ORIGINAL ARTICLE

Interobserver reproducibility of a grading system for chromophobe renal cell carcinoma

A. Pérez-Pedrosa, J.A. Ortiz-Rey, Y. Lorenzo-Mahía, B. Iglesias-Rodríguez, A. Peteiro-Cancelo, J. González-Carreró

Servicio de Anatomía Patológica, Hospital Xeral-Cíes, Complejo Hospitalario Universitario de Vigo, Vigo, Pontevedra, Spain
Servicio de Anatomía Patológica, Hospital do Meixoeiro, Complejo Hospitalario Universitario de Vigo, Vigo, Pontevedra, Spain
Servicio de Anatomía Patológica, Hospital POVISA, Vigo, Pontevedra, Spain

Received 8 March 2012; accepted 5 April 2012
Available online 5 October 2013

KEYWORDS
Carcinoma; Renal; Chromophobe; Grading; Fuhrman

Abstract
Objectives: To evaluate interobserver reproducibility of a grading system proposed by Paner et al. for chromophobe renal cell carcinoma.

Material and methods: After selecting 23 cases of chromophobe renal cell carcinoma from the Xeral-Cíes Hospital, Meixoeiro Hospital and POVISA Hospital from the last 15 years, an informative meeting on the Paner et al. grading system criteria was held. After, the participating pathologists applied the system to each case, evaluating one slide selected. Kappa index for interobserver reproducibility was calculated, and it was classified according to the Landis and Koch scale.

Results: The grading distribution was similar for most of the 6 participating observers, with grade 1 predominance. The remaining 2 observers considered a relatively higher proportion of grade 2. Kappa index values ranged from 0.136 to 0.674, with a discrete-moderate reproducibility index predominance (0.21–0.60). Highest Kappa value (0.674) was obtained between the most novel and the most expert interobservers. The lowest Kappa value was obtained among the most veteran pathologists (0.136).

Conclusions: Interobserver reproducibility for chromophobe renal cell carcinoma is discrete-moderate in our institutions when the novel grade proposed by Paner et al. is used. Labeling of grades 1 and 2 is not homogeneous among 6 participating observers. While awaiting a grading consensus on a new classification by the scientific societies, we consider that the routine use of a grading system for chromophobe renal cell carcinoma should not be used.

© 2012 AEU. Published by Elsevier España, S.L. All rights reserved.


* Corresponding author.
E-mail address: appedrosa77@gmail.com (A. Pérez-Pedrosa).

2173-5786/$ - see front matter © 2012 AEU. Published by Elsevier España, S.L. All rights reserved.
Interobserver reproducibility of a grading system for chromophobe renal cell carcinoma

PALABRAS CLAVE
Carcinoma; Renal; Cromófobo; Gradación; Fuhrman

Reproducibilidad interobservador de un sistema de grado para el carcinoma de células renales tipo cromófobo

Resumen
Objetivos: Valorar la reproducibilidad interobservador y evaluar el sistema de gradación propuesto por Paner et al. para el carcinoma de células renales cromófobo.

Material y métodos: Tras seleccionar 23 casos de carcinoma renal de tipo cromófobo de los hospitales Xeral-Cies, Meixoeiro y POVISA de Vigo de los últimos 15 años se ha realizado una sesión informativa de los criterios del sistema de gradación de Paner et al. Posteriormente los patólogos observadores han aplicado dicho sistema a cada caso, valorando una lamíllula seleccionada. Se ha calculado el índice Kappa de reproducibilidad interobservador, ponderado según la escala de Landis y Koch.

Resultados: La distribución de grados en la mayoría de los 6 observadores participantes es similar, con predominio del grado 1 en 4 de los mismos. Los 2 observadores restantes consideraron una mayoría relativa de casos como grado 2. Los valores de Kappa oscilan entre 0,136 y 0,674, observándose un predominio de valores indicadores de reproducibilidad discreta-moderada (0,21-0,60). El mayor valor de Kappa (0,674) se ha dado entre un observador novel y el patólogo más experto. Entre los 2 observadores más veteranos se ha obtenido el índice más bajo (0,136).

Conclusiones: La reproducibilidad interobservador en nuestros centros para el grado propuesto por Paner et al. es discreta-moderada. La asignación de los grados 1 y 2 no es homogénea entre los 6 observadores participantes. En espera de la existencia de una gradación consensuada por las sociedades científicas, creemos prudente no utilizar ningún sistema de gradación en los carcinomas de células renales de tipo cromófobo.

© 2012 AEU. Publicado por Elsevier España, S.L. Todos los derechos reservados.

