* Address for correspondence: Department of Oncology-Pathology, Karolinska Institutet Radiumhemmet P1:02, Karolinska University Hospital, 171 76 Stockholm, Sweden.
E-mail address: lars.egevad@ki.se (L. Egevad).

https://doi.org/10.1016/j.pathol.2024.12.633

Available online 17 January 2025

sufficient data on the detailed prognostic relevance of contemporary clinical grading.[11] Since a large number of Swedish pathologists have been trained in prostate pathology by one of the authors (LE), it can be assumed that the reporting to this database is done in a relatively uniform manner.

We found a prognostic heterogeneity between Gleason scores 3+5, 4+4 and 5+3 and between Gleason scores 4+5, 5+4 and 5+5 (Fig. 1). This heterogeneity was lost when the grades were collapsed into ISUP grades 4 and 5, respectively. Not only are the long-term outcomes different within the ISUP grades, there is also an overlap between ISUP grades 4 and 5, as both the prostate-specific and all-cause mortality of Gleason scores 5+3 (ISUP grade 4) and 4+5 (ISUP grade 5) were almost identical (Fig. 1). The prostate-specific mortality of Gleason scores 5+3 and 4+5 was 0.32 (95% confidence interval 0.27–0.36) and 0.30 (0.29–0.31), respectively, after 5 years and 0.44 (0.39–0.49) and 0.45 (0.44–0.46), respectively, after 10 years.[5,6]

