Clinical Utility of Quantitative Gleason Grading in Prostate Biopsies and Prostatectomy Specimens

Guido Sauter a,*, Stefan Steurer a, Till Sebastian Clauditz a, Till Krech b, Corinna Wittmer a, Florian Lutz a, Maximilian Lennartz a, Tim Janssen a, Nayira Hakimi a, Ronald Simon a, Mareike von Petersdorff-Campen a, Frank Jacobsen a, Katharina von Loga a, Waldemar Wilczak a, Sarah Minner a, Maria Christina Tsourlakis a, Viktoria Chirico a, Alexander Haese b, Hans Heinzer b, Burkhard Beyea b, Markus Graefen b, Uwe Michlb, Georg Salomon b, Thomas Steuber b, Lars Henrik Budås b, Elena Hekeler a, Julia Malsy-Mink a, Sven Kutzera a, Christoph Fraune a, Cosima Göbel a, Hartwig Hulana b, Thorsten Schlommc

a Institute of Pathology, University Medical Center Hamburg-Eppendorf, Germany; b Martini-Klinik, Prostate Cancer Center, University Medical Center Hamburg-Eppendorf, Germany; c Department of Urology, Section for translational Prostate Cancer Research, University Medical Center Hamburg-Eppendorf, Germany

Article info

Article history:
Accepted October 15, 2015

Associate Editor:
James Catto

Keywords:
Prostate cancer
Gleason score
Quantitative Gleason grade
Prognosis

Abstract

Background: Gleason grading is the strongest prognostic parameter in prostate cancer. Gleason grading is categorized as Gleason ≤6, 3 + 4, 4 + 3, 8, and 9–10, but there is variability within these subgroups. For example, Gleason 4 components may range from 5–45% in a Gleason 3 + 4 + 7 cancer.

Objective: To assess the clinical relevance of the fractions of Gleason patterns.

Design, setting, and participants: Prostatectomy specimens from 12,823 consecutive patients and of 2,971 matched preoperative biopsies for which clinical data with an annual follow-up between 2005 and 2014 were available from the Martini-Klinik database.

Outcome measurements and statistical analysis: To evaluate the utility of quantitative grading, the fraction of Gleason 3, 4, and 5 patterns seen in biopsies and prostatectomies were recorded. Gleason grade fractions were compared with prostatectomy findings and prostate-specific antigen recurrence.

Results and limitations: Our data suggest a striking utility of quantitative Gleason grading. In prostatectomy specimens, there was a continuous increase of the risk of prostate-specific antigen recurrence with increasing percentage of Gleason 4 fractions with remarkably small differences in outcome at clinically important thresholds (0% vs 5%; 40% vs 60% Gleason 4), distinguishing traditionally established prognostic groups. Also, in biopsies, the quantitative Gleason scoring identified various intermediate risk groups with respect to Gleason findings in corresponding prostatectomies. Quantitative grading may also reduce the clinical impact of interobserver variability because borderline findings such as tumors with 5%, 40%, or 60% Gleason 4 fractions and very small Gleason 5 fractions (with pivotal impact on the Gleason score) are discarded.

Conclusions: Quantitative Gleason pattern data should routinely be provided in addition to Gleason score categories, both in biopsies and in prostatectomy specimens.

Patient summary: Gleason score is the most important prognostic parameter in prostate cancer, but prone to interobserver variation. The results of our study show that morphological aspects that define the Gleason grade in prostate cancer represent a continuum. Quantitation of Gleason patterns provides clinically relevant information beyond the traditional Gleason grading categories ≤3 + 3, 3 + 4, 4 + 3, 8, 9–10. Quantitative Gleason scoring can help to minimize variations between different pathologists and substantially aid in optimized therapy decision-making.

© 2015 European Association of Urology. Published by Elsevier B.V. All rights reserved.

* 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).
1. Introduction

The Gleason grade is the strongest prognostic parameter in prostate cancer. It is purely based on cancer architecture and not influenced by cellular morphology. Five different patterns with increasing abnormalities of glandular structure have initially been defined, but patterns 1 and 2 are hardly used in contemporary praxis. A maximum of three different patterns can be part of the Gleason grade. These include the most prevalent pattern, the second most prevalent pattern, and the worst or least differentiated pattern (if different from the most or second most frequent pattern). The rules for defining the Gleason score vary between large and small specimens. By convention, the most and worst Gleason patterns form the Gleason score in needle biopsies. In radical prostatectomy specimens and in transurethral resections, the Gleason score is defined by the most and the second most frequent pattern, while a third (worst) pattern can be added as a tertiary Gleason pattern.

