Review Article | Iran J Pathol. 2016; 11(2): 120 - 126

http://www.ijp.iranpath.org/

Correlation between Gleason Scores in Needle Biopsy and Corresponding Radical Prostatectomy Specimens: A Twelve-Year Review

Maliheh Khoddami[1], Yassaman Khademi[2], Maryam Kazemi Aghdam[1], Haleh Soltanghoraee[3]

1. Pediatric Pathology Research Center, Shahid Beheshti University of Medical Sciences Tehran, Iran

2. APCP, Pathobiology Laboratory Center, Tehran, Iran

3. APCP, Reproductive Biotechnology Research Center, Avicenna Research Institute, ACECR, Tehran, Iran

KEY WORDS

Adenocarcinoma
Gleason
Overgrading
Prostatic neoplasms
Undergrading

ARTICLE INFO

Received 02 Apr 2015;
Accepted 10 Oct 2015;

ABSTRACT

***Background*:** Presence of discordance between the Gleason score on needle biopsy and the score of radical prostatectomy specimen is common and universal. In this study, we determined the accuracy of Gleason grading of biopsies in predicting histological grading of radical prostatectomy specimens and the degree of overgrading and undergrading of prostatic adenocarcinoma in our center, which is one of the referral centers in Tehran.

***Methods*:** In this retrospective study, we analyzed the results of prostate needle biopsies and subsequent prostatectomies diagnosed at the Pathobiology Laboratory Center, Tehran, Iran in 45 patients between 2002 and 2013. Preoperative clinical data and the information from biopsy and prostatectomy specimens were collected. The accuracy, sensitivity, specificity, and positive and negative predictive values of different grades and groups were assessed. Pearson and Spearman correlation coefficient were used to determine the relation of different variables.

***Results*:** The biopsy Gleason score was identical to the scores in prostatectomy specimens in 68.2% cases, while 31.8% were discrepant by 1 or 2 Gleason score. We had 9.1% downgrading and 22.7% cases upgraded after prostatectomy. The sensitivity and positive predictive value was 86% and 79% for low grade, 67% and 75% for moderate grade, and 80% and 80% for high-grade tumors, respectively.

***Conclusion*:** Overall, the reliability of Gleason grading of needle biopsies in predicting final pathology was satisfavory. Moderate grade group was the most difficult to diagnose in needle biopsy.

©Iran J Pathol. All rights reserved.

***Corresponding Information**:* Dr. Yassaman Khademi, MD, Pathobiology Laboratory Center, Tehran, Iran. Email: yassamankhademi@yahoo.com Tel: +98-21-22666563

Copyright © 2016, IRANIAN JOURNAL OF PATHOLOGY. This is an open-access article distributed under the terms of the Creative Commons Attribution-noncommercial 4.0 International License which permits copy and redistribute the material just in noncommercial usages, provided the original work is properly cited.

Introduction

The most common malignancy of the prostate is adenocarcinoma, accounting for more than 25% of all malignancies in male. Over 90% of prostatic carcinomas are conventional acinar adenocarcinoma, the majority of which are multifocal (60%-90%) (1). After more than four decades, the Gleason score (GS) still remains the most widely accepted grading system in the evaluation

of prostatic adenocarcinoma (2). Gleason grading system is based on a morphologic continuum of architectural dedifferentiation and is directly correlated with response to therapy, mortality rate, as well as a predictor of time to recurrence. Multifocality is common in prostate cancer and is expected to see different grades in different foci and/or within the same focus of prostatic carcinoma (3, 4). This is one of the sources of intra observer variability in reporting and the common discrepancy between the GS needle biopsy (NB) and radical prostatectomy (RP) specimens with exact match in only 41 to 43% in large series (4, 5). When the treatment decision is based only on the prostatic NB specimen, the accuracy of NB GS is of outmost importance.

We evaluated the degree of the discrepancies between GS on NB and RP in the Pathobiology Laboratory Center, which is one of the large private referral centers for evaluation of prostatic tissue in Tehran, Iran.

