Diagnosis of "Poorly Formed Glands" Gleason Pattern 4 Prostatic Adenocarcinoma on Needle Biopsy

An Interobserver Reproducibility Study Among Urologic Pathologists With Recommendations

Ming Zhou, MD, PhD, Jianbo Li, PhD,† Liang Cheng, MD, PhD,‡ Lars Egevad, MD,§ Fang-Ming Deng, MD,* Lakshmi Priya Kunju, MD,‖ Cristina Magi-Galluzzi, MD, PhD,† Jonathan Melamed, MD,* Rohit Mehra, MD,‖ Savvas Mendrinos, MD,¶ Adeboye O. Osunkoya, MD,# Gladell Paner, MD,** Steve S. Shen, MD, PhD,†† Toyonori Tsuzuki, MD,‡‡ Kiril Trpkov, MD,§§ Wei Tian, MD,¶ Ximing Yang, MD, PhD,‖‖ and Rajal B. Shah, MD¶*

Abstract: Accurate recognition of Gleason pattern (GP) 4 prostate carcinoma (PCa) on needle biopsy is critical for patient management and prognostication. "Poorly formed glands" are the most common GP4 subpattern. We studied the diagnostic reproducibility and the quantitative threshold of grading GP4 "poorly formed glands" and the criteria to distinguish them from tangentially sectioned GP3 glands. Seventeen urologic pathologists were first queried for the definition of "poorly formed glands" using cases representing a spectrum of PCa glandular differentiation. Cancer glands with no or rare lumens, elongated compressed glands, and elongated nests were considered "poorly formed glands" by consensus. Participants then graded a second set of 23 PCa cases that potentially contained "poorly formed glands" with a fair interobserver agreement ($\kappa = 0.34$). The consensus diagnoses, defined as agreement by $>70\%$ participants, were then correlated with the quantitative (≤ 5, 6 to 10, >10) and topographic features of poorly formed glands (clustered, immediately adjacent to, and intermixed with other well-formed PCa glands) in each case. Poorly formed glands immediately adjacent to other well-formed glands regardless of their number and small foci of ≤ 5 poorly formed glands regardless of their location were not graded as GP4. In contrast, large foci of >10 poorly formed glands that were not immediately adjacent to well-formed glands were graded as GP4. Grading "poorly formed glands" is challenging. Some morphologic features are, however, reproducible for and against a GP4 diagnosis. This study represents an important step in standardization of grading of "poorly formed glands" based on quantitative and topographic morphologic features.

Key Words: Gleason grade, Gleason pattern 4, poorly formed glands, prostate cancer, prostate needle biopsy, interobserver reproducibility

(*Am J Surg Pathol* 2015;39:1331–1339)

BACKGROUND

Gleason score assigned to a prostate carcinoma (PCa) on needle biopsy is one of the most powerful indicators of clinical outcomes and is also used to determine patient management and therapy.[1–5] Of particular importance is the correct identification of Gleason pattern (GP) 4 and 5 PCa and distinction from GP3 PCa. The presence of any GP4 constitutes a clinically significant cancer[5,6] and may disqualify patients for and prompt termination of active surveillance by most protocols.[7]

The definition of GP4 has evolved significantly since the inception of the Gleason grading system. It was first described as "large clear cells growing in a diffuse pattern resembling hypernephroma"[8] and was later expanded to include "raggedly infiltrating, fused-glandular tumor"[2] and "glands that are not single and separate, but coalesce and branch."[9] The 2005 International Society of Urological Pathology (ISUP) consensus conference on Gleason grading of prostatic carcinoma further expanded GP4 to include ill-defined glands with poorly formed glandular lumens in addition to 3 previously recognized subpatterns (fused microacinar glands, large cribriform

From the *New York University Medical Center, New York, NY; †Cleveland Clinic, Cleveland, OH; ‡Indiana University, Indianapolis, IN; ‖The University of Michigan, Ann Arbor, MI; ¶Division of Urologic Pathology, Miraca Life Sciences Research Institute, Miraca Life Sciences, Irving; ††Houston Methodist Hospital, Houston, TX; #Emory University School of Medicine, Atlanta, GA; **University of Chicago; ‖‖Northwestern Medical Center, Chicago, IL; §Karolinska Institutet, Stockholm, Sweden; ‡‡Japanese Red Cross Nagoya Daini Hospital, Nagoya, Japan; and §§Calgary Laboratory Services and University of Calgary, Calgary, AB, Canada.

Conflicts of Interest and Source of Funding: The authors have disclosed that they have no significant relationships with, or financial interest in, any commercial companies pertaining to this article.

Correspondence: Ming Zhou, MD, PhD, Department of Pathology, New York University Medical Center, 560 First Avenue, TCH-461, New York, NY 10016 (e-mail: ming.zhou@nyumc.org).

glands, and cribriform glands with an irregular border, and hypernephromatoid cells).[10]

"Poorly formed glands" morphology has been recognized as one of the most common GP4 subpatterns in current practice. In a recent study of PCa diagnosed on prostate biopsy,[11] the "poorly formed glands" subpattern was identified in 57% of cases, followed by fused glands (53%), cribriform glands (25%), and the hypernephroid pattern (0.3%). The "fused glands" subpattern is relatively well defined[12] and has substantial to excellent interobserver agreement.[13] Practically all cribriform glands are now considered GP4.[5] The GP4 "poorly formed glands" subpattern, however, suffers definitional ambiguity. The 2005 ISUP consensus stated that "only a cluster of such glands, where a tangential section of Gleason pattern 3 glands cannot account for the histology, would be acceptable as Gleason pattern 4," but did not elaborate on the minimum number of "poorly formed glands" in a "cluster" and how the possibility of tangential sectioning can be ruled out with reasonable certainty. In a recent report studying the potential impact of reproducibility of Gleason grading in men being managed by active surveillance, such a definitional ambiguity may have accounted for the low interobserver reproducibility for the histologic distinction of tangentially sectioned GP3 from GP4 ($\kappa = 0.27$).[14]

Because of its prognostic and therapeutic implications, it is critical to accurately recognize the GP4 "poorly formed glands" subpattern and to distinguish it from GP3. However, several important issues pertaining to this particular GP4 subpattern have not been adequately addressed: (1) definition of "poorly formed glands"; (2) the diagnostic reproducibility of the GP4 "poorly formed glands" subpattern; (3) the quantitative threshold for grading a small focus of poorly formed glands as GP4; and (4) the criteria to distinguish GP4 "poorly formed glands" from tangentially sectioned GP3. The ultimate aim of this study was to identify histologic features that pathologists can use to reproducibly diagnose "poorly formed glands" as GP4.