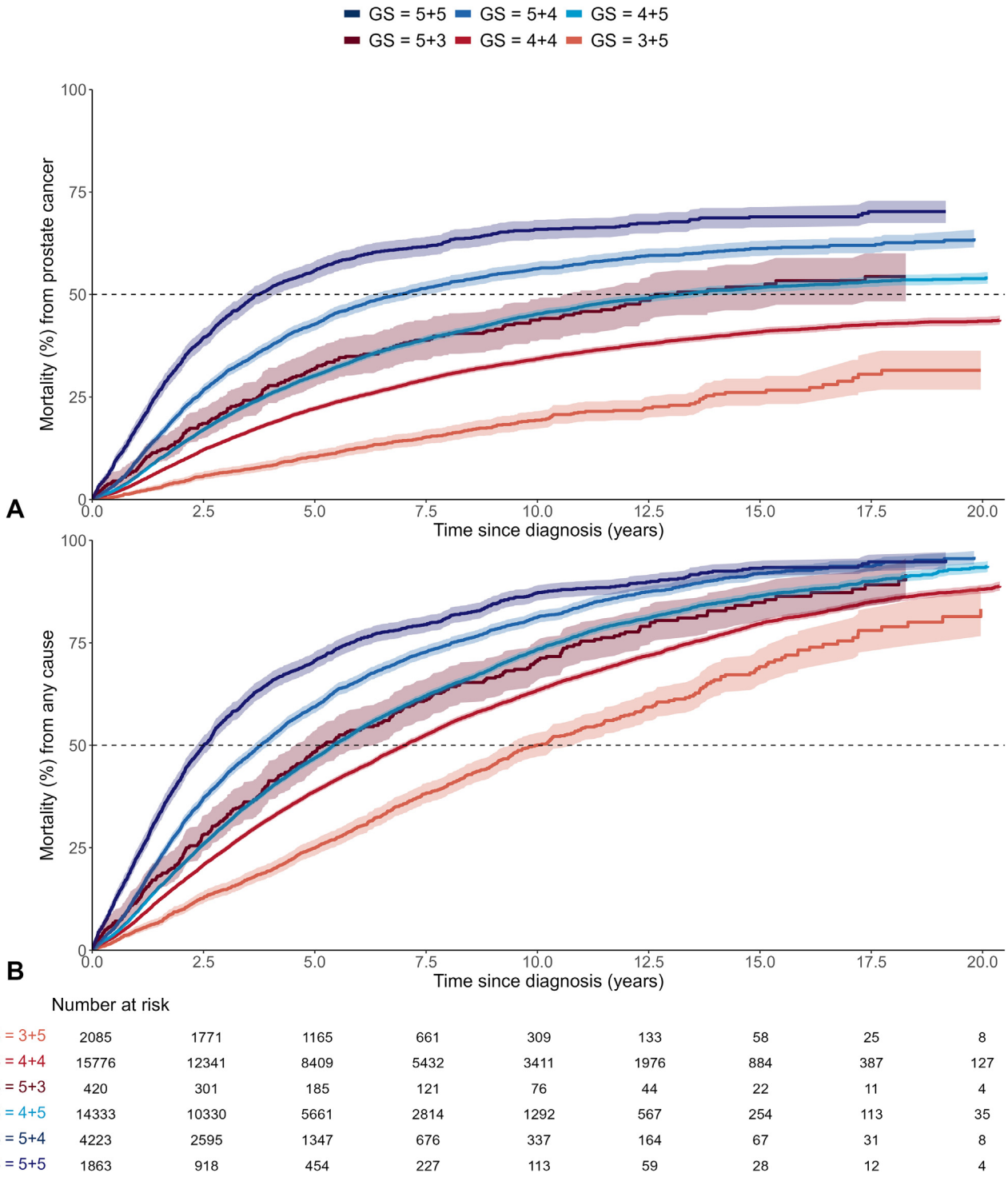

Fig. 1. (A) Prostate cancer-specific mortality (PCSM) of men with Gleason score (GS) 3+5, 4+4, 5+3, 4+5, 5+4 and 5+5 considering death from other causes as a competing risk. (B) All-cause mortality (ACM) of men with Gleason scores of 3+5, 4+4, 5+3, 4+5, 5+4 and 5+5. Curves are truncated at the time point of the last event.

ISUP grade 2 is Gleason score 3+4=7, while ISUP grade 3 is Gleason score 4+3=7, and one of the arguments for the implementation of ISUP grades has been that Gleason score 7 is not a uniform prognostic category. It was recognised in the late 1990s that there was a prognostic heterogeneity among Gleason score 7 cancers, with Gleason score 4+3 cancer having a less favourable outcome than Gleason score 3+4 cancer.[12] When prostate cancer grade is reported using Gleason scores, there is an awareness that Gleason score 4+3 is prognostically different from 3+4. Indeed, it was an early recommendation that not only the Gleason sum be reported, but also the Gleason pattern components.[13] It has also been recommended that Gleason score 3+4=7 tumours be designated Gleason score 7a and Gleason score 4+3=7 tumours be designated Gleason score 7b.[14]

As an extension of these earlier studies, it was demonstrated that the outcome of Gleason score 7 cancer was not only dependent on whether Gleason grade 3 or 4 was the dominant pattern, but also that the percentage of Gleason pattern 4/5 correlated with outcome both after radical prostatectomy[15] and in men on watchful waiting.[16] This has been confirmed by Sauter *et al.* in a cohort of 12,823 men who underwent radical prostatectomy[17] and in men whose tumour was inoperable,[18] where the proportion of Gleason pattern 4 was shown to be of prognostic significance. Thus, there seems to be a prognostic continuum with increasing amounts of Gleason grades 4 and 5; therefore, it may be misleading to assume that the most critical step is from <50% to >50% of such patterns. In a series of men who underwent radical prostatectomy, McNeal *et al.* only found lymph node metastases in cases with at least 3.2 mL of Gleason pattern 4 or 5 cancers, suggesting a volume threshold of high-grade cancer for metastatic capacity.[19] Recently, it was suggested that quantification of Gleason pattern 4 on needle biopsy would be a better predictor of prostatectomy grade and outcome than the Gleason score.[20] Similar to the Gleason score, quantification of Gleason pattern 4 suffers from problems with reproducibility.[21] Artificial intelligence may offer decision support that helps achieving a standardised assessment.[22]

There has been a remarkable shift in the distribution of Gleason scores over the past decades.[23] This has undoubtedly been driven by some decisions taken at the ISUP consensus conferences on prostate cancer grading in 2005 and 2014.[1,24] Most decisive was the recommendation at the ISUP 2005 meeting to always include the highest Gleason grade in the Gleason score of needle biopsies regardless of its volume.[24] This has resulted in a greater proportion of cancers assigned the highest scores. For better prognostic separation of these cases it has become increasingly important to identify all individual Gleason scores, including their Gleason pattern components. It appears from Fig. 1 that the majority of the highest Gleason scores have unique prognostic values. Not only is this separation lost by grouping of these scores into ISUP groups, but the overlap between ISUP grades 4 and 5 will also make the current score grouping system misleading.