The Gleason grading was first described in 1966 [1]. It has undergone various modifications and clarifications since then. The last clarification appeared in 2005, when the contemporary practice was summarized by a group of more than 70 prostate cancer pathologists [2]. The Gleason grade is a key parameter for determining appropriate patient treatment according to all major guidelines but it is subject to substantial interobserver variability. Many of the cancers differently interpreted by multiple pathologists represent borderline cases between $3 + 3 = 6$ and $3 + 4 = 7$ or between $3 + 4 = 7$ and $4 + 3 = 7$ [3].

The current study is based on the hypothesis that a more detailed description of the Gleason grade, including quantification of all components (ie, Gleason grade patterns) would result in a more precise outcome prediction than traditional categorization. To test this hypothesis we compared quantitative Gleason grading data with prostate-specific antigen (PSA) recurrence in 12 823 consecutive patients treated with a prostatectomy between 2005 and 2014, and compared biopsy with prostatectomy findings in 2971 patients from which corresponding preoperative biopsies were available.

2. Material and methods

2.1. Patients

Radical prostatectomy specimens were available from 20 828 consecutive patients, undergoing surgery between 1992 and 2014 at the Department of Urology and the Martini-Klinik, Prostate Cancer Center at the University Medical Center Hamburg-Eppendorf. Follow-up data were obtained from the Martini-Klinik database, which was initiated in 1992 (H. Huland) with yearly patient reported outcome measurements and a typical annual follow-up rate of more than 90% (Supplementary Fig. 1). For the current study, 12 823 patients were selected who were diagnosed in 2005 and later, and had complete follow-up data ($n = 12 150$) and/or from which Gleason data were available from matched biopsies and radical prostatectomy specimens ($n = 2971$). There was a median follow-up of 38.3 mo (range: 1–121 mo; Table 1). PSA values were measured following surgery and PSA recurrence was defined as a postoperative PSA $> 0.2$ ng/ml confirmed by a second determination with a serum PSA $> 0.2$ ng/ml. All prostate specimens were analyzed according to a standard procedure, including complete embedding of the entire prostate for histological analysis [4]. The starting point of the study was selected, because the pathology department at the University Medical Center Hamburg-Eppendorf has been led by G. Sauter since January 2005 until the present day. Gleason grading attitudes did not change at our institution. Gleason grading was largely done according to the International Society of Urological Pathology recommendations describing the contemporary practice in 2005 [2]. We did, however, not change our conservative position towards “small irregular glands” as a Gleason 4 feature because these are accompanied by unequivocal gland fusions (a traditional Gleason 4 feature) in the vast majority of cases [2]. This position was subsequently also endorsed by others [5]. We also did not limit our Gleason score to the index tumor (which is not always obvious to identify) but did a “total” Gleason scoring including all cancer foci of a prostate. Since 2005, the percentages of each Gleason pattern were routinely recorded for each prostate cancer diagnosed at our institution. For this purpose, for each core of every available prostate cancer biopsy and for every prostatectomy specimen, the percentages of Gleason $\leq 3$, $4$, and $5$ patterns in cancerous tissue were estimated during the regular process of biopsy interpretation. This does not require special efforts because every pathologist always quantitates the percentages of each Gleason pattern in the normal procedure of prostate biopsy evaluation. The only difference in our