MATERIALS AND METHODS

Twenty-nine cases of PCa that potentially contained "poorly formed glands" were selected from the pathology archives of Miraca Life Sciences and New York University Medical Center and were confirmed by 2 of the authors (R.B.S. and M.Z.). The selection of these cases was primarily based on morphologic criteria for GP4 "poorly formed glands" as outlined by the 2005 ISUP consensus.[10] In addition, some of the selected cases represented morphologies not described in the 2005 ISUP consensus but might be considered as GP4 "poorly formed glands." Overall, these cases ranged from being very likely to be considered as GP4 "poorly formed glands" to being borderline between GP4 "poorly formed glands" and tangentially sectioned GP3. Representative digital images of hematoxylin and eosin–stained slides from 29 cases were captured. All cases were represented by 1 image of $\times 100$ magnification, and 1 case was represented by an additional image of $\times 200$ magnification. The images were distributed electronically to 17 participating urologic pathologists.

The first 6 cases (images) contained cancer glands that represented a morphologic spectrum of glandular differentiation ranging from glands with well-formed lumens (Fig. 1A arrows #1), glands with well-formed but diminutive lumens that were smaller than the lining nuclei (Fig. 1A arrows #2), mixed well-formed and diminutive glands (Fig. 1B), glands with vague lumens (Fig. 1C), glands with no or rare discernible lumens (Fig. 1A arrows #3, D), elongated compressed glands (Fig. 1E), and elongated nests/small cords with no or rare lumens (Fig. 1F). Participants were only asked whether they would consider the cancer glands shown in the images as "poorly formed" without considering either the possibility of "tangential sectioning" nor rendering a Gleason score for the case.

For the next set of 23 cases, each participant was asked whether a component of GP4 "poorly formed glands" was present using criteria applied in the participants' respective clinical practice. In addition, if one felt that no GP4 "poorly formed glands" subpattern was present in a given case, he/she was asked to provide an alternative Gleason grade. Finally, each participant was asked about the number of tissue levels required to classify a small focus of poorly formed glands as GP4. To ensure the anonymity of their responses, all participants were de-identified and randomly assigned a number (1 to 17).

The second set of 23 cases was classified on the basis of the number of poorly formed glands in the focus (≤ 5, 6 to 10, and >10) and their topographic relation to the adjacent well-formed cancer glands. If poorly formed glands formed lobulated architecture without intermixing with other well-formed glands, they were considered "clustered" (Fig. 2). When poorly formed glands were intermixed with and immediately adjacent to (with <1 gland distance from) other well-formed glands, they were considered "adjacent to" well-formed cancer glands (Fig. 3). If poorly formed glands were intermixed with but were at least 1 gland distant from the adjacent well-formed cancer glands, they were considered "intermixed" with well-formed cancer glands (Fig. 4). Accordingly, these 23 cases were placed into 9 morphologic categories (Table 1).

The morphologic features (number and topographic location of poorly formed glands) were correlated with consensus diagnoses, which were defined as agreement of a diagnosis by at least 12 of 17 (70%) participants. This percentage was chosen because the probability of 70% of participants rendering the same diagnosis by chance is $<5\%$. Therefore, 1 of the following 3 diagnoses was assigned to each of 23 cases: *Consensus for GP4 "poorly formed glands"* (≥ 12 participants made the diagnosis of GP4 "poorly formed glands"); *Consensus against GP4 "poorly formed glands"* (≥ 12 participants opposed the diagnosis of GP4 "poorly formed glands"); and *No consensus* (< 12 participants agreed whether the case was

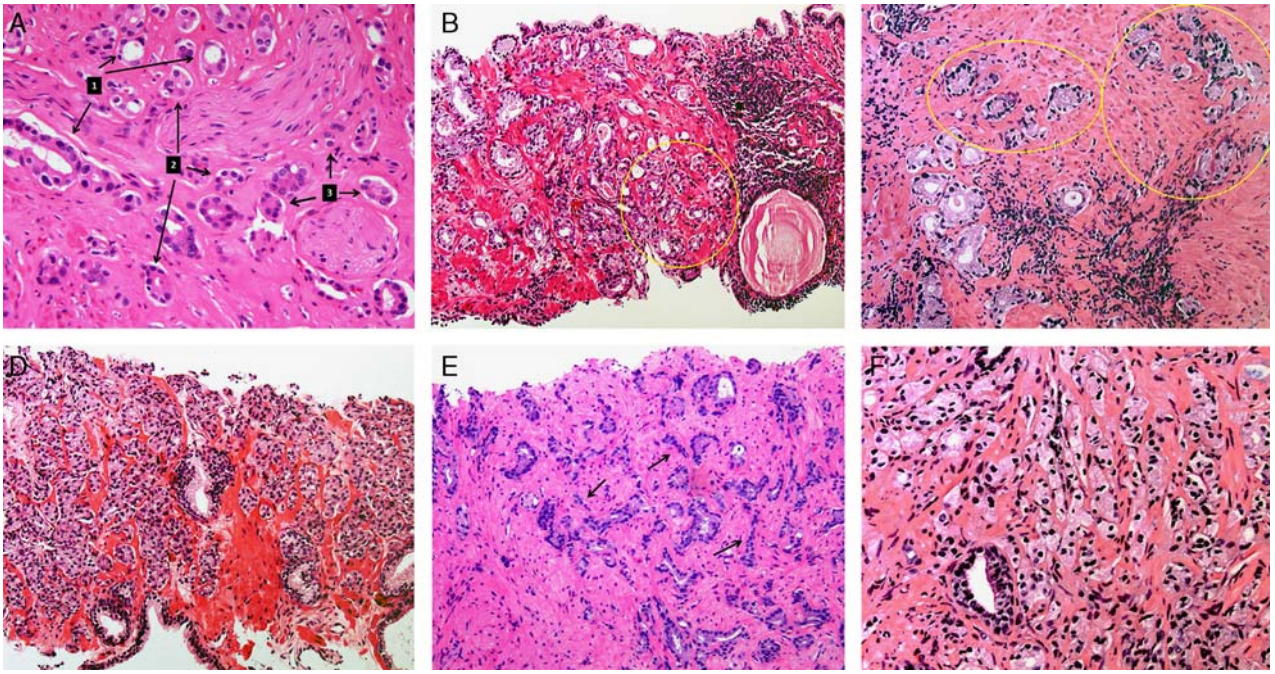

FIGURE 1. Prostate cancer glands with varying degree of glandular luminal differentiation, from glands with well-formed lumens (A arrows #1), glands with well-formed but diminutive lumens that are smaller than the lining nuclei (A arrows #2), mixed glands with well-formed and diminutive lumens (B, circle), glands with vague lumens (C, circles), glands with no or rare discernible lumens (A arrows #3 and D), elongated compressed glands (E, arrows), and elongated nests with no or rare lumens (F).

