Interobserver variations in reporting of prostatic adenocarcinoma using core biopsy specimens: a retrospective study from a tertiary referral hospital in Saudi Arabia

A.C. Al-Rikabi[1] *and H. Alkhalidi*[1]

تقييم الاختلافات بين الفاحصين في الإبلاغ عن السرطانة الغدية في البروستات باستخدام عينات خزعات لبية: دراسة استعادية من مستشفى إحالة جامعي في المملكة العربية السعودية

عمّار الركابي، هشام الخالدي

الخلاصة: يتم في السنوات الأخيرة إرسال أعداد أكبر من الخزعات اللبية للبروستات من أجل التقييم الهيستوباثولوجي، مع ما يصاحب ذلك من زيادة في عبء العمل بالنسبة لاختصاصي الباثولوجيا. وقد هدفت هذه الدراسة الاستعادية إلى تقييم التوافق والاختلاف بين الفاحصين من اختصاصيي الهيستوباثولوجيا في الإبلاغ عن السرطانة الغدية في البروستات باستخدام مواد تم الحصول عليها من عينات خزعات لبية من البروستات. فقد تم جلب ما مجموعه 810 عينات لخزعات لبية بالإبرة من البروستات تم الحصول عليها من 100 مريض يشتبه بإصابتهم بسرطانة غدية في البروستات، وذلك من المواد الأرشيفية في مستشفى جامعة الملك خالد بالرياض، وتم تصنيفها – بشكل مستقل – من قِبَل ثلاثة من اختصاصيي الهيستوباثولوجيا ذوي الخبرة، والذين عُمِّي عليهم التشخيص الأصلي. فكان هناك اتفاق كبير بين الفاحصين من اختصاصيي الباثولوجيا، ترافَقَ مع أحراز كابا غير مرجَّحة تتراوح ما بين 0.69 و0.85. وخلص الباحثان إلى تشجيع اختصاصيي الباثولوجيا في مستشفيات أخرى على أن يراجعوا – بشكل دوري – تَوَافُق التشخيصات في محاولة لتحسين ممارساتهم في مجال باثولوجيا غدة البروستات.

ABSTRACT In recent years, greater numbers of prostate biopsy cores are being submitted for histopathological assessment, with a concomitant increase in workload for the pathologist. This retrospective study aimed to assess the concordance and interobserver variation between histopathologists in reporting prostatic adenocarcinoma using material obtained from prostatic core biopsy specimens. A total of 810 prostatic needle core biopsy specimens obtained from 100 patients with suspected prostatic adenocarcinoma were retrieved from the archival material at King Khalid University Hospital, Riyadh, and classified independently by 3 experienced histopathologists who were blinded to the original diagnosis. There was considerable interobserver agreement between the pathologists, with unweighted kappa scores ranging from 0.69–0.85. We would encourage other hospital pathologists to review periodically the uniformity of diagnoses in an attempt to improve their practices in prostate gland pathology.

Évaluation des variations inter-observateurs dans la notification des adénocarcinomes prostatiques à partir d'échantillons de microbiopsie : étude rétrospective dans un hôpital de soins tertiaires en Arabie saoudite

RÉSUMÉ Ces dernières années, le nombre d'échantillons de microbiopsie de la prostate soumis à une analyse histopathologique a augmenté, ainsi que la charge de travail concomitante du pathologiste. La présente étude rétrospective visait à évaluer la concordance et la variation inter-observateurs parmi les histopathologistes dans la notification de l'adénocarcinome prostatique à partir d'échantillons de microbiopsie prostatiques. Au total, 810 échantillons prostatiques de microbiopsie au trocart de 100 patients chez qui un adénocarcinome prostatique était suspecté, ont été extraits des archives de l'hôpital universitaire King Khalid à Riyad, puis classifiés indépendamment par trois histopathologistes expérimentés qui ignoraient le diagnostic initial. La concordance entre les observateurs pathologistes était élevée, avec des scores kappa non pondérés compris entre 0,69 et 0,85. Nous encourageons d'autres pathologistes hospitaliers à examiner périodiquement l'uniformité des diagnostics afin d'améliorer leurs pratiques dans la pathologie de la prostate.