Materials and methods

The files of Pathobiology Laboratory Center, Tehran, Iran one of the referral laboratories for evaluation of prostatic tissue, were searched for patients who had prostate needle biopsy and subsequent prostatectomy during a period of 12 years (2002 -2013). Patients who had only needle biopsy or prostatectomy specimens were excluded from the study. Biopsies were taken by different physicians using thin-core (18-gauge) needles used in conjunction with biopsy guns attached to transrectal ultrasound. Sextant, octant, or other modes of systematic sampling were performed.

The entire received specimen was embedded in different capsules according to the labels and was paraffin embedded. At least three slides were prepared from each specimen (at least a total of 10-12 step sections of 2-3 μ thick tissue) and stained by Hematoxylin and Eosin (H&E).

The prostatectomy specimens were examined grossly by a staff pathologist, divided into right and left as well as anterior and posterior quandrants. Each quadrant, then, was serially sectioned at approximately 4-5 mm intervals from apex to base, examined for the presence of tumor. Slices with grossly identifiable tumor were submitted. If no tumor was grossly detected, every other slice was submitted. The proximal urethral and vas deferentia margins as well as the entire distal urethra and seminal vesicles, if present, were also embedded for routine processing.

The slides were examined under the light microscope by one or more general pathologists. Immunohistochemical stains were used in suspect cases. Biopsies and prostatectomy specimens were graded by different staff pathologists or by a single pathologist and the final gradings were made regardless of the NB results. The specimens were graded using Gleason's grading system, which is based on histologic patterns; i.e. "extent of glandular differentiation and the pattern of growth of tumor in prostatic stroma" and GS, defined as the sum of the primary and secondary predominant patterns, were determined (2). GS of 2 to 6 is considered low grade, 7 as moderate grade, and 8 to 10 as high grade.

Tumor within the prostatic capsule was designated as organ confined and tumor at the inked resection margin was considered margin positive. Presence or absence of vascular and/or perineural invasion, as well as extraprostatic involvement were also evaluated.

Preoperative clinical data, including patients' age, PSA, free PSA, free PSA/ total PSA ratio and PSAD were also collected, as well as the information from biopsy and prostatectomy specimens (Gleason's primary and secondary tumor grades, scores, capsular and perineural invasion, extraprostatic extension, prostatic weight).

A comparison was made between the Gleason grades and scores of the needle biopsy and prostatectomy specimens. Because the prostatectomy GS is based on examination of the entire gland and reflects the underlying biology of the disease more accurately, it is considered as "gold standard".

The accuracy, sensitivity, specificity, and positive (PPV) and negative predictive values (NPV) of different grades and groups were assessed. Pearson and Spearman correlation coefficient were used to determine the relation of different variables. P value >0.05 was considered statistically insignificant and P value ≤ 0.05 was considered as statistically significant.

Results

Forty-five patients with mean age of 64.09 yr (49 to 74 years) were included in this study. In prostatectomy specimens, most patients had low-grade tumors (22 cases, 48.8%), 18 had moderate grade (40%), and 5 were high grade (17.7%). The frequency of first and second dominant grades and different scores in prostatectomy specimens are shown in Tables 1-3. Grade 3 and score 7 were the most frequent grade and score.

In the first grading, five cases changed from 4 to 3, one case from 5 to 4 (total of 6 down grading); four cases changed from 2 to 3, two cases from 3 to 4 and one case from 2 to 4 (total of 7 up grading) in prostatectomy specimens. In the second grading, four changed from 4 to 3, and one form 3 to 2 (total of 5 down grading); one changed from 2 to 3, eight from 3 to 4, and three from 4 to 5 (total of 12 up grading) in prostatectomy specimens. Change from 3 to 4 was the most frequent change, followed by change from 4 to 3. The sensitivity, specificity, PPV, NPV and accuracy of different grades are shown in Table 4.

