Integrating Tertiary Gleason 5 Patterns into Quantitative Gleason Grading in Prostate Biopsies and Prostatectomy Specimens

Guido Sauter a,*, Till Clauditz a, Stefan Steurer a, Corinna Wittmer a, Franziska Büscheck a, Till Krech a, Florian Lutz a, Maximilian Lennartz a, Luisa Harms a, Lisa Lawrenz a, Christina Möller-Koop a, Ronald Simon a, Frank Jacobsen a, Waldemar Wilczak a, Sarah Minner a, Maria Christina Tsourlakis a, Viktoria Chirico a, Sören Weidemann a, Alexander Haese b, Thomas Steuber b, Georg Salomon b, Michael Matiu c, Eik Vettorazzi d, Uwe Michl b, Lars Budäus b, Derya Tilki b, Imke Thederan b, Dirk Pehrke b, Burkhard Beyer b, Christoph Fraune a, Cosima Göbel a, Marie Heinrich a, Manuela Juhnke a, Katharina Möller a, Ahmed Abdulwahab Abdullah Bawahab a, Ria Uhlig a, Hartwig Huland b, Hans Heinzer b, Markus Graefen b, Thorsten Schlomms b,e

a Department of Pathology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; b Martini-Klinik, Prostate Cancer Center, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; c Department of Ecoclimatology, Technical University Munich, Munich, Germany; d Department of Medical Biometry and Epidemiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; e Department of Urology, Section for Translational Prostate Cancer Research, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Abstract

Background: Presence of small (tertiary) Gleason 5 pattern is linked to a higher risk of biochemical recurrence in prostate cancer. It is unclear, however, how to integrate small Gleason 5 elements into clinically relevant Gleason grade groups.

Objective: To analyze the prognostic impact of Gleason 5 patterns in prostate cancer and to develop a method for integrating tertiary Gleason 5 patterns into a quantitative Gleason grading system.

Design, setting, and participants: Prostatectomy specimens from 13 261 consecutive patients and of 3295 matched preoperative biopsies were available. Percentages of Gleason 3, 4, and 5 had been recorded for each cancer.

Outcome measurements and statistical analysis: Results and limitations: Our data demonstrate that minimal Gleason 5 areas have strong prognostic impact in Gleason 7 carcinomas, while further expansion of the Gleason 5 pattern population has less impact. We thus defined an integrated quantitative Gleason score (IQ-Gleason) by adding a lump score of 10 to the percentage of unfavorable Gleason pattern (Gleason 4/5) if any Gleason 5 was present and by adding another 7.5 points in case of a Gleason 5 fraction >20%. There was a continuous increase of the risk of prostate-specific antigen recurrence with increasing IQ-Gleason. This was also true for subgroups with identical Cancer of the Prostate Risk Assessment Postsurgical scores (p < 0.0001) or Gleason grade groups (p < 0.0001).

Conclusions: The IQ-Gleason represents a simple and efficient approach for combining both quantitative Gleason grading and tertiary Gleason grades in one highly prognostic numerical variable.

* Corresponding author. Institute of Pathology, Martinistraße 52, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. Tel. +49 40 7410 55500; Fax: +49 40 7410 55997. E-mail address: g.sauter@uke.de (G. Sauter).

http://dx.doi.org/10.1016/j.eururo.2017.01.015

0302-2838/© 2017 Published by Elsevier B.V. on behalf of European Association of Urology.
Patient summary: Prostatectomy specimens (13 261) were analyzed to estimate the relevance of small Gleason 5 elements in prostate cancers. Even the smallest Gleason 5 areas markedly increased the risk of prostate-specific antigen recurrence after surgery. Larger fractions of Gleason 5 patterns had less further impact on prognosis. Based on this, a numerical Gleason score (integrated quantitative Gleason score) was defined by the percentages of Gleason 4 and 5 patterns, enabling a refined estimate of patient prognosis.