diagnostic for or against GP4 "poorly formed glands"). For cases that did not reach a consensus, the diagnosis for or against GP4 was "favored" when 10 or 11 of 17 participants agreed on a diagnosis. Statistical analyses were performed by 1 of the authors (J.L.) using R software for statistical computing. Interobserver reproducibility was assessed by Fleiss κ, implemented using the package "psy." Logistic regression analysis was performed to analyze the influence of quantity and topographic location of poorly formed glands with consensus diagnosis.

RESULTS

Definition of Poorly Formed Glands

Participants were presented with 6 images of discrete cancer glands with varying degree of glandular luminal differentiation and were asked whether they would consider any of the images as "poorly formed glands," without considering the possibility of tangential sectioning and actual GP. "Glands with well-formed lumens" (Fig. 1A #1) were regarded as "well formed" by consensus (Table 2). "Glands with no or only rare discernible lumens" (Figs. 1A #3, C), "elongated compressed glands" (Fig. 1E), and "elongated nests with no or rare lumens" (Fig. 1F) were considered "poorly formed" by consensus. Cancer glands with well-formed but diminutive lumens (Fig. 1A #2), mixed well-formed and diminutive glands (Fig. 1B), and glands with vague lumens (Fig. 1C) were not considered "poorly formed glands" by the majority of participants, but no consensus was attained. The unweighted κ value for the agreement of the definition of "poorly formed glands" was fair ($\kappa = 0.35$).

Grading "Poorly Formed Glands" as GP4

Table 3 summarizes the consensus diagnoses of 23 cases with potentially "poorly formed glands" by 17 urologic pathologists. Ten (43%) cases attained a consensus diagnosis for GP4 "poorly formed glands." However, a uniform (100%) diagnostic agreement was not achieved in any of the cases by the participants. Six cases achieved 94% diagnostic agreement (16 of 17 pathologists agreed on GP4). Six (26%) cases attained a consensus against GP4 "poorly formed glands," whereas the remaining 7 (31%) had no consensus either for or against GP4 "poorly formed glands." For the latter 7 cases in which a consensus was not reached, 4 were "favor GP4," and the remaining 3 were equivocal. The agreement for diagnosing GP4 "poorly formed glands" in these 23 cases was fair ($\kappa = 0.34$). When GP4 was considered not to be present, the alternative GP considered was 3. GP5 was not considered in any of the cases.

Correlation of Histologic Features With Consensus Diagnoses

Table 1 shows the correlation of histologic features (number and topographic location of poorly formed glands) with consensus diagnoses in 23 cases.

Correlation of the Number of Poorly Formed Glands With Consensus Diagnoses

Of 5 cases with ≤ 5 poorly formed glands, none attained a consensus for GP4 regardless of their topographic

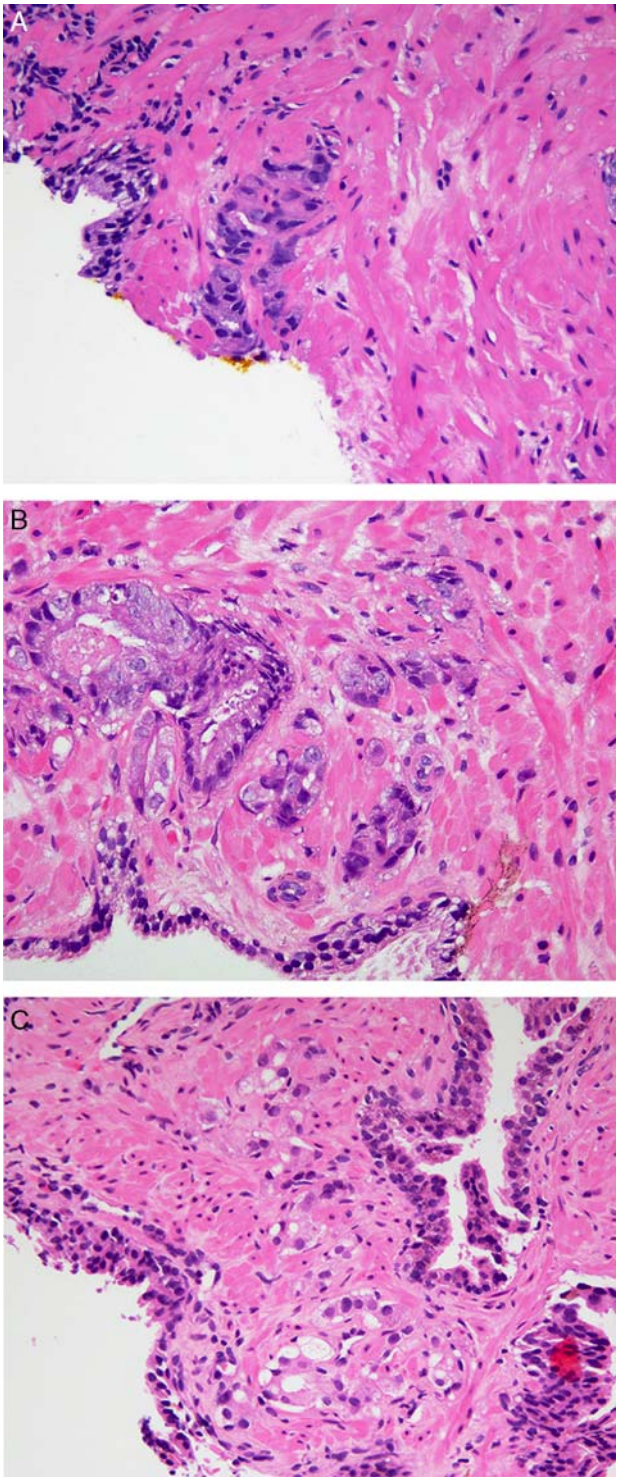

FIGURE 2. "Clustered" poorly formed glands. Cancer glands form lobulated architecture and do not intermix with other well-formed glands. A, ≤5 poorly formed glands. This case had no consensus diagnosis but was favored to be GP4 (10/17) among participants. B, 6 to 10 poorly formed glands. This case had a consensus of GP4 (15/17). C, >10 poorly formed glands. This case has a consensus of GP4 (16/17).