[1]*Department of Pathology, College of Medicine, King Saud University, Riyadh, Saudi Arabia (Correspondence to A.C. Al-Rikabi: ammar_rikabi12@yahoo.com).*

Received: 20/06/12; accepted: 08/07/12

Introduction

Apart from skin cancer, carcinoma of the prostate is the most common internal malignancy among men in Western countries. It is responsible for 10% of cancer deaths (*1*) and is on the increase in most countries (*2*). The rate of prostate cancer in Saudi Arabia ranked sixth among male patients with a crude annual incidence of 5.7 per 100 000 (*3*).

The diagnosis of prostate cancer requires the estimation of prostate-specific antigen and multiple cores obtained by thin-bore needle biopsies. Increasing the number of needle cores analysed to between 6 to 12 has been shown to improve prostate cancer detection by 29% (*4,5*). For this reason, a greater number of prostate biopsies are obtained nowadays and more biopsy cores are being submitted than ever before and this has created a huge interpretive burden for the diagnostic histopathologist (*6*), a burden that is exacerbated by the difficulties of prostate biopsy interpretation (*6,7*).

The subject of interobserver variation in cytological and histological diagnoses of cancer and its epidemiological implications has become increasingly relevant in the last 20 years. Various studies have shown highly reproducible diagnoses of uterine cervical neoplasia (*8*), while there was considerable interobserver variation in the reporting of anal intraepithelial neoplasia (*9*) and in certain types of breast carcinomas (*10*). With this in mind, this study in Saudi Arabia aimed to assess the diagnostic reproducibility and interobserver variation in histopathological reporting of prostatic adenocarcinoma among 3 histopathology consultants.

Methods

Sample

In this retrospective study a total of 810 prostatic needle biopsies obtained from 100 patients from 2005 until the end of 2011 were retrieved from the archives of the histopathology unit at King Khalid University Hospital, Riyadh. The number of needle cores obtained from each patient varied from 6 to 12 cores. This was due to an evidence-based change in the policy (in the last 10 years) regarding the number of needle cores that should be obtained per patient.

Data collection

All biopsies were processed and stained using haematoxylin and eosin stain in our laboratory. We routinely obtain 6 cuts (2 adjacent sections from 3 separate levels) from the paraffin block for routine staining.

Three experienced general histopathologists were asked to examine each of the 810 stained needle biopsies without their knowledge of the previous clinical, radiological or histopathological findings of the 100 patients from whom those biopsies were obtained. The results obtained from each pathologist (A, B and C) were independently documented as being either negative for malignancy (normal), positive for malignancy (abnormal) or inconclusive, i.e. in need of further immunohistochemical stains or repeat biopsies.

Analysis

The results were tabulated and analysed using the multiple-reader Cohen kappa statistical analysis method (*7,11*) using *SPSS*, version 18. The aim was to assess the precision pertaining to agreement between observers (interobserver agreement). The Gleason score (grade of malignant cases) was not, however, recorded as this parameter was outside the scope of this retrospective study.

Results

The results obtained by the 3 participating pathologists are summarized in Table 1. The kappa score for interobserver agreement between pathologists A and B was 0.71 ($P < 0.001$), between pathologists A and C was 0.69 ($P < 0.001$) and between pathologist B and C was 0.85 ($P < 0.001$).

Our results also showed that almost all cases of diagnostic uncertainty and interobserver differences in interpretation were due to the presence of small atypical or atrophic acini areas consisting of 5 or fewer acini. The percentages of such cases varied between 2% and 6% and were mainly due to the presence of small atypical acinar proliferation.