### Table 1 – Patient cohort

<table>
<thead>
<tr>
<th>Study cohort ($n = 12 823$)</th>
<th>No. of patients (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Follow-up (mo)</td>
<td></td>
</tr>
<tr>
<td>$n$</td>
<td>12 150</td>
</tr>
<tr>
<td>Mean</td>
<td>38.3</td>
</tr>
<tr>
<td>Median</td>
<td>36.0</td>
</tr>
<tr>
<td>Age (yr)</td>
<td></td>
</tr>
<tr>
<td>$&lt;50$</td>
<td>440 (3.4%)</td>
</tr>
<tr>
<td>51–59</td>
<td>2882 (22.5%)</td>
</tr>
<tr>
<td>60–69</td>
<td>7045 (54.9%)</td>
</tr>
<tr>
<td>$&gt;70$</td>
<td>2456 (19.2%)</td>
</tr>
<tr>
<td>Pretreatment PSA (ng/ml)</td>
<td></td>
</tr>
<tr>
<td>$&lt;4$</td>
<td>1551 (12.1%)</td>
</tr>
<tr>
<td>4–10</td>
<td>8002 (62.5%)</td>
</tr>
<tr>
<td>10–20</td>
<td>2354 (18.4%)</td>
</tr>
<tr>
<td>$&gt;20$</td>
<td>893 (7%)</td>
</tr>
<tr>
<td>pT stage (AJCC 2002)</td>
<td></td>
</tr>
<tr>
<td>pT2</td>
<td>8582 (67%)</td>
</tr>
<tr>
<td>pT3a</td>
<td>2602 (20.3%)</td>
</tr>
<tr>
<td>pT3b</td>
<td>1588 (12.4%)</td>
</tr>
<tr>
<td>pT4</td>
<td>45 (0.4%)</td>
</tr>
<tr>
<td>Gleason grade</td>
<td></td>
</tr>
<tr>
<td>$&lt;6$</td>
<td>2277 (17.8%)</td>
</tr>
<tr>
<td>3 + 4</td>
<td>6849 (53.4%)</td>
</tr>
<tr>
<td>3 + 4 TG5</td>
<td>655 (5.1%)</td>
</tr>
<tr>
<td>4 + 3</td>
<td>1176 (9.2%)</td>
</tr>
<tr>
<td>4 + 3 TG5</td>
<td>1060 (8.3%)</td>
</tr>
<tr>
<td>8</td>
<td>72 (0.6%)</td>
</tr>
<tr>
<td>9–10</td>
<td>734 (5.7%)</td>
</tr>
<tr>
<td>pN stage</td>
<td></td>
</tr>
<tr>
<td>pN0</td>
<td>7777 (60.3%)</td>
</tr>
<tr>
<td>pN+</td>
<td>1028 (11.7%)</td>
</tr>
<tr>
<td>Surgical margin</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>10 442 (82.8%)</td>
</tr>
<tr>
<td>Positive</td>
<td>2171 (17.2%)</td>
</tr>
<tr>
<td>Cancers with matched preoperative needle biopsies</td>
<td>2971</td>
</tr>
</tbody>
</table>

AJCC = American Joint Committee on Cancer; PSA = prostate specific antigen.

* Numbers do not always add up to 12 823 in the different categories because of cases with missing data.
procedure is that these “raw data” of cancer interpretation are recorded and reported to clinicians instead of only using them to select the right Gleason category, for example 3 + 4 = 7 (<50% Gleason 4) or 4 + 3 = 7 (>50% Gleason 4). Another (historical) prostatectomy cohort operated between 1992 and 2004 was available for comparison. For these patients, Gleason grading criteria might have differed slightly from our current (since 2005) practice and quantitative Gleason information was only recorded in a fraction of cases (1484 of 3915; 38%).

2.2. Biopsies

Preoperative needle biopsies containing samples from at least eight locations were available from 2971 of our patients. Preoperative biopsies were unavailable from the remaining patients, as their initial cancer diagnosis had been established in other institutions. For each biopsy location the length of the cancer involvement was recorded and a quantitative Gleason score with estimated percentages of Gleason 3, 4, and 5 was assigned. In case of discrepant Gleason findings between biopsies from different locations of a patient, an average quantitative Gleason score was calculated based on tumor length and Gleason grade percentages of all biopsies. Gleason grading rules were applied as for prostatectomy specimens in these analyses including assignment of a tertiary Gleason grade where appropriate. Biopsy findings were compared with results from corresponding prostatectomy specimens employing both the worst biopsy and average Gleason findings.

2.3. Statistics

Statistical calculations were performed with JMP 9 software (SAS Institute Inc., NC, USA). Survival curves were calculated according to Kaplan–Meier. The Log-rank test was applied to detect significant survival differences between groups. COX proportional hazards regression analysis was performed to test the statistical independence and significance between pathological and clinical variables. Chi-square test was used to compare Gleason patterns in biopsies and radical prostatectomies.