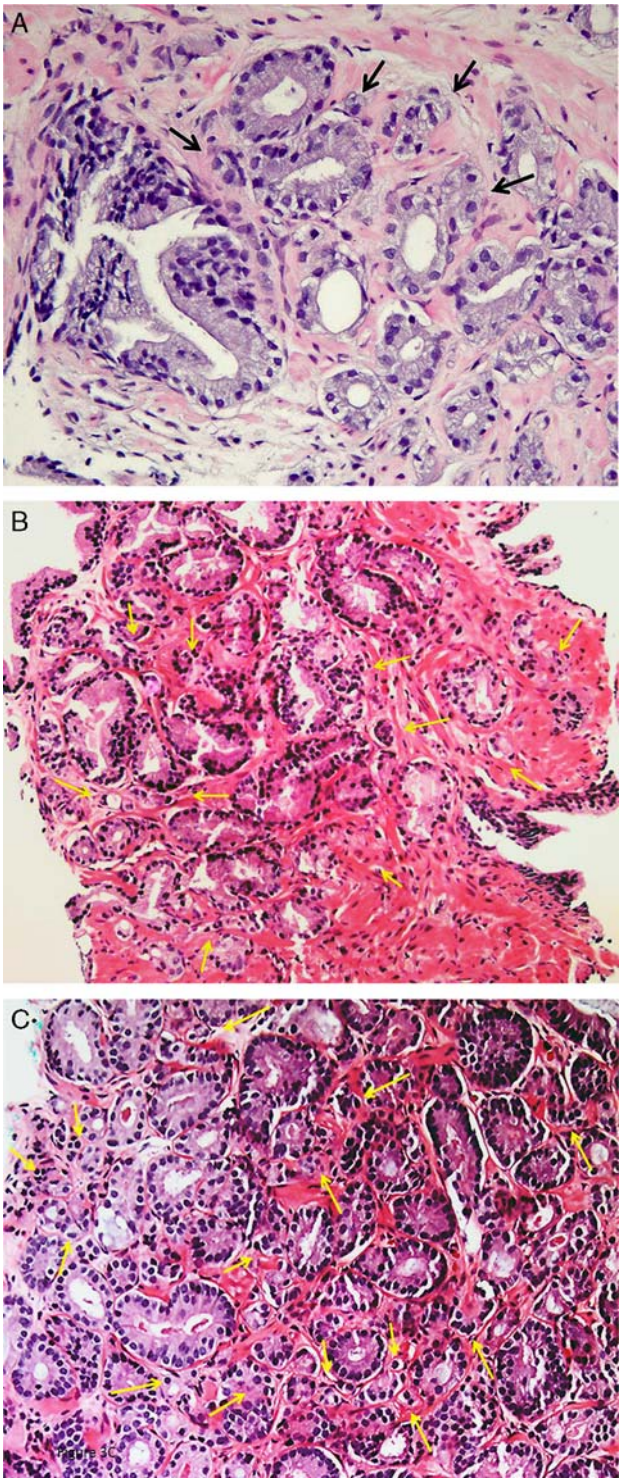

FIGURE 3. "Adjacent" poorly formed glands. Cancer glands are intermixed with and immediately adjacent to (with <1 gland distance from) other well-formed glands. A, ≤5 poorly formed glands. This case had a consensus against GP4 (14/17). B, 6 to 10 poorly formed glands. This case had a consensus against GP4 (16/17). C, >10 poorly formed glands. This case has a consensus against GP4 (13/17). Arrows denote poorly formed glands.