In the National Comprehensive Cancer Network (NCCN) guidelines, Gleason scores 8–10 in more than four cores, or the presence of a primary Gleason pattern 5, are used as some of the criteria for moving a localised prostate cancer into the 'very high risk' group.[25,26] The current study indeed justifies the use of primary Gleason pattern 5 as an indicator of poor outcome, although a Gleason score of 4+5=9 portends an equally bad prognosis as a Gleason score of 5+3=8.

A disadvantage of registry studies is the lack of central review. Reviewing individual cases in a large registry study would not be practically feasible. For example, it has been claimed that when biopsies with a Gleason score of 5+3 undergo central review, the vast majority are in fact either downgraded or upgraded.[27] In our series only 0.2% (420/172,112) of all prostate cancers diagnosed on needle biopsies were Gleason score 5+3. It can be assumed that some of the cases of unusual Gleason scores might have been re-graded in a central review. However, the purpose of our study was to assess the prognostic impact of grading as it is currently performed in clinical practice.

It has been claimed that Gleason score 5+3 cancers have a worse outcome than other Gleason score 8 cancers and have a behaviour more similar to Gleason score 9 cancers.[28] We here show that the prognosis of Gleason score 5+3 cancers is very similar to that of Gleason score 4+5 cancers, but there is also a significantly different prostate cancer-specific mortality in Gleason score 5+4 and 5+5 cancers compared to Gleason score 5+3 or 4+5 cancers. Thus, it may not be justified to lump Gleason score 5+3 with all Gleason score 9–10 cancers either. The most appropriate approach seems to be to report not only an ISUP grade but also the full Gleason score, including its patterns.

One of the reasons behind the challenges in grading of prostatic adenocarcinoma is the pronounced grade heterogeneity of this cancer. In more than 50% of radical prostatectomy specimens, at least three Gleason grades are seen.[29] In the past, tertiary patterns of higher grade were reported by 91% of urological pathologists in addition to the Gleason score components of needle biopsies.[30] One of the effects of the ISUP 2005 decision to include such patterns in the Gleason score of biopsies was that the concept of 'tertiary pattern of higher grade' in needle biopsies by definition was eliminated.[24] This has undoubtedly been one of the drivers behind the significant Gleason inflation seen over the past decades.[23] For radical prostatectomy specimens, separate reporting of tertiary patterns of higher grade was still recommended by the 2005 ISUP consensus group,[24] while currently practices regarding reporting of tertiary Gleason pattern 5 components vary.[31] Men with a tertiary pattern of higher grade have a higher tumour stage at radical prostatectomy and progress more often than those without such a pattern.[32,33]

Another characteristic feature of prostate cancer is its striking multifocality. In 87% of radical prostatectomy specimens, two or more foci are reported.[34] In multifocal cancer, tumour foci often have different Gleason scores.[34] This is naturally reflected in a grade heterogeneity in needle biopsies. It has been discussed whether the optimal method of reporting Gleason scores of needle biopsies is to give a composite global score based on all cancer found in the cores or if the radical prostatectomy grade is best reflected by the highest score of an individual core.[35–39] Reporting of the highest Gleason score decreases undergrading, but this comes at the expense of an even more prevalent overgrading.[36] In a series of 2,527 prostate biopsies, the global grade showed a better correlation with the final prostatectomy grade than the highest grade.[38] The two grade measures differed in only 7% of cases, and in those cases the global Gleason score predicted prostatectomy score better than the highest Gleason score (62% and 19%, respectively). In a series of men who underwent hormonal treatment, the global Gleason score, but not the highest Gleason score, was an independent predictor of outcome.[37] In a report from the European Randomised Study of Screening for Prostate Cancer (ERSPC) group, the use of global Gleason score was recommended in order to avoid overtreatment.[39] A survey among pathologists from 22 countries showed that biopsy reporting practices vary among European pathologists, but reporting of a global score was most commonly used (77%).[35] Thus, there is still no consensus as to which of these methods is optimal, but an advantage with the NPCR registry is that it uniformly reports the global Gleason score.