Down grading in score happened from 7 to 6

Table 1
Frequency of first dominant grade in prostatectomy specimens

Frequency	First Grade	Percent
2	2	4.4
35	3	77.8
8	4	17.8
45	Total	100.0

Table 2
Frequency of second dominant grade in prostatectomy specimens

Percent	Frequency	Score
15.6	7	2
42.2	19	3
31.1	14	4
11.1	5	5
100.0	45	Total

Table 3
Frequency of different scores in prostatectomy specimens

Percent	Frequency	Score
4.4	2	4
11.1	5	5
33.3	15	6
40.0	18	7
2.2	1	8
8.9	4	9
100.0	45	Total

Table 4
Sensitivity, specificity, PPV, NPV and accuracy of different grades

Grade	Sensitivity (percent)	Specificity (percent)	PPV (percent)	NPV (percent)	Accuracy (percent)
2	89	92	57	98	92
3	73	74	79	67	73
4	46	81	43	82	72
5	60	98	67	96	95

in three cases and from 8 to 7 in one case (one score change). Upgrading in scores were as follows: 8 to 9 in one case, 7 to 8 in one, 6 to 7 in three, 5 to 7 in two, 5 to 6 in two, and 4 to 5 in one case (one score change in 8 cases and 2 score change in 2 cases). Most patients had low-grade tumors (22 cases, 48.8%), 18 had moderate grade (40%), and 5 were high grade (17.7%). Group changed from high to moderate in one case, and from moderate to low in three cases. In upgraded group, four had no change in group (3 low grades and 1 high grade), five moved from low to moderate and one from moderate to high group.

The biopsy score was identical to the prostatectomy specimen score in 30 (68.2%) of cases, while 14 (31.8%) were discrepant by 1 or 2 Gleason scores. Overall, 9.1% of biopsies were overgraded, while 22.7% were undergraded. In prostatectomy specimens, Group change from high to moderate happened in one case, and from moderate to low in three cases. In upgraded group, four had no change in group (3 low grades and 1 high grade), five moved from low to moderate and one from moderate to high group. The sensitivity, specificity, PPV, NPV and accuracy of different groups are shown in Table 5.

Capsular invasion and/or extra prostatic extension was seen in 4 of 5 high grade tumors, in 9 of 18 moderate grade and 9 in 22 low grade tumors, which indicates a positive relation between tumor score and capsular invasion and/or extra prostatic extension (*P*-value= 0.009). No meaningful statistically correlation was present between age, PSA level, PASD, prostatic weight and tumor score (*P*-value= 0.099, 0. 541, 0.857, and 0.227, respectively).

Discussion

The Gleason grading system remains the most widely used grading system and one of the most important prognostic predictors for prostatic adenocarcinoma. Relying on the biopsy grade to make clinical management decisions needs a high degree of correlation between histological grading of the biopsy and RP specimens. The accuracy of NB GS to predict the GS in prostatectomy varies tremendously in the literature and discrepancy is universal (4, 6-16).

Table 5
Sensitivity, specificity, PPV, NPV and accuracy of different groups

Grade	Sensitivity (percent)	Specificity (percent)	PPV (percent)	NPV (percent)	Accuracy (percent)
Low	86	78	79	85	82
Moderate	67	85	75	79	78
High	80	97	80	97	98

Perineural invasion was present in 41 (91.1%) cases (19 of 22 low grade, 16 of 18 moderate grade and 5 of 5 high grade tumors). Tumor score was associated with perineural invasion (statistically significant, *P*-value= 0.032).

Vascular invasion was identified in 3 of 5 high grade tumors, in 1 of 18 moderate grade and none in 22 low grade tumors, which indicates a positive relation between tumor score and vascular invasion (*P*-value= 0.00).

Humphrey (5) reviewed 18 articles, including 3789 patients, and found 43% exact correlation, and 77% correlation plus or minus one Gleason score. We had identical scores in 68.2% cases, while 31.8% were discrepant by 1 or 2 Gleason scores.

Undergrading of carcinoma in needle biopsy is more common than overgrading; undergrading

was reported in 42-48% and overgrading was 15-17% (3-5)The analysis including14839 patients in 16 studies from 6 different countries demonstrated overall accuracy of only 63%. Upgrading was reported in 30% and downgrading in 7% of patients (6). Our data showed 22.7% upgrading, and 9.1% downgrading; 68.2% were accurately predicted.

Low, moderate, and high-grade tumors on biopsy do not have the same accuracy; high-grade cancer was the least accurate on biopsy (it was downgraded in 50% of patients) followed by moderate grade prostate cancers (6). However, for a biopsy Gleason 5–6, King (4) had a sensitivity and PPV of 70% and 52%, respectively, while for Gleason 7 these were 38% and 53%, respectively, and for Gleason 8–10 were 47% and 58%, respectively. Similar to King's results, moderate grade tumor was the most difficult to find on biopsy, followed by high-grade cancer in our study.