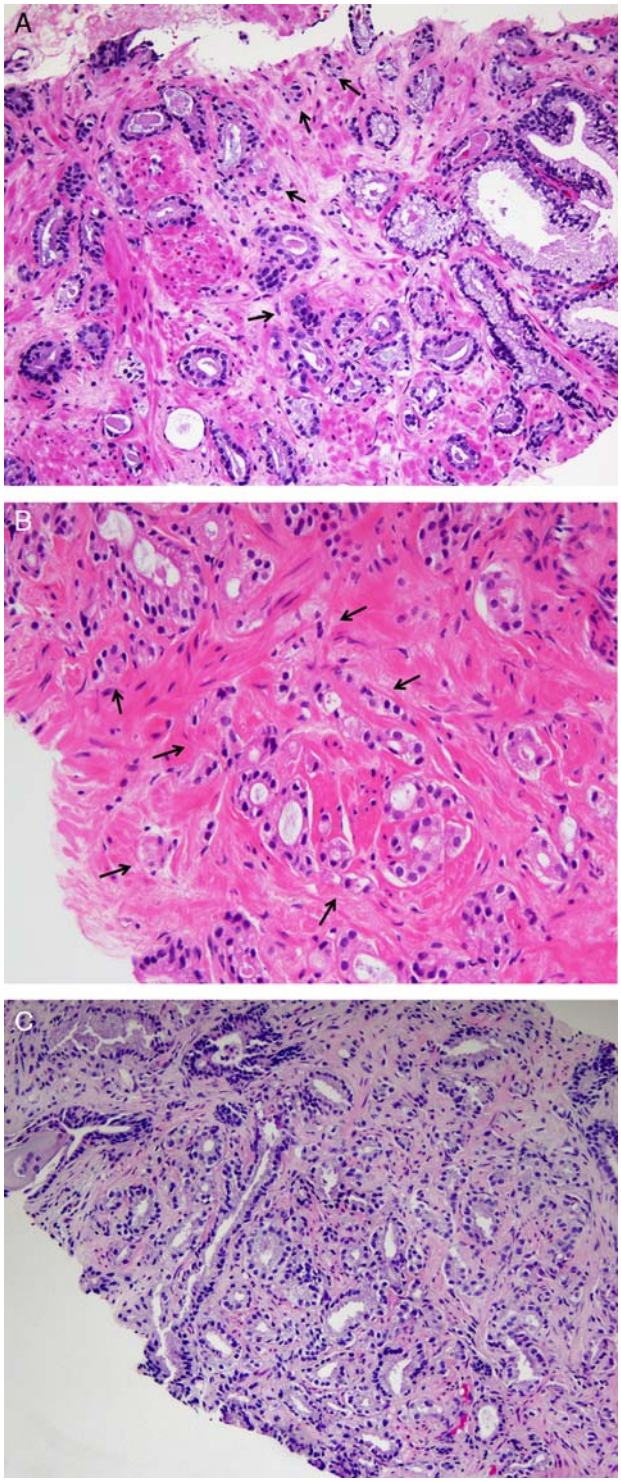

FIGURE 4. "Intermixed" poorly formed glands. Cancer glands are intermixed with but are at least 1 gland distant from the adjacent well-formed cancer glands. A, ≤5 poorly formed glands. This case had a consensus against GP4 (14/17). B, 6 to 10 poorly formed glands. This case had a consensus for GP4 (16/17). C, >10 poorly formed glands. This case has a consensus for GP4 (14/17). Arrows denote poorly formed glands.

locations (Fig. 5). A consensus against GP4 was reached in 4 cases, and no consensus was achieved in the fifth case. Cases with 6 to 10 poorly formed glands attained a consensus for GP4 in 4/9, a consensus against GP4 in 1/9, and no consensus in 4/9 cases. Cases with ≥10 poorly formed glands attained a consensus for GP4 in 6/9, a consensus against GP4 in 1/9, and no consensus in 2/9 cases.

Correlation of the Topographic Location With Consensus Diagnoses

Of 11 cases in which poorly formed glands were clustered together (clustered) (Fig. 2), consensus for GP4 was attained in 6, a consensus against GP4 was reached in 1, and there was no consensus in 4 cases. Of 5 cases in which poorly formed glands were intermixed with other well-formed cancer glands (intermixed) (Fig. 3), a consensus for GP4 was attained in 4 and a consensus against GP4 in 1. Of 7 cases in which poorly formed glands were adjacent to other well-formed glands (adjacent) (Fig. 4), none attained a consensus for GP4. A consensus against GP4 was, however, attained in 4, and there was no consensus in 3 cases. In particular, all 9 cases with >10 poorly formed glands with "clustered" and "intermixed" topographic location had a consensus for GP4.

In general, cases with more "poorly formed glands" that are located away from other adjacent well-formed glands are more likely to achieve a consensus for GP4 (Fig. 5). A case with a topographic location of "clustered" or "intermixed" and a quantity of ≥6 had an odd ratio of 15.9 (95% confidence interval, 1.9-131.9, $P = 0.01$) of being classified as GP4 when compared with a case with a topographic location of "adjacent" and of any quantity, or a case of ≤5 in quantity and of any topographic location.

The last 3 of 23 study cases, including 1 case of ≤5 clustered "poorly formed glands" and 2 cases of 6 to 10 clustered "poorly formed glands," did not attain a consensus, as the "poorly formed glands" were at the edge of the biopsy specimens. The majority of participants (15/17) would grade "poorly formed glands" as GP4 only when they retain "poorly formed glands" morphology in at least 2 levels. The overall agreement in grading 23 cases was fair ($\kappa = 0.34$).

DISCUSSION

The 2005 ISUP consensus on Gleason grading of prostatic carcinoma expanded the morphologic scope of GP4 to include an additional subpattern consisting of "ill-defined glands with poorly formed glandular lumina where tangential sectioning is ruled out,"[10] This "poorly formed glands" morphology, according to several studies, is now the most common GP4 subpattern in clinical practice, identified in 47% to 57% of prostate biopsies.[11,15] In radical prostatectomy specimens, the "poorly formed glands" subpattern was the exclusive component or was mixed with other GP4 subpatterns in 75.5% cases.[16] Grading "poorly formed" and cribriform glands as

TABLE 1. Categorization of 23 Study Cases Based on the Number and Topographic Location of Poorly Formed Glands and Their Consensus Diagnoses

	Topographic Location (Relation With Adjacent Well-formed Cancer Glands)			
	Clustered	Adjacent	Intermixed	Total
No. poorly formed glands				
≤5	2 (0, 1, 1)*	2 (0, 2, 0)	1 (0, 1, 0)	5 (0, 4, 1)
6-10	5 (2, 0, 3)	2 (0, 1, 1)	2 (2, 0, 0)	9 (4, 1, 4)
>10	4 (4, 0, 0)	3 (0, 1, 2)	2 (2, 0, 0)	9 (6, 1, 2)
Total	11 (6, 1, 4)	7 (0, 4, 3)	5 (4, 1, 0)	23 (10, 6, 7)

*Numbers in parenthesis (N1, N2, N3): N1—consensus for GP4; N2—consensus against GP4; N3—no consensus.

GP4 and inclusion of high-grade PCa of minute amount in the final Gleason score have led to a significant upward shift of Gleason score in prostate biopsies.[17–19] For example, Helpap and Egevad[18] showed that the number of biopsies diagnosed as GS6 decreased from 48% to 22%, whereas the number of biopsies diagnosed as GS7 increased from 26% to 68%, after adopting the ISUP modified Gleason grading system. Such a shift in Gleason grade on biopsies has significantly impacted its prognostic value. Pierorazio et al[20] showed that GS6 PCa diagnosed on biopsies has excellent prognosis despite sampling error and potential upgrading at radical prostatectomy, with a 5-year biochemical recurrence-free survival of 94.7% (vs. 82.7% for GS7). More importantly, the therapeutic modalities are different for GS6 and GS7 PCa, with the latter in general being ineligible for active surveillance and focal therapy in the vast majority of practices.[5]

Given the critical prognostic and therapeutic difference between GP3 and GP4 PCa, as well as the very frequent presence of "poorly formed glands" as a GP4 component in prostate biopsies, it is critical to investigate the diagnostic reproducibility and the histologic features that are associated with reproducible grading of this GP4 subpattern. However, most of the previous studies on grading reproducibility did not separate various GP4 subpatterns.[13,15,21–27] Only 1 study, conducted by McKenney et al,[14] looked specifically at the interobserver reproducibility for the histologic distinction of tangentially sectioned GP3 from GP4 in 17 PCa patients who were managed by active surveillance. To our knowledge, no study so far has attempted to identify the histologic features of "poorly formed glands" that are reproducibly associated with grading "poorly formed glands" as GP4. This current study set out to study: (1) the definition of "poorly formed glands"; (2) the diagnostic reproducibility of the GP4 "poorly formed glands" subpattern; (3) the quantitative threshold for grading a small focus of poorly formed glands as GP4; and (4) the criteria to distinguish GP4 "poorly formed glands" from tangentially sectioned GP3.