3. Results

3.1. Gleason grading in prostatectomy samples

The impact of different categorizations of Gleason Scores on patient prognosis is shown in Figure 1 for the 12,150 patients treated between 2005 and 2014. These data demonstrate a striking prognostic impact of the Gleason grading, irrespective of how subgroups are formed. This holds true for the traditional classification into 6, 7, 8, and 9–10 (data not shown) or for the subgroups of 6, 3 + 4 = 7, 4 + 3 = 7, 8, and 9–10 (Fig. 1A). However, the more subgroups are delineated, the more subgroups with a different outcome can be defined. For example, the data show that the groups of 3 + 4 = 7 and 4 + 3 = 7 can be divided into further subgroups with distinct prognostic difference \( p < 0.0001 \) each for 3 + 4 = 7 low [1–24% Gleason 4] vs 3 + 4 = 7 high [25–49% Gleason 4] and for 4 + 3 = 7 low [50–74% Gleason 4] vs 4 + 3 = 7 high [75–94% Gleason 4]; Fig. 1B). The definition of further subgroups results in an even finer distinction of the patient risk (Fig. 1C). Criteria for subgroup selection for these analyses were in order to have significant numbers of patients in every subgroup and avoiding subgroups that would include 3 + 4 and 4 + 3 cancers (such as 40–60% Gleason 4). It is noteworthy that these 126 cancers with a Gleason 4 + 3 = 7 and >80% Gleason 4 areas behaved closely similar as in Gleason 4 + 4 = 8 carcinomas. In all calculations, presence of a tertiary Gleason grade identified further prognostic relevant subgroups. All these associations remain highly significant in the subgroups of pT2, pT3a, pT3b, and nodal positive carcinomas (Supplementary Fig. 2–5). The impact of quantitative Gleason grading on patient outcome remained constant during the 10 yr study. There was no obvious difference in the Kaplan–Meier curves for patients treated and analyzed from 2005–2008 versus 2009–2014 (Supplementary Fig. 6 A–B). Even the data derived from a historical cohort treated and analyzed from 1992–2004, where quantitative Gleason data were only occasionally
recorded, showed a similar impact of percentage Gleason 4 on PSA recurrence (Supplementary Fig. 6C).

3.2. Gleason grade in biopsies

The impact of different kinds of Gleason scores on the likelihood of finding an unfavorable Gleason grade in a subsequent prostatectomy specimen is shown in Figure 2. Our data are shown for categorization according to the worst positive biopsy (Fig. 2A) and according to subgroups defined by the average Gleason score on all biopsies (Fig. 2B–D). The latter calculations include a distinction according to traditional subgroups, that is, Gleason ≤ 6, 3 + 4 = 7, 4 + 3 = 7, ≥ 8 (Fig. 2B), further separation of Gleason 3 + 4 = 7 and Gleason 4 + 3 = 7 in two subgroups each (Fig. 2C), and an even finer distinction of Gleason 7 tumors into six different subgroups (Fig. 2D). In all calculations based on average data from all tumor-containing biopsies, cancers with a “tertiary” Gleason 5 pattern are shown as additional separate subgroups. For all calculations the data demonstrate a strong predictive impact of the biopsy Gleason grade with a more subtle distinction if average Gleason grades are used. The extra value provided by the average (quantitative) Gleason grade is also demonstrated in a separate analyses of patients having a “worst” biopsy with a Gleason score of 3 + 4 = 7 \((p < 0.0001)\), 4 + 3 = 7 \((p = 0.0072)\), 4 + 4 \((p = 0.0004)\), or 4 + 5 \((p < 0.0001)\) is shown in Figure 3.