© 2017 Published by Elsevier B.V. on behalf of European Association of Urology.

1. Introduction

In radical prostatectomies, the Gleason score is defined by the most and the second-most frequent architectural pattern, while an additional third pattern is added as a tertiary Gleason grade if it is worse than the primary and secondary grade. The presence of a tertiary Gleason grade is significantly linked to unfavorable tumor features [1–4]. It is unclear, however, what minimal fraction of Gleason 5 patterns is needed to exert a prognostic role and how the tertiary Gleason pattern could optimally be included into a summary Gleason grade [1,5,6]. Considering the lack of convincing data in this area, the recent consensus paper on Gleason Grading by the International Society of Urologic Pathology (ISUP) postponed suggestions on the integration of a tertiary Gleason 5 pattern [7]. It is thus not surprising that many urologists have difficulties in interpreting the clinical impact of tertiary Gleason patterns and may even disregard this important additional grading information [8].

At our institution, we have analyzed more than 20 000 radical prostatectomy specimens between 2005 and 2015. For all these patients, we had recorded the relative quantities of Gleason patterns present in their tumors. Using the percentage of Gleason 4 patterns as a classifier we had earlier shown that the risk of prostate-specific antigen (PSA) recurrence continuously increases with a rising percentage of Gleason 4 patterns [9]. While this quantitative Gleason grading enables a finer prognostic discrimination within Gleason 7 score prostate cancers than traditional grade grouping, it does not further classify cancers with a tertiary Gleason grade. In this present study, we thus investigated the relative role of the percentage of Gleason 4 and 5 patterns in a cohort of 13 261 patients with complete follow-up data and developed a method for integrating both patterns into one continuous numerical score.

2. Material and methods

2.1. Patients

Prostatectomy specimens from 13 261 consecutive patients operated between 2005 and 2015 at the University Medical Center Hamburg-Eppendorf had complete follow-up data. Preoperative biopsies from at least eight locations were available from 3295 patients. For each cylinder the length of the cancer was recorded as well as the percentages of Gleason 3, 4, and 5 patterns. For details see the Supplementary data.

2.2. Integrated quantitative Gleason score

The integrated quantitative Gleason score (IQ-Gleason) combines all Gleason pattern data of a prostate cancer in one continuous numerical value. It ranges from 0 to 117.5 and is calculated as follows: percentage of unfavorable Gleason pattern (Gleason 4 + Gleason 5) + 10 score points if any Gleason 5 pattern was seen + another 7.5 score points in case of Gleason 5 quantities >20%. For example, the IQ-Gleason of a Gleason 3 + 4 = 7 cancer with 40% Gleason 4 is 40, the IQ-Gleason of a Gleason 3 + 4 + 7/tertiary grade (TG) 5 cancer with 40% Gleason 4 and 5% Gleason 5 is 40 + 5 + 10 + 7.5. This represents the basis value (percentage of unfavorable Gleason pattern = 45) plus 10 score points because there is <20% Gleason 5 pattern. The IQ-Gleason of a (Gleason 4 + 5 = 9) cancer with 60% Gleason 4 and 40% Gleason 5 is 60 + 40 + 10 + 7.5 = 117.5. For details see Supplementary data.

2.3. Cancer of the Prostate Risk Assessment, Postsurgical score

The Cancer of the Prostate Risk Assessment, Postsurgical score (CAPRA-S; including preoperative PSA, Gleason score, surgical margin status, extracapsular extension, seminal vesicle invasion, and lymph node invasion) was calculated for a subset of 12 992 cancers as described [10].

2.4. Statistics

See Supplementary data.

3. Results

3.1. Gleason Grade and tumor phenotype

The relationship of classical Gleason categories with other clinical and pathological parameters is shown in Table 1. Gleason 5 pattern occurred in 2883 (22%) of our patients including 1967 (19%) patients with Gleason 3 + 4 or 4 + 3 cancers where the Gleason 5 component was small enough to result in a tertiary pattern.