In a study, 42- 69% discrepancy was found between NB GS and that of RP (7). Almost 95% of cases lacking exact correlation were within 1 GS group. Up to 45% of GS 8-10 at NB were downgraded to 7 or less at RP (attributed to incorrect interpretation or incomplete RP block sampling). NB GS of 2-4 was usually higher in RP (7, 17). In our study, no change happened in score 9; both cases with score 8 were changed. Twenty five percent of score 7, 33.3% of scores 5 and 6, and 33.3% of scores 2-4 were changed. The results of correlation between core biopsy and RP grades on over 1000 patients at the Johns Hopkins Hospital in 2006 were as follows: 81% for score 5 to 6, 68% for score 7 and 70% for score 8-10 (17). Our results included 66.6% for score 4, 66.6% for score 5-6, 75% for score 7, zero percent for score 8 and 100% for score 9.

68.8% of their 6625 patients had NB=RP GS, 25.0% had NB<RP GS, and 6.2% had NB>RP GS. In this study, the rate of discordance was 32–73% (9).

These differences may or may not be of clinical significance. When there is no change in the scoring group, an under grading of 2 or less GS might have little effect on the treatment. On the other hand, in some cases, a change of only 1 GS could alter treatment recommendations (7).

Different factors cause grading error, the most important one is sampling error, attributed to the heterogeneous nature of prostatic adenocarcinoma (3, 9). Grading error was twice as frequent in heterogenic cancers compared to non-heterogenic tumors (18). Borderline cases, in which a NB may be graded in two different, but not necessarily incorrect, ways, and biopsies taken from transition zone tumors are other causes of grading error (9). Failure to recognize an infiltrating growth pattern or small areas of gland fusion, and attempt to grade very tiny areas of carcinoma, so-called minimal or limited adenocarcinoma are other causes of undergrading (13).

Another factor that can also influence overall correlation of NB with RP results is the pathologists' experience. On reexamination, exact duplication of histologic scores occurred approximately 50% of the time and were within 1 score point approximately 85% of the time (19). Ruijter et al. reported 17% of errors in 187 cases resulted from pathologists' misinterpretation (18).

The recent increased accuracy of NB GS could be due to use of thin-core (18-gauge) needles attached to transrectal ultrasound, for sextant or other forms of systematic sampling (9-13). Divrik et al. showed improvement in the agreement between the Gleason score from core biopsy and radical prostatectomy specimens with an increase in the number of core biopsies obtained (56% vs. 41%) (20). To minimize grading error, taking at least 6 cores each 1.5 cm in length is recommended (4, 18). Other factors contributed in increasing accuracy are modernization of the Gleason scoring system, as well as dissemination of the literature via multimedia specialty courses at meetings and online web sites/atlases

(7, 9, 18). The significance of perineural invasion (PNI) and lymphovascular (LVI) invasion in prostatic NB in predicting prostate cancer recurrence is debated. Presence of PNI suggests correlation with risk of extraprostatic extension in most studies, which is a significant prognostic factor in predicting cancer recurrence.

However, the value of LVI in NB independent of Gleason grade or other pathological variables remains uncertain as a significant factor in predicting recurrence or survival (21). We found a positive relation between tumor score and vascular invasion (*P*-value= 0.00), perineural invasion (*P*-value= 0.032) and capsular invasion and/or extra prostatic extension (*P*-value= 0.009).

Conclusion

Overall, the reliability of Gleason grading of needle biopsies in predicting final pathology was good. However, pathologists and urologists must consider the phenomenon of undergrading and overgrading in reporting prostate specimens and managing patients.

Acknowledgments

We thank Abolfazl Ghoodjani for statistical analysis of the study.

Conflict of interest

The authors declare that there is no conflict of interests.

References

1. Ro JY, Kim K-R, Shen SS, Amin MB, Ayala AG. Tumors and tumor-like conditions of the male genital tract. In: Fletcher CDM, editor. Diagnostic Histopathology of Tumors. 4th ed. Philadelphia: Saunders; 2013.