4. Discussion

Gleason 7 category is still part of many clinical recommendations and guidelines including D’Amico criteria, German Association of Urology-, American Urological Association-, and National Comprehensive Cancer Network guidelines [6–9]. A variety of studies have meanwhile shown that Gleason 7 tumors should be further subdivided, with most of them suggesting a categorization into Gleason 3 + 4 and Gleason 4 + 3 tumors depending on the predominant Gleason pattern seen in a cancer [10–14]. Our data are derived from a single high-throughput prostate cancer center performing >2000 prostatectomies annually and a quality management system resulting in continuous follow-up data in >90% of our patients (the Martini-Klinik database). The large size of our cohort \(n = 12,823\) enabled us to see that the percentage of adverse Gleason patterns in prostatectomy specimens has striking prognostic impact in prostate cancer patients and enables a further clinically relevant risk assessment within and beyond Gleason 3 + 4 and also 4 + 3 cancers. It is noteworthy that we are not suggesting a particular relevance of the subgroups that
we have selected for our analysis. We are rather demonstrating that the “percentage of unfavorable Gleason” is a meaningful parameter for assessing individual prostate cancers. These data are in agreement with earlier studies proposing a direct importance of the Gleason grade percentages in cohorts of 209 [15], 305 [16], 379 [17], and 504 [18] cancers. Overall, these findings fit well with the high diversity of clinical courses of prostate cancer rather representing a continuum than a property that can be classified in only a few subgroups.

To assess the value of a “quantitative” Gleason grade in pretreatment decision-making, the average Gleason grade on all cancer containing biopsies of 2971 patients was compared with the findings in subsequent prostatectomy specimens. The strong and statistically significant differences between subgroups going beyond the traditional “Gleason subgroups” of ≤6, 7, ≥8, ≤6, 3 ≤4 = 7, 4 + 3 = 7, ≥8 further corroborates our view of prostate cancer aggressiveness (and Gleason grading) as a continuous rather than a categorical variable. It is remarkable, that this also applies for patients having a Gleason sum of 3 + 4, 4 + 3, 4 + 4, or 4 + 5 in their individual biopsy cylinder with the “worst Gleason grade”. This observation challenges the contemporary practice to classifying patients according to their “worst biopsy cylinder”. It is indeed intuitive that a patient with prostate cancer in six of eight biopsies including five cores with Gleason 3 + 3 and one core with a small carcinoma classified as Gleason 4 + 5 (because of only a few percent of pattern Gleason 5) might be in a better situation than a patient with a Gleason 4 + 5 carcinoma in six of eight biopsies showing 60% Gleason 4 and 40% Gleason 5.

A more subtle risk prediction using quantitative aspects of the Gleason grading is of particular interest within the subgroups of “Gleason 7” or “Gleason 3 + 4 = 7” carcinomas as these patients face the widest variety of treatment options including active surveillance (AS). While some guidelines such as of the German Association of Urology, European Association of Urology, as well as American Urological Association- and National Comprehensive Cancer Network guidelines, limit the option of AS to Gleason 3 + 3 = 6 carcinomas, other authors and also clinical trials consider Gleason 3 + 4 = 7 carcinomas as eligible for conservative therapy [19–22]. Recent data from two large AS cohorts suggest that AS is equally safe in low and selected early intermediate-risk cancers [20,23]. Large ongoing prospective AS trials like the European “The Prostate cancer Research International: Active Surveillance Study” and the German PREFERE study are open for selected patients with a Gleason 3 + 4 tumor [24,25]. Given the uncertainty on whether or not a Gleason 3 + 4 = 7 carcinoma is still suitable for AS, it is tempting to speculate, that 3 + 4 = 7 carcinomas with a low Gleason 4 fraction may be these that can be conservatively treated while those with a high Gleason 4 fraction should be treated more aggressively. Our data on the impact of Gleason 4 percentage on needle biopsies strongly support this notion. Gleason 3 + 4 = 7 cancers with only 5% or 10% of Gleason 4 pattern in biopsies bear an almost identical risk of unfavorable Gleason scores in prostatectomies as Gleason 3 + 3 = 6 cancer.