3.2. Prognostic impact of the fractions of Gleason 4 and 5 patterns in Gleason 3 + 4 cancers with a tertiary Gleason 5 grade

The separate analysis of the percentage of Gleason 4 (Fig. 1A) and Gleason 5 patterns (Fig. 1B) within the subgroup of 742 Gleason 3 + 4/TG 5 cancers revealed a fundamental difference in their prognostic impact. For the fraction of Gleason 4 patterns, there was a continuous increase of the risk of PSA recurrence with increasing percentage of Gleason 4 patterns (Fig. 1A). In contrast, prognosis deteriorated sharply in the presence of minimal amounts of Gleason 5 patterns and only worsened slightly with increasing Gleason 5 percentage.

3.3. Prognostic impact of the fractions of Gleason 4 and 5 patterns in Gleason 4 + 3 cancers with a tertiary Gleason 5 grade

A similar analysis for the subgroup of 1225 Gleason 4 + 3 = 7/TG 5 cancers revealed a continuous increase of
the risk of PSA recurrence with increasing percentage of Gleason 4 patterns (Fig. 1C). In contrast, prognosis again deteriorated sharply if only minimal amounts of Gleason 5 patterns were present and further worsened only slightly with increasing Gleason 5 percentage (Fig. 1D).

3.4. Role of tertiary Gleason grades in subgroups defined by comparable quantitative Gleason grades

The impact of the fraction of Gleason 5 patterns within cancers with comparable percentages of unfavorable Gleason (patterns 4/5) is shown in Fig. 2A–D. In this representation, for example, cancers in the group “21–50%” contain a total of 21–50% Gleason 4 and 5 patterns. The group “21–50%/no Gleason 5” contain 21–50% pure Gleason 4 pattern, while the group “21–50%/1–5% Gleason 5” contains 21–50% Gleason 4/5 patterns including 1–5% Gleason 5. The group “21–50%/>5% Gleason 5” contains 21–50% Gleason 4/5 including >6% Gleason 5 pattern. These data show that also within the quantitative Gleason system, the presence of minimal quantities of Gleason 5 patterns already deteriorates prognosis significantly, while larger fractions of Gleason 5 patterns have less additional impact.

3.5. IQ-Gleason

The relationship between the IQ-Gleason and the risk of recurrence is shown in Fig. 2E and F. These data show a continuous increase of the risk for PSA recurrence with increasing IQ-Gleason \( p < 0.0001 \). To estimate the gain of predictive power achieved by quantification of Gleason 4–5 patterns, we performed receiver operating characteristic/area under the curve (AUC) analysis. We took into account that finding a small fraction (smaller than the fraction of Gleason 3 and of Gleason 4) of Gleason 5 patterns in a Gleason 7 score cancer is not equally handled by pathologists. Some would always call this a territory Gleason 5 (Method A). Others would make Gleason 5 the second pattern in their Gleason score if it involves at least 5% of the cancer (Method B). The following methods for reporting Gleason grade were compared: (1) Gleason grade groups 1–5 recommended by the ISUP/World Health Organization (WHO) in 2016 [7] regarding any Gleason 5 patterns as tertiary pattern (Method A), (2) Gleason grade groups 1–5 recommended by the ISUP/WHO in 2016 [7] regarding small Gleason 5 patterns >5% as a secondary pattern (Method B), (3) similar groups for for 1 but factoring in the percentage of Gleason 5 patterns, (4) similar groups as for 2 but factoring in the percentage of Gleason 5 patterns, (5) quantitative Gleason grading as earlier defined [9], (6) IQ Gleason (Supplementary Fig. 1). This analysis demonstrated a marked increase of the predictive power for both IQ Gleason \((AUC = 0.77)\) and quantitative Gleason \((AUC = 0.77)\) as compared with traditional grade groups recommended by the ISUP/WHO in 2016 irrespective of whether Gleason 5 patterns percent were factored in \((AUC 0.72/0.73)\) or not \((AUC = 0.72/0.73)\).