To clarify the definition of "poorly formed glands," we queried the participants with PCa glands exhibiting a spectrum of glandular differentiation ranging from sharply delineated lumens to well-formed but diminutive lumens (luminal diameter smaller than the nuclei of the lining nuclei) to glands with vague or ill-defined lumens, and finally to glands with no or rare luminal formation (Fig. 1). The κ value for the agreement on the definition of "poorly formed glands" is fair (0.35), suggesting a significant definitional variation of "poorly formed glands" among the participating urologic pathologists. The lack of substantial agreement on the definition of "poorly formed glands" among urologic pathologists is perhaps understandable, as the glandular luminal differentiation represents a morphologic continuum. Binary division of cancer glands into "well formed" and "poorly formed" categories is rather arbitrary. In this study,

TABLE 2. Morphologic Definition of "Poorly Formed Glands"

Image	Morphology	Diagnosis as "Poorly Formed Glands"
Figure 1A #1	Cancer glands with well-formed lumens	No (15/17)*
Figure 1A #2	Cancer glands with well-formed but diminutive lumens that were smaller than the lining nuclei	No (11/17)
Figure 1A #3	Cancer glands with no lumens	Yes (13/17)*
Figure 1B	Mixed glands with well-formed and diminutive lumens	No (10/17)
Figure 1C	Cancer glands with vague lumens	No (9/17)
Figure 1D	Cancer glands with no or only rare discernible lumens	Yes (17/17)*
Figure 1E	Elongated compressed glands	Yes (13/17)*
Figure 1F	Elongated nests with no or rare lumens	Yes (16/17)*

*Attained consensus.

TABLE 3. Consensus Diagnosis of 23 Cases With "Poorly Formed Glands" Among 17 Participants

Consensus Diagnosis	Group Results Number
For GP4	10/23 (43%)
100% participants agreement	0
≥70% participants agreement	10
Against GP4	6/23 (26%)
100% participants agreement	0
≥70% participants agreement	6
No consensus	7/23 (31%)
Favor GP4	4
Favor against GP4	0
Equivocal	3

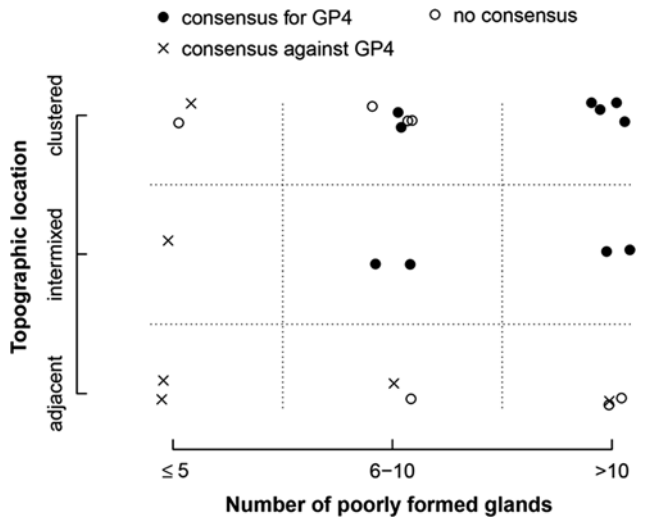

FIGURE 5. A graphic representation demonstrating how quantitative and qualitative features of poorly formed glands influenced the consensus diagnoses.

however, we established by consensus that the "glands with no or only rare discernible lumens" should be considered "poorly formed." In addition, elongated compressed glands without discernible lumens, a pattern previously not described (Fig. 1E), and elongated nests/small cords with no or rare lumens, were also considered "poorly formed glands" by consensus. It is equally important to point out that there was no consensus when glands were less well-differentiated, including glands with well-formed but diminutive lumens, glands with vague or ill-defined lumens and mixed glands with well-formed and poorly formed lumens. These findings suggest that urologic pathologists use a conservative approach and default to lower grade (GP3) when cancer glands are considered borderline between well and poorly formed. We therefore recommend that only glands with no or rare lumen formation, elongated compressed glands, and elongated nests or small cords with no or rare lumens be considered "poorly formed." Glands with diminutive or vague lumens should not be considered "poorly formed."

It is also important to remember that not all "poorly formed glands" should be graded as GP4, as a minute focus of poorly formed glands may not be quantitatively sufficient to grade as GP4, and one has to rule out the possibility of tangentially sectioned GP3. The distinction between GP4 "poorly formed glands" and tangentially sectioned GP3 is often difficult and arbitrary. A recent study found that the interobserver reproducibility for the histologic distinction of tangentially sectioned GP3 from GP4 was low ($\kappa = 0.27$); however, it did not investigate which morphologic features contributed to such grading ambiguity.[14] The current study, in contrast, attempted to address this issue by looking at 2 histologic features, that is, the number and topographic location of "poorly formed glands," that may help distinguish GP4 "poorly formed glands" from tangentially sectioned GP3.

Poorly formed glands as the result of tangential sectioning are expected to be immediately adjacent to the well-formed "mother" glands from which the former are derived. In contrast, many poorly formed glands located away from other well-formed glands are unlikely to result from tangential sectioning. In this study, 7 cases were categorized as "adjacent to" (poorly formed glands are immediately adjacent to [with <1 gland distance from]) other well-formed glands. Four cases had a consensus against GP4, and 3 cases had a diagnosis of "favor a diagnosis against GP4." No case had a consensus for GP4. In contrast, of 11 cases that were categorized as "clustered" (poorly formed glands that form a lobulated architecture without intermixed well-formed glands) and 5 cases that were categorized as "intermixed" (poorly formed glands are intermixed with, but at least 1 gland away from, the adjacent well-formed cancer glands), a consensus for GP4 was made in 10/16 cases. In particular, all 6 cases that were "clustered" or "intermixed" and had >10 poorly formed glands had a consensus for GP4. This would suggest that any amount of "poorly formed glands" that are immediately adjacent to other well-formed glands regardless of their number should be considered the result of tangential sectioning and graded as GP3 instead of GP4.