2. Gleason DF. Classification of prostatic carcinoma. Cancer Chemother Rep 1966; 50(3):125-8.

3. Amin MB, Grignon DJ, Humphery PA, Srigley JR. Gleason grading of prostate cancer. Philadelphia: Lippincott Williams & Wilkins; 2004.

4. King CR. Patterns of prostate cancer biopsy grading: Trends and clinical implications. IJC 2000; 90(6):305-11.

5. Humphrey PA. Prostate pathology. Chicago: ASCP Press; 2003.

6. Cohen MS, Hanley RS, Kurteva T, Ruthazer R, Silverman ML, Sorcini A, et al. Comparing the Gleason prostate biopsy and Gleason prostatectomy grading system: The Lahey Clinic Medical Center Experience and an International Meta-Analysis. Eur Urol 2008; 54(2):371-81.

7. Fine SW, Epstein JI. A Contemporary Study Correlating Prostate Needle Biopsy and Radical Prostatectomy Gleason Score. J Urol 2008; 179(4):1335-9.

8. Memis A, Ugurlu O, Ozden C, Oztekin CV, Aktas BK, Akdemir AO. The correlation among the percentage of positive biopsy cores from the dominant side of prostate, adverse pathology, and biochemical failure after radical prostatectomy. Kaohsiung J Med Sci 2011; 27(8):307-13.

9. Müntener M, Epstein JI, Hernandez DJ, Gonzalgo ML, Mangold L, Humphreys E, et al. Prognostic significance of Gleason score discrepancies between needle biopsy and radical prostatectomy. Eur Urol 2008; 53(4):767-76.

10. Bostwik DG. Gleason grading of prostatic needle biopsies. Correlation with grade in 316 matched prostatectomies. Am J Surg Pathol 1994; 18(8):796-803.

11. Steinberg DM, Sauvageot J, Piantadosi S, Epstein JI. Correlation of prostate needle biopsy and radical prostatectomy Gleason grade in academic and community settings. Am J Surg Pathol 1997; 21(5):566-76.

12. Cookson MS, Fleshner NE, Soloway SM, Fair RF. Correlation between Gleason score of needle biopsy and radical prostatectomy specimens: Accuracy and clinical implications. J Urol 1997; 157(2):559-62.

13. Rubin MA, Dunn R, Kambham N, Misick CP, O'Toole KM. Should a Gleason score be assigned to a

minute focus of carcinoma on prostate biopsy? Am J Surg Pathol 2000; 24(12):1634-40.

14. Lange PH, Narayan P. Understaging and undergrading of prostate cancer. Urology 1983; 21(2):113-8.

15. Garnett JE, Oyasu R, Grayhack JT. The accuracy of diagnostic biopsy specimen in predicting tumor grades by Gleason's classification of radical prostatectomy specimens. J Urol 1984; 131(4):690-3.

16. Mills SE, Fowler JE. Gleason histologic grading of prostatic carcinoma. Correlation between biopsy and prostatectomy specimens. Cancer 1986; 57(2):346-9.

17. Epstein JI, Netto GJ. Biopsy interpretation of the prostate. Philadelphia: Lippincott Williams & Wilkins; 2008.

18. Ruijter E, Van Leenders G, Miller G, Debruyne F, Van de Kaa C. Errors in histological grading by prostatic needle biopsy specimens: frequency and predisposing factors. J Pathol 2000; 192(2):229-33.

19. Gleason DF. Histologic grading of prostate cancer: a perspective. Hum Pathol 1992; 23(3):273–79.

20. Divrik RT, Eroğlu, A, Şahin A, ZorluF, Özen H. Increasing the number of biopsies increases the concordance of Gleason scores of needle biopsies and prostatectomy specimens. Eur Urol 2007; 25(5):376-82.

21. Freeman A. Perineural and lymphovascular invasion on prostatic biopsy: Pathological assessment and significance. Surg Oncol 2009; 18(3):200-2.

How to cite this article:

Khoddami M, Khademi y, Kazemi Aghdam M, Soltanghoraee H. Correlation between Gleason scores in needle biopsy and corresponding radical prostatectomy specimens. A twelve-year review. Iran J Pathol. 2016;11(2):120-6.