3.6. Comparison to Gleason grade groups

Fig. 3 shows comparisons of the IQ-Gleason with ISUP/WHO 2016 [7] Gleason grade groups 2–5 determined according to Method A (panel a) or according to Method B (panel b). For these analyses, tertiary Gleason grades were disregarded as these are not included in the five grade groups. Additional calculations show the impact of IQ-Gleason in grade groups 2 and 3 after exclusion of cases with a tertiary grade 5 and
for the grade group 2 and 3 cancers with a tertiary grade 5 (Supplementary Fig. 2). All these analyses show that the IQ-Gleason provides substantial prognostic information beyond the conventional grade groups in almost all instances. This is particularly true if Gleason 5 fractions of ≥5% are considered secondary patterns and if smaller fractions of Gleason 5 patterns are disregarded (Fig. 3B; \( p < 0.0001 \) each).

3.7. **Comparison with a clinical nomogram (CAPRA-S)**

Applying the IQ-Gleason in cancers defined by identical CAPRA-S categories revealed that the IQ-Gleason provided substantial additional prognostic information beyond this nomogram. This not only holds true for the subsets of low-, intermediate-, and high-risk cancers (Supplementary Fig. 3; \( p < 0.0001 \) each) but even for subsets defined by identical CAPRA-S scores as long as the tumors belonged to the low- and intermediate-risk categories (Fig. 4; \( p < 0.0001 \) each).

3.8. **IQ-Gleason in biopsies**

The impact of the biopsy IQ-Gleason on the likelihood of finding an unfavorable Gleason grade in a subsequent prostatectomy specimens is shown in Fig. 5. The data demonstrate that the risk of having an unfavorable result in the prostatectomy continuously increases with higher IQ-Gleason scores. This also applies within patient groups defined by identical worst biopsy result (Supplementary Fig. 4).

4. **Discussion**

In 10,483 consecutive Gleason 3 + 4 and 4 + 3 cancers, we found a tertiary Gleason 5 pattern in 1967 patients (19%). This is in the range of earlier studies describing a tertiary Gleason 5 pattern in 7.1–29% of cancers [2,11–14]. Minor discrepancies between these studies are explained by interobserver variability on diagnosing Gleason pattern 5.
Fig. 2 – Prognostic impact of the fraction of Gleason 5 patterns in cancers with (a) 21–50% Gleason 4 + 5 patterns, (B) 51–65% Gleason 4 + 5 patterns, (C) 66–85% Gleason 4 + 5 patterns, and (D) >86% Gleason 4 + 5 patterns. (E) Relationship between the continuous integrated quantitative Gleason score (IQ-Gleason score) and the probability for 5-yr recurrence-free survival (RFS). The IQ-Gleason ranges from 0 (no Gleason 4/5 patterns) to 117.5 (100% Gleason 4/5 patterns including > 20% Gleason 5 patterns). (F) Prognostic impact of the IQ-Gleason after categorization in 10 groups. Gl.4 = ant Gleason 4 pattern; Gl.5 = any Gleason 5 pattern; Gl.4 + 5 = any Gleason 4 + 5 pattern; PSA = prostate-specific antigen.

Remarkably, even minimal amounts of Gleason 5 patterns lead to a significant deterioration of patient prognosis, while a further expansion of Gleason 5 areas was clinically much less relevant. The nature of the Gleason grading offers a speculative explanation for this observation. The Gleason grading is solely based on architectural features and strictly disregards cytological atypia. This is not undisputed. Data suggest that additional consideration of nuclear changes could improve the Gleason grading [17–21]. Indeed, it is our perception that nuclear atypia is often pronounced in Gleason 4 patterns with transition into Gleason 5. It is thus possible, that Gleason 5 pattern is not only prognostic because of its mere existence but also because it preferably arises in particularly unfavorable Gleason 4 patterns with severe cellular atypia, which under Gleason grading criteria is not distinguishable from less dangerous Gleason 4 patterns. In this context, the rapid evolution of automated digital...
Fig. 3 – Prognostic impact of the integrated quantitative Gleason score in subsets of cancers of identical Gleason grade groups using two different methods of defining a tertiary Gleason 5 grade. (A) Tertiary grade 5 was assumed if Gleason 5 was the third-most pattern. (B) Tertiary grade 5 was assumed only if the fraction of Gleason 5 glands was less than 5% in cancers with Gleason 3+4 or 4+3.