Interobserver variability is one of the most striking limitations of the Gleason grading [26]. Available data from studies that also distinguish the categories 3 + 4 + 7 and 4 + 3 = 7 on needle biopsies show that >25% of Gleason 3 + 3 carcinomas are upgraded and >15% of Gleason 3 + 4 = 7 carcinomas are downgraded if the slides are reviewed by a second pathologist [3,27–29]. Remarkably, in one of these studies 15 consensus cases were selected from 25 initial cases based on the ability of at least 2/3 of expert pathologists to agree on one Gleason score [27]. That a 2/3 agreement could not be achieved in 10 out of these 25 cases

Fig. 3 – Extra value provided by the average (quantitative) Gleason grade. Comparison of the quantitative Gleason score with the radical prostatectomy score in subsets of cancers with identical worst (maximal) Gleason score seen in the biopsy. Max = maximum; RPE = radical prostatectomy.
exhibits that interobserver variability also occurs at very relevant levels among specialized pathologists. It is our experience that quantifying Gleason patterns has the potential to markedly reduce the clinical impact of interobserver variability. For example, we would expect that patients diagnosed as a Gleason 3 + 3 = 6 by one pathologist but as a Gleason 3 + 4 with 5% or 10% Gleason 4 by another pathologist could still be considered for AS. It is noteworthy that quantitative Gleason grading is not a complicated process for a pathologist. Identifying and quantifying the individual Gleason patterns is the core element of every reading of a prostate cancer specimen. For example, the only difference between Gleason 3 + 4 = 7 and 4 + 3 = 7 is whether or not the percentage of Gleason pattern 4 exceeds 50%. The only additional work in quantitative Gleason grading is to record the “raw data” of the biopsy interpretation and calculating a “patient average” in case of multiple biopsies with variable findings.

The most significant limitation to our study is their origin in only one high-throughput center, where the biopsy process, pathology, and surgery are highly standardized. It is thus possible that the absolute numbers presented in our study may 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 agree in their assessments of Gleason pattern percentages. Our data do suggest, however, that quantitative Gleason grade information provides valuable additional information to classical Gleason grade categories. For the future, it can be hoped that the number of institutions having budget and capabilities for a systematic follow-up of their patients will sharply increase. We anticipate that the ability to compare histology with clinical outcome will eventually improve the quality of biopsy interpretation. With respect to conventional pathology, training sets of biopsies are needed that not only have “expert grading” but also clinical outcome as endpoints. Given the progress in digital image analysis, it further appears likely that outcome-controlled biopsy cohorts will eventually enable the development of novel “automated” computer based prostate cancer grading systems. The use of PSA recurrence as an endpoint is another study limitation. However, PSA recurrence is an excellent surrogate parameter for unfavorable disease outcome. Every feature that has so far been linked to prostate cancer mortality has also been found associated with PSA recurrence. This also includes the Gleason grade [5].

In summary, our data demonstrate that clinically relevant morphologic information represents a continuum in prostate cancer and that histologic information can be provided going beyond classical Gleason categories. A detailed description of the percentages of Gleason patterns in a cancer might not only provide more subtle prognostic information but also reduce the clinical impact of interobserver variability as pathologists automatically identify borderline findings as such.

**Study concept and design:** Guido Sauter, Hartwig Huland, Thorsten Schlomm.

**Acquisition of data:** Stefan Steurer, Till Sebastian Clauditz, Till Krehc, Corinna Wittmer, Florian Lutz, Tim Janssen, Nayira Hakimi, Mareike von Petersdorff-Campen, Frank Jacobsen, Katharina von Loga, Waldemar Wilczak, Sarah Minner, Maria Christina Tsourlakis, Viktoria Chirico, Alexander Haese, Hans Heinzer, Burkhard Beyer, Markus Graefen, Uwe Michl, Georg Salomon, Thomas Steuber, Lars Henrik Budäus, Elena Hekeler, Julia Malsy-Mink, Sven Kutzera, Christoph Fraune, Cosima Göbel.

**Analysis and interpretation of data:** Guido Sauter, Florian Lutz, Maximilian Lennartz, Tim Janssen, Nayira Hakimi, Ronald Simon, Thorsten Schlomm.

**Drafting of the manuscript:** Guido Sauter, Maximilian Lennartz, Ronald Simon, Thorsten Schlomm.

**Critical revision of the manuscript for important intellectual content:** Guido Sauter, Stefan Steurer, Ronald Simon, Hartwig Huland, Thorsten Schlomm.

**Statistical analysis:** Maximilian Lennartz, Ronald Simon.

**Obtaining funding:** None.

**Administrative, technical, or material support:** Florian Lutz, Tim Janssen, Nayira Hakimi.

**Supervision:** Guido Sauter, Hartwig Huland, Thorsten Schlomm.

**Other:** None.

**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.2015.10.029.

**References**


**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.