The real strength of the Gleason grading system is that it factors in the grade heterogeneity present in the tumour that is quite characteristic for prostatic adenocarcinoma. We have shown that this complexity is obscured by grouping several of these Gleason grade combinations into five ISUP grades. It would appear that reporting of Gleason scores with an assessment of the proportion of the highest grade present provides superior prognostic information.

Ethical approval

This study was approved by the Research Ethics Authority Dnr 2020–03437.

Conflicts of interest and sources of funding

The authors state that there are no conflicts of interest to disclose. This study was funded by grants from The Swedish Cancer Foundation

(grant no. CAN 20 1358 PjFs and 22 22051 Pj) and The Stockholm Cancer Society (grant no. 204043).

References

1. 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.
2. WHO Classification of Tumours Editorial Board. *Urinary and male genital tumours*. 5th ed. Lyon: IARC; 2022.
3. Epstein JI, Zelefsky MJ, Sjoberg DD, et al. A contemporary prostate cancer grading system: a validated alternative to the Gleason score. *Eur Urol* 2015;**69**(3):428–35.
4. Egevad L, Delahunt B, Bostwick DG, et al. Prostate cancer grading, time to go back to the future. *BJU Int* 2021;**127**:165–8.
5. Egevad L, Micoli C, Delahunt B, et al. Prognosis of Gleason score 8 prostatic adenocarcinoma in needle biopsies: a nationwide population-based study. *Virchows Arch* 2024;**484**:995–1003.
6. Egevad L, Micoli C, Samaratunga H, et al. Prognosis of Gleason score 9-10 prostatic adenocarcinoma in needle biopsies: a nationwide population-based study. *Eur Urol Oncol* 2024;**7**:213–21.
7. Harding-Jackson N, Kryvenko ON, Whittington EE, et al. Outcome of Gleason 3+5=8 prostate cancer diagnosed on needle biopsy: prognostic comparison with Gleason 4+4=8. *J Urol* 2016;**196**:1076–81.
8. Huynh MA, Chen MH, Wu J, et al. Gleason score 3+5 or 5+3 versus 4+4 prostate cancer: the risk of death. *Eur Urol* 2016;**69**:976–9.
9. Lu TC, Moretti K, Beckmann K, et al. ISUP Group 4 - a homogenous group of prostate cancers? *Pathol Oncol Res* 2018;**24**:921–5.
10. van den Bergh RC, van der Kwast TH, de Jong J, et al. Validation of the novel International Society of Urological Pathology 2014 five-tier Gleason grade grouping: biochemical recurrence rates for 3+5 disease may be overestimated. *BJU Int* 2016;**118**:502–5.
11. Rider JR, Sandin F, Andren O, et al. Long-term outcomes among noncuratively treated men according to prostate cancer risk category in a nationwide, population-based study. *Eur Urol* 2013;**63**:88–96.
12. 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.
13. Eble JN, Sauter G, Epstein JI, et al. *Pathology and genetics of tumours of the urinary system and male genital organs*. Lyon: IARC Press; 2004.
14. Helpap B, Egevad L. The significance of modified Gleason grading of prostatic carcinoma in biopsy and radical prostatectomy specimens. *Virchows Arch* 2006;**449**:622–7.
15. Stamey TA, McNeal JE, Yemoto CM, et al. Biological determinants of cancer progression in men with prostate cancer. *JAMA* 1999;**281**:1395–400.
16. Egevad L, Granfors T, Karlberg L, et al. Percent Gleason grade 4/5 as prognostic factor in prostate cancer diagnosed at transurethral resection. *J Urol* 2002;**168**:509–13.
17. Sauter G, Steurer S, Clauditz TS, et al. Clinical utility of quantitative Gleason grading in prostate biopsies and prostatectomy specimens. *Eur Urol* 2016;**69**:592–8.
18. Delahunt B, Steigler A, Atkinson C, et al. Percentage grade 4 tumour predicts outcome for prostate adenocarcinoma in needle biopsies from patients with advanced disease: 10-year data from the TROG 03.04 RADAR trial. *Pathology* 2022;**54**:49–54.