PSA = prostate-specific antigen.
image recognition holds substantial promise. We anticipate that self-learning imaging systems will soon be able to integrate all architectural, cytological, and nuclear information on a prostate cancer and to define patient outcome predictors that go way beyond established Gleason patterns.

To combine all available morphologic information into one numerical prognostic value, we used the fraction of non-Gleason 3 patterns as the basis score, because of its pivotal role for recurrence prediction. Two simple mathematical modifications were applied to integrate Gleason 5 pattern findings. To account for the significant deterioration of outcome induced by even minimal amounts of Gleason 5 patterns we first added an extra score of 10 to the percentage of unfavorable Gleason pattern (Gleason 4/5) if any Gleason 5 pattern was present. The value of 10 was selected because presence of any Gleason 5 pattern increased the risk of PSA recurrence in a comparable way as 10% more Gleason 4 patterns. Second, we added another 7.5 points to the percentage of unfavorable Gleason pattern in case of a very high fraction of Gleason 5 patterns (>20%) because of the somewhat worse outcome in such cases in some analyses. Arguments supporting the utility of our IQ-Gleason for prognosis prediction include: (1) fine (continuous) discrimination of the risk for PSA recurrence after prostatectomy, (2) substantial prognostic information provided beyond established clinical nomograms such as the CAPRA-S score, (3) substantial prognostic information provided beyond Gleason grade groups [7], and (4) discrimination of multiple prognostic groups within Gleason 3 + 4/TG5 and Gleason 4 + 3/TG5 cancers.

To evaluate the IQ-Gleason for pretreatment decision-making, findings in cancer containing biopsies of 3295 patients were compared with those in subsequent prostatectomy specimens. The continuous increase of unfavorable Gleason grade in prostatectomies with increasing biopsy IQ-Gleason further validates the IQ-Gleason concept. Although describing prostate cancer malignancy by a score ranging from 0 to 117.5 might appear more complicated than just employing five categories, this concept is in line with our clinical experience that prostate cancer aggressiveness is a continuous rather than a categorical variable. It is important to note that all thresholds for defining IQ-Gleason subgroups in our study were artificially selected to visualize prognostic impact. We do not propose to use these subgroups as categorical prognostic groups. A rapidly increasing number of investigators are suggesting that at least a subset of Gleason 3 + 4 cancers can be treated by active surveillance [22,23]. We postulate that these are the cancers with a relatively low IQ-Gleason, for example ranging below 15. The significant prognostic impact of the IQ-Gleason in patients having a Gleason score of 3 + 4, 4 + 3, 8, or 9–10 in their worst individual biopsy cylinder challenges the frequently used practice of classifying patients solely according to their worst biopsy result.

It is of note, that applying the IQ-Gleason to prostate biopsies could reduce the fraction of cancers where
Fig. 5 – Comparison of the integrated quantitative Gleason score (IQ-Gleason) in prostate biopsies with (A) the classical Gleason score or (B) the IQ-Gleason in subsequent prostatectomy specimens of 3432 patients. (C) Plot of the probabilities of having Gleason ≤ 3+4 (red line), 4+3 (green line), or ≥ 8 (blue line) in the subsequent prostatectomy specimens in comparison to the IQ-Gleason.