Discussion

It may sometimes be challenging for the pathologist to deliver a definite diagnosis of adenocarcinoma in prostate biopsies, particularly if the size of the lesion is too small to judge the presence of an infiltrative pattern. This issue has become more pertinent in recent years due to clinical stage reduction of

Table 1 Frequency distribution of prostatic specimen observations among the 3 pathologists (*n* = 810)

Observation	Pathologist A		Pathologist B		Pathologist C	
	No.	%	No.	%	No.	%
Normal	558	68.9	575	71.0	562	69.4
Abnormal	200	24.7	214	26.4	231	28.5
Inconclusive	52	6.4	21	2.6	17	2.1

prostate cancer, which occurred as a consequence of widespread prostate-specific antigen testing and increased numbers of biopsies leading to "early" diagnosis of smaller cancerous foci (*6*). Also repeat biopsies in the context of active surveillance treatment might lead to an increased frequency of small foci of adenocarcinoma, high-grade prostatic intraepithelial neoplasia and lesions reported as suspicious for malignancy or atypical small acinar proliferation.

In this study of prostate core biopsy specimens the degree of concordance among the 3 histopathologists was estimated using the kappa coefficient. This is the most commonly reported measure in the medical literature, and can provide more information than a simple calculation of the raw proportion of agreement. The method is excellent for comparing results obtained by individuals but is slightly affected by prevalence (*11*). The study showed a good degree of concordance in the interpretation of prostatic needle biopsies among the 3 participating histopathologists, with interobserver agreements which varied between kappa 0.69 and 0.85. There was substantial interobserver agreement between pathologists A and B (0.71) and A and C (0.69) (kappa values 0.61–0.80 are generally interpreted as substantial agreement) (*11*), while the interobserver agreement between pathologist B and C (0.85) fell within the kappa range of 0.81–1, which is generally interpreted as almost perfect agreement (*11*). Inconclusive interpretations were mostly due to the presence of small atypical or atrophic acini. These findings are keeping with those reported in the literature (*7,12–15*).

Despite the interpretative difficulties and the burden of an increasing workload, experienced surgical pathologists have a high level of accuracy in prostatic needle biopsy interpretation and Gleason grading. Interobserver reproducibility of Gleason grading among urologic pathologists has been shown to acceptable (*5,16–19*). The greater differences of interpretation result from low-grade cancers (*6*), cancers with small cribriform pattern (*20*) and cancers whose histology is on the border between Gleason patterns (*16,17*). The false-negative rate (missed prostate cancer) was 0.6–1% and the false-positive rate (overdiagnosis of prostate cancer) was 0.3 (*6,18*). These numbers indicate a small but significant error level that could be avoided by secondary pathology review (*18,19*). The findings obtained from the biopsies in the current study were not compared with those seen in the excision specimens (prostatectomies) as we aimed to assess the reproducibility of the results and not measure the accuracy of the initial diagnoses.

Conclusions

This study show good interobserver agreement in the interpretation of prostatic and needle biopsies among the participating histopathologists (kappa ranges 0.69–0.85). Inconclusive interpretations were mostly due to the presence of small atypical or atrophic acini. The establishment of an intradepartmental system of consultation with joint reporting and signing out of prostatic carcinoma by at least 2 experienced histopathologists will help maintain a high degree of diagnostic concordance. Based on the results presented, we would encourage other hospital pathologists, in collaboration with their urologists, to review periodically the uniformity of their diagnoses in an attempt to improve their prostate gland pathology practices.

Acknowledgements

The authors would like to express their gratitude to Dr Abdulmalik Al-Sheikh who kindly agreed to participate in this study. The secretarial help of Ms Roxanne Alamares during the typing of this manuscript is also greatly appreciated.

Funding: None.

Competing interests: None declared.