The minimum number of poorly formed glands in a prostate biopsy that can be graded as GP4 "poorly formed glands," in a reproducible manner, has not been firmly established. In this study, of 5 cases with ≤ 5 poorly formed glands, none had a consensus for GP4, and in 4 cases there was a consensus against GP4. In contrast, 7 cases with 6 to 10 poorly formed glands categorized as "clustered" or "intermixed" had a consensus diagnosis for GP4 in 4 cases and "favor GP4" in 3 cases. All 6 cases with >10 poorly formed glands categorized as "clustered" or "intermixed" had a consensus for GP4. These results suggest that the minimum number of "poorly formed glands" that can be graded as GP4 is 6. Any case with ≤ 5 poorly formed glands, regardless of the topographic location, should not be graded as GP4. Any case with >10 poorly formed glands, as long as the poorly formed glands are not immediately adjacent to other well-formed glands, can be confidently graded as GP4. Histologic features that are "diagnostic of" and "against" GP4 "poorly formed glands" are summarized in Table 4.

Even though we identified several histologic features that urologic pathologists rely on to diagnose GP4

TABLE 4. Histologic Features That Are Diagnostic of and Against GP4 Poorly Formed Glands by Urologic Pathologists

Histologic features that are "diagnostic of" GP4 "poorly formed glands"
>10 poorly formed glands that are not immediately adjacent to other well-formed glands
Histologic features that are "against" GP4 "poorly formed glands"
Poorly formed glands intermixed with and immediately adjacent to (with <1 gland distance from) well-formed glands regardless of their number
≤ 5 poorly formed glands regardless of their location

"poorly formed glands" and to distinguish them from tangentially sectioned GP3, the reproducibility of grading GP4 "poorly formed glands" among 17 participants was fair ($\kappa = 0.34$). This finding suggests that a significant variation in grading GP4 "poorly formed glands" among urologic pathologists, which is inherent due to the subjective nature of morphologic evaluation. In fact, none of the morphologic parameters achieved uniform consensus for or against GP4 among all participants. Similarly, tangentially sectioned GP4 poorly formed glands may be misinterpreted as individual cells or infiltrating cords of GP5 specifically when the focus of concern is quantitatively small (Shah et al, unpublished data), resulting in erroneous Gleason grade inflation. It is therefore logical to use a conservative approach and a higher diagnostic threshold to diagnose GP4 and GP5 when there is grading uncertainty due to potential tangential sectioning or when the focus of concern is small. Multiple sections should be reviewed. In this study, the majority of participants (15/17) would grade poorly formed glands as GP4 only when they retained poorly formed morphology in at least 2 levels. Sharing difficult cases with colleagues is encouraged to reduce grading uncertainty. If uncertainty still remains, however, it is recommended to default to GP3.[28] This conservative approach is supported by several recent studies showing potentially different prognostic significance of different GP4 subpatterns. In a study to compare the prognostic value of the classical and modified Gleason scores using posttreatment PSA nadir as the endpoint in a cohort of PCa patients treated with radiation and hormonal ablation, Delahunt and colleagues found that the classical Gleason grading system provided slightly better prognostic discrimination than the modified system. More importantly, there was no significant difference between modified Gleason scores 6 and 7[29] in terms of prognostic discrimination, suggesting that in the modified Gleason grading system GP3 and GP4 seemed to overlap in their survival characteristics. Several other studies have demonstrated that subpatterns of GP4 in radical prostatectomy specimens had different prognostic implications. The cribriform subpattern predicted worse clinical outcomes than other subpatterns such as "poorly formed" and "fused glands" subpatterns.[16,30,31] It is conceivable that specific subpatterns of GP4 in a PCa with Gleason score 7 in prostate biopsy may also have similar prognostic implications and therefore merits documentation in biopsy reports.

We acknowledge that the current study has several limitations. First, this study was conducted from a set of highly selective images representing difficult-to-classify areas in the spectrum of GP4 "poorly formed glands" rather than cases routinely seen in practice. As a result, the interobserver reproducibility of grading "poorly formed glands" as GP4 was only fair, and the κ (0.34) is lower than those reported in the previous studies on interobserver reproducibility of Gleason grading many of which utilized consecutive cases encountered in daily practice.[13,15,21–27] Unlike other ones, this study investigated the diagnostic reproducibility of grading GP4 "poorly formed glands," a subpattern that has been shown to have lower diagnostic reproducibility than other GP4 subpatterns.[9,19] Regardless, this study highlights the difficulty and problem areas in grading GP4 "poorly formed glands." Another potential limitation is the correlation of various histologic features of poorly formed glands with the consensus diagnoses, rather than correlation with clinical outcomes. Nonetheless, the consensus diagnosis was correlated with objective morphologic criteria (number and topographic location of poorly formed glands). Despite these limitations, this study demonstrated that certain morphologic features are reproducible for grading GP4 "poorly formed glands." We hope that the findings of this study represent an important step in standardization of grading of "poorly formed glands" based on quantitative and topographic morphologic features.

In summary, the overall reproducibility for grading GP4 "poorly formed glands" was fair among a group of urologic pathologists. However, a set of histologic criteria was reproducibly associated with the diagnosis of GP4 "poorly formed glands." We recommend that only cancer glands with no or rare luminal formation be considered "poorly formed." Furthermore, "poorly formed glands" intermixed with and immediately adjacent to other well-formed glands regardless of their number and ≤ 5 "poorly formed glands" regardless of their location should not be graded as GP4. In contrast, > 10 "poorly formed glands" that are not immediately adjacent to well-formed glands are considered GP4. When there is uncertainty whether a focus of "poorly formed glands" represent GP4 or GP3, we recommend use of a conservative approach with default to GP3.