RPE = radical prostatectomy; T5 = tertiary 5.
pathologists largely differ in their Gleason grading. Extreme interobserver variations are often due to small fractions of Gleason 5 patterns that are called by one but not by the other pathologist. For example, 2% Gleason 5 pattern called by one pathologist in a cancer with otherwise 30% Gleason 4 pattern would change the Gleason score from 3 + 4 (which is grade group 2 of 5) to 3 + 5 (grade group 4 of 5), which is substantial, but would change the IQ-Gleason only from 30 to 40. A tumor called 4 + 3 (grade group 3 of 5) with 55% Gleason 4 would change to 4 + 5 (grade group 5 of 5) if another pathologist considered 2% of the cancer Gleason 5 instead of Gleason 4. This discrepancy would only change the IQ-Gleason from 55 to 65, which again constitutes a minor difference.

The most significant limitation of our study is their origin in one high-throughput center, where the biopsy process, pathology, and surgery are highly standardized. The absolute numbers presented in our study may thus not be 1:1 transferable to other centers. Given the well-known interobserver variability of Gleason grading, it is obvious that pathologists will not always concur in their Gleason pattern percentage estimate [20,24–27]. Our approach of not limiting Gleason grade assessment to the index tumor and our conservative position towards irregular glands as Gleason 4 pattern is in disagreement with criteria suggested by ISUP.

The IQ-Gleason—statistically seen—is a superior predictor of outcome as compared with the traditional Gleason grade groups. It does, however, not result in better receiver operating characteristic values than the previously reported quantitative Gleason grading approach. Some might thus argue, that minor fractions of Gleason 5 could as well be disregarded. However, the strong prognostic impact of minor fractions of Gleason 5 pattern in virtual every subgroup defined by identical percentages of non-Gleason 3 cancer argues against such models. The IQ-Gleason system fixes a fundamental weakness of the quantitative Gleason grade. The quantitative Gleason grade [9] represents a mixed system combining continuous and categorical values to classify prostate cancers. It defines about 80% of cancers by a continuous score ranging from 0 to 100, while the remaining 20% (those with a TG5) are put in only two categories (3 + 4[TG5] and 4 + 3[TG5]). Besides being purely continuous the IQ-Gleason solves the issue of variable definitions used in the pathology community for tertiary Gleason 5 grades.

5. Conclusions

In summary, our data indicate a paramount importance of the percentage of unfavorable Gleason patterns for assessing prostate cancer aggressiveness. Gleason 5 areas are the morphologic correlate for advanced tumor progression and herald—almost irrespective of their quantity—further deterioration of prognosis. The IQ-Gleason represents a simple and efficient approach for combining both quantitative Gleason grading and tertiary Gleason grades in one highly prognostic numerical variable. For the future, it can be hoped, that the number of institutions collecting follow-up data of their patients will increase. The ability to systematically compare histology with clinical outcome will eventually improve the quality of biopsy interpretation at such centers.

Author contributions: Guido Sauter 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: Sauter, Schlomn, Heinzer.

Acquisition of data: Sauter, Clauditz, Steurer, Wittmer, Büscheck, Krech, Lutz, Harms, Lawrenz, Jacobson, Wilczak, Minner, Tsurulakis, Weidemann, Fraune, Göbel, Heinrich, Möller, Bawahab, Uhlig.

Analysis and interpretation of data: Sauter, Simon, Matiu.

Drafting of the manuscript: Sauter, Clauditz, Steurer, Simon, Schlomn. Critical revision of the manuscript for important intellectual content: Heinzer, Beyer, Minner.

Statistical analysis: Simon, Matiu, Pehrke, Vetterazzi.

Obtaining funding: None.

Administrative, technical, or material support: Möller-Koop.

Supervision: Sauter, Schlomn.

Other (provided clinical data): Haese, Schlomn, Sauter, Michl, Budäus, Tilk, Heinzer, Graefen.

Financial disclosures: Guido Sauter 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.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.eururo.2017.01.015.

References