19. McNeal JE, Villers AA, Redwine EA, et al. Histologic differentiation, cancer volume, and pelvic lymph node metastasis in adenocarcinoma of the prostate. *Cancer* 1990;**66**:1225–33.
20. Vickers AJ, Assel M, Cooperberg MR, et al. Amount of Gleason pattern 3 is not predictive of risk in grade group 2-4 prostate cancer. *Eur Urol* 2024;**86**:1–3.
21. Glaessgen A, Hamberg H, Pihl CG, et al. Interobserver reproducibility of percent Gleason grade 4/5 in prostate biopsies. *J Urol* 2004;**171**:664–7.
22. Strom P, Kartasalo K, Olsson H, et al. Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study. *Lancet Oncol* 2020;**21**:222–32.
23. Danneman D, Drevin L, Robinson D, et al. Gleason inflation 1998-2011: a registry study of 97,168 men. *BJU Int* 2015;**115**:248–55.
24. Epstein JI, Allsbrook Jr WC, Amin MB, et al. The 2005 International Society of Urological Pathology (ISUP) consensus conference on gleason grading of prostatic carcinoma. *Am J Surg Pathol* 2005;**29**:1228–42.
25. National Comprehensive Cancer Network. NCCN clinical practice guidelines in oncology - prostate cancer version 4.2024. https://www.nccn.org/professionals/physician_gls/pdf/prostate.pdf [accessed Aug 2024].
26. Schaeffer EM, Srinivas S, Adra N, et al. NCCN guidelines insights: prostate cancer, version 3.2024. *J Natl Compr Canc Netw* 2024;**22**:140–50.
27. Kryvenko ON, Williamson SR, Schwartz LE, et al. Gleason score 5+3=8 (grade group 4) prostate cancer-a rare occurrence with contemporary grading. *Hum Pathol* 2020;**97**:40–51.
28. Zhou Y, Lin C, Hu Z, et al. Differences in survival of prostate cancer Gleason 8-10 disease and the establishment of a new Gleason survival grading system. *Cancer Med* 2021;**10**:87–97.
29. Aihara M, Wheeler TM, Ohori M, et al. Heterogeneity of prostate cancer in radical prostatectomy specimens. *Urology* 1994;**43**:60–6.
30. Egevad L, Allsbrook Jr WC. Epstein JI. Current practice of Gleason grading among genitourinary pathologists. *Hum Pathol* 2005;**36**:5–9.
31. Kench JG, Judge M, Delahunt B, et al. Dataset for the reporting of prostate carcinoma in radical prostatectomy specimens: updated recommendations from the International Collaboration on Cancer Reporting. *Virchows Arch* 2019;**475**:263–77.
32. Hattab EM, Koch MO, Eble JN, et al. Tertiary Gleason pattern 5 is a powerful predictor of biochemical relapse in patients with Gleason score 7 prostatic adenocarcinoma. *J Urol* 2006;**175**:1695–9.
33. Pan CC, Potter SR, Partin AW, et al. The prognostic significance of tertiary Gleason patterns of higher grade in radical prostatectomy specimens: a proposal to modify the Gleason grading system. *Am J Surg Pathol* 2000;**24**:563–9.
34. Arora R, Koch MO, Eble JN, et al. Heterogeneity of Gleason grade in multifocal adenocarcinoma of the prostate. *Cancer* 2004;**100**:2362–6.
35. 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.
36. Kunju LP, Daignault S, Wei JT, et al. Multiple prostate cancer cores with different Gleason grades submitted in the same specimen container without specific site designation: should each core be assigned an individual Gleason score? *Hum Pathol* 2009;**40**:558–64.
37. Tolonen TT, Kujala PM, Tammela TL, et al. Overall and worst gleason scores are equally good predictors of prostate cancer progression. *BMC Urol* 2011;**11**:21.
38. Trpkov K, Sangkhamanon S, Yilmaz A, et al. Concordance of "case level" global, highest, and largest volume cancer grade group on needle biopsy versus grade group on radical prostatectomy. *Am J Surg Pathol* 2018;**42**:1522–9.
39. Verhoef EI, Kweldam CF, Kummerlin IP, et al. Characteristics and outcome of prostate cancer patients with overall biopsy Gleason score 3+4=7 and highest Gleason score 3+4=7 or > 3+4=7. *Histopathology* 2018;**72**:760–5.