REFERENCES

1. Bostwick DG, Grignon DJ, Hammond ME, et al. Prognostic factors in prostate cancer. College of American Pathologists Consensus Statement 1999. *Arch Pathol Lab Med*. 2000;124:995–1000.
2. Gleason DF, Mellinger GT. Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging. *J Urol*. 1974;111:58–64.
3. Mellinger GT. Prognosis of prostatic carcinoma. *Recent Results Cancer Res*. 1977;60:61–72.
4. Shah RB. Current perspectives on the Gleason grading of prostate cancer. *Arch Pathol Lab Med*. 2009;133:1810–1816.
5. Brimo F, Montironi R, Egevad L, et al. Contemporary grading for prostate cancer: implications for patient care. *Eur Urol*. 2013;63:892–901.
6. Ellis CL, Partin AW, Han M, et al. Adenocarcinoma of the prostate with Gleason score 9-10 on core biopsy: correlation with findings at radical prostatectomy and prognosis. *J Urol*. 2013;190:2068–2073.
7. Dall'Era MA, Albertsen PC, Bangma C, et al. Active surveillance for prostate cancer: a systematic review of the literature. *Eur Urol*. 2012;62:976–983.
8. Gleason DF. Classification of prostatic carcinomas. *Cancer Chemother Rep*. 1966;50:125–128.
9. Gleason DF. Histological grading and clinical staging of prostatic carcinoma. In: Tannenbaum M, ed. *Urologic Pathology*. Philadelphia: Lea & Feibiger; 1977:171–198.
10. Epstein JI, Allsbrook WC Jr, 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;29:1228–1242.
11. Gottipati S, Warncke J, Vollmer R, et al. Usual and unusual histologic patterns of high Gleason score 8 to 10 adenocarcinoma of

the prostate in needle biopsy tissue. *Am J Surg Pathol*. 2012;36: 900–907.
12. Eble J, Sauter G, Epstein JI, et al. *World Health Organization Classification of Tumors: Pathology and Genetics of Tumours of the Urinary System and Male Genital Organs*. Lyons: IARC; 2004.
13. Helpap B, Kristiansen G, Beer M, et al. Improving the reproducibility of the Gleason scores in small foci of prostate cancer—suggestion of diagnostic criteria for glandular fusion. *Pathol Oncol Res*. 2012;18:615–621.
14. McKenney JK, Simko J, Bonham M, et al. The potential impact of reproducibility of Gleason grading in men with early stage prostate cancer managed by active surveillance: a multi-institutional study. *J Urol*. 2011;186:465–469.
15. Egevad L, Algaba F, Berney DM, et al. Interactive digital slides with heat maps: a novel method to improve the reproducibility of Gleason grading. *Virchows Arch*. 2011;459:175–182.
16. Dong F, Yang P, Wang C, et al. Architectural heterogeneity and cribriform pattern predict adverse clinical outcome for Gleason grade 4 prostatic adenocarcinoma. *Am J Surg Pathol*. 2013;37:1855–1861.
17. Danneman D, Drevin L, Robinson D, et al. Gleason inflation 1998-2011. A registry study of 97 168 men. *BJU Int*. 2014;115:248–255.
18. Helpap B, Egevad L. The significance of modified Gleason grading of prostatic carcinoma in biopsy and radical prostatectomy specimens. *Virchows Arch*. 2006;449:622–627.
19. Zareba P, Zhang J, Yilmaz A, et al. The impact of the 2005 International Society of Urological Pathology (ISUP) consensus on Gleason grading in contemporary practice. *Histopathology*. 2009;55: 384–391.
20. Pierorazio PM, Walsh PC, Partin AW, et al. Prognostic Gleason grade grouping: data based on the modified Gleason scoring system. *BJU Int*. 2013;111:753–760.
21. Allsbrook WC Jr, Mangold KA, Johnson MH, et al. Interobserver reproducibility of Gleason grading of prostatic carcinoma: urologic pathologists. *Hum Pathol*. 2001;32:74–80.
22. Allsbrook WC Jr, Mangold KA, Johnson MH, et al. Interobserver reproducibility of Gleason grading of prostatic carcinoma: general pathologist. *Hum Pathol*. 2001;32:81–88.
23. Carlson GD, Calvanese CB, Kahane H, et al. Accuracy of biopsy Gleason scores from a large uropathology laboratory: use of a diagnostic protocol to minimize observer variability. *Urology*. 1998; 51:525–529.
24. Egevad L. Reproducibility of Gleason grading of prostate cancer can be improved by the use of reference images. *Urology*. 2001;57: 291–295.
25. Lessells AM, Burnett RA, Howatson SR, et al. Observer variability in the histopathological reporting of needle biopsy specimens of the prostate. *Hum Pathol*. 1997;28:646–649.
26. Melia J, Moseley R, Ball RY, et al. A UK-based investigation of inter- and intra-observer reproducibility of Gleason grading of prostatic biopsies. *Histopathology*. 2006;48:644–654.
27. Rodriguez-Urrego PA, Cronin AM, Al-Ahmadie HA, et al. Interobserver and intraobserver reproducibility in digital and routine microscopic assessment of prostate needle biopsies. *Hum Pathol*. 2011;42:68–74.
28. Amin MB, Lin DW, Gore JL, et al. The critical role of the pathologist in determining eligibility for active surveillance as a management option in patients with prostate cancer: consensus statement with recommendations supported by the College of American Pathologists, International Society of Urological Pathology, Association of Directors of Anatomic and Surgical Pathology, the New Zealand Society of Pathologists, and the Prostate Cancer Foundation. *Arch Pathol Lab Med*. 2014;138:1387–1405.
29. Delahunt B, Lamb DS, Srigley JR, et al. Gleason scoring: a comparison of classical and modified (international society of urological pathology) criteria using nadir PSA as a clinical end point. *Pathology*. 2010;42:339–343.
30. Iczkowski KA, Torkko KC, Kotnis GR, et al. Digital quantification of five high-grade prostate cancer patterns, including the cribriform pattern, and their association with adverse outcome. *Am J Clin Pathol*. 2011;136:98–107.
31. Kweldam CF, Wildhagen MF, Steyerberg EW, et al. Cribriform growth is highly predictive for postoperative metastasis and disease-specific death in Gleason score 7 prostate cancer. *Mod Pathol*. 2015; 28:457–464.