References

1. Nelson WG, De Marzo AM, Isaacs WB. Prostate cancer. N Engl J Med. 2003 Jul 24;349(4):366–81. PMID:12878745
2. Zaridze DG, Boyle P, Smans M. International trends in prostatic cancer. Int J Cancer. 1984 Feb 15;33(2):223–30. PMID:6693200
3. Cancer incidence and survival report. Riyadh: National Cancer Registry, Ministry of Health; 2007:56–7.
4. Fink KG, Hutarew G, Pytel A, Schmeller NT. Prostate biopsy outcome using 29 mm cutting length. Urol Int. 2005;75(3):209–12. PMID:16215306
5. Zeng J, Bauer J, Zhang W, Sesterhenn I, Connelly R, Lynch J, et al. Prostate biopsy protocols: 3D visualization-based evaluation and clinical correlation. Comput Aided Surg. 2001;6(1):14–21. PMID:11335955
6. Bostwick DG, Meiers I. Prostate biopsy and optimization of cancer yield. Eur Urol. 2006 Mar;49(3):415–7. PMID:16442209
7. Van der Kwast TH, Evans A, Lockwood G, Tkachuk D, Bostwick DG, Epstein JI, et al. Variability in diagnostic opinion among pathologists for single small atypical foci in prostate biopsies. Am J Surg Pathol. 2010 Feb;34(2):169–77. PMID:20061936
8. McCluggage WG, Walsh MY, Thornton CM, Hamilton PW, Date A, Caughley LM, et al. Inter- and intra-observer variation in the histopathological reporting of cervical squamous intraepithelial lesions using a modified Bethesda grading system. Br J Obstet Gynaecol. 1998 Feb;105(2):206–10. PMID:9501788
9. Carter PS, Sheffield JP, Shepherd N, Melcher DH, Jenkins D, Ewings P, et al. Interobserver variation in the reporting of the histopathological grading of anal intraepithelial neoplasia. J Clin Pathol. 1994 Nov;47(11):1032–4. PMID:7829679
10. Pedersen L, Holck S, Schiødt T, Zedeler K, Mouridsen HT. Inter- and intraobserver variability in the histopathological di-

agnosis of medullary carcinoma of the breast, and its prognostic implications. *Breast Cancer Res Treat*. 1989 Oct;14(1):91–9. PMID:2605345

11. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med*. 2005 May;37(5):360–3. PMID:15883903
12. Epstein JI, Herawi M. Prostate needle biopsies containing prostatic intraepithelial neoplasia or atypical foci suspicious for carcinoma: implications for patient care. *J Urol*. 2006 Mar;175(3 Pt 1):820–34. PMID:16469560
13. Herawi M, Parwani AV, Irie J, Epstein JI. Small glandular proliferations on needle biopsies: most common benign mimickers of prostatic adenocarcinoma sent in for expert second opinion. *Am J Surg Pathol*. 2005 Jul;29(7):874–80. PMID:15958851
14. Iczkowski KA, MacLennan GT, Bostwick DG. Atypical small acinar proliferation suspicious for malignancy in prostate needle biopsies: clinical significance in 33 cases. *Am J Surg Pathol*. 1997 Dec;21(12):1489–95. PMID:9414193
15. Novis DA, Zarbo RJ, Valenstein PA. Diagnostic uncertainty expressed in prostate needle biopsies. A College of American Pathologists Q-probes Study of 15,753 prostate needle biopsies in 332 institutions. *Arch Pathol Lab Med*. 1999 Aug;123(8):687–92. PMID:10420224
16. Melia J, Moseley R, Ball RY, Griffiths DF, Grigor K, Harnden P, et al. A UK-based investigation of inter- and intra-observer reproducibility of Gleason grading of prostatic biopsies. *Histopathology*. 2006 May;48(6):644–54. PMID:16681679
17. Coard KC, Freeman VL. Gleason grading of prostate cancer: level of concordance between pathologists at the University Hospital of the West Indies. *Am J Clin Pathol*. 2004 Sep;122(3):373–6. PMID:15362366
18. Barqawi AB, Turcanu R, Gamito EJ, Lucia SM, O'Donnell CI, Crawford ED, et al. The value of second-opinion pathology diagnoses on prostate biopsies from patients referred for management of prostate cancer. *Int J Clin Exp Pathol*. 2011 Jun 20;4(5):468–75. PMID:21738818
19. Oxley J. Reviewing negative prostatic core biopsies for the multidisciplinary team meeting. *Histopathology*. 2005 Dec;47(6):643–4. PMID:16324205
20. Latour M, Amin MB, Billis A, Egevad L, Grignon DJ, Humphrey PA, et al. Grading of invasive cribriform carcinoma on prostate needle biopsy: an interobserver study among experts in genitourinary pathology. *Am J Surg Pathol*. 2008 Oct;32(10):1532–9. PMID:18724248