Introduction

The chromophobe renal carcinoma is a malignancy that, in most cases, has a good prognosis. The classification of the degree of differentiation has been made following the criteria established for other neoplasms of kidney cells, the so-called Fuhrman grading. This classification establishes the different degrees based primarily on the nuclear features of the neoplastic cells (nuclear pleomorphism, chromatin features, visibility, and prominence of nucleoli). The Fuhrman grading has the advantage of its convenience and familiarity to the general pathologist.

However, the scientific literature, for years, has questioned the value thereof for some tumor subtypes.1,2 Recently, some authors (Paner et al.) have pointed out its poor reproducibility and limited predictive value on the prognosis in chromophobe renal cell carcinomas, while they have suggested an alternative classification.

The 3-degree system proposed by Paner et al. assumes the nuclear atypia inherent to chromophobe carcinoma and establishes criteria based on the architecture (such as the nuclear overlap and its layout pattern) and identifying criteria of frank anaplasia (sarcomatoid change, polylobulation, etc.), as well as the chromatin features involving frank atypia.

The new classification proposed would not significantly affect the tumors labeled as Fuhrman 1–2 or Fuhrman 4, as these, no doubt, would be classified as grades 1 and 3, respectively, in the new classification, with similar prognostic implications.

However, the series studied by Paner et al. provide mixed results in the carcinomas labeled as Fuhrman 3. Thus, a subgroup of them would meet the criteria of anaplasia and would be classified as new grade 2, while others would correspond to a new grade 1.

In our study, we retrospectively analyzed a series of cases applying this novel system, and we intend to analyze their applicability from our experience.

Material and methods

We reviewed 23 cases of the last 15 years with a diagnosis of chromophobe carcinoma of the Complejo Hospitalario Universitario de Vigo (CHUVI) and the POVISA Hospital, which have been graduated with the new classification. We intend to assess the interobserver reproducibility of this system in the team of pathologists of our departments.

After the selection of cases diagnosed with undoubted diagnosis of chromophobe carcinoma, a slide was given to each of the 6 participating pathologists, 2 of them residents and the remaining 4 experienced associate physicians. They all previously attended a briefing in which the classification criteria of Paner et al. had been exposed before, accompanied by demonstrative images of each grade. Subsequently, an assessment of the degree was required according to these criteria. We performed a statistical analysis of results, based on descriptive, epidemiological, and analytical data, the latter consisting in the calculation of the Kappa index for interobserver reproducibility. The results
Table 1  Scale of Landis and Koch for interpreting the Kappa value (grade of interobserver reproducibility).

<table>
<thead>
<tr>
<th>Kappa value</th>
<th>Grade of agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;0.00</td>
<td>No agreement</td>
</tr>
<tr>
<td>0.00-0.20</td>
<td>Insignificant</td>
</tr>
<tr>
<td>0.21-0.40</td>
<td>Discrete</td>
</tr>
<tr>
<td>0.41-0.60</td>
<td>Moderate</td>
</tr>
<tr>
<td>0.61-0.80</td>
<td>Substantial</td>
</tr>
<tr>
<td>&gt;0.80</td>
<td>Nearly perfect</td>
</tr>
</tbody>
</table>

Figure 1  Fuhrman grades assigned in the pathology report.

Figure 2  Pathological stage assigned in the pathology report.

The majority of cases studied had been classified with Fuhrman grades 2 or 3 (39 and 30% respectively), grades 1 (9%) and 4 (4%) being minority. Most cases at diagnosis had a low pathologic stage (48% in stage I) (Figs. 1 and 2).

Assigning grades according to Paner et al. follow a similar distribution in the 6 participating pathologists, although the main differences are observed in grades 1 and 2. In this regard, the 'inversion' in the pattern of assignments of these grades in 2 of the pathologists stands out with regard to the other 4 participants (Fig. 3).

Regarding the Kappa indexes of interobserver concordance, 'insignificant' concordance indexes were obtained in 4 of the possible interobserver combinations. The index is 'discreet' interobserver concordance also in 4 combinations. There have been 5 interobserver combinations in 'moderate' concordance ranges. Only one of the possible combinations has been of 'substantial' concordance, and none of them has been 'no agreement' or 'almost perfect agreement' (Table 2). Therefore, we obtained 60% interobserver combinations in the discrete-moderate range.

The best grade of agreement, with a 0.674 Kappa index ('substantial' agreement according to the Landis and Koch scale) was obtained in the combination of one of the participating residents and the oldest associate pathologist. By contrast, the minimum degree of agreement obtained (Kappa = 0.136 /'insignificant' agreement) is the combination between the 2 most experienced pathologists.

Figure 3  New grade of Paner et al. Distribution of grades by observer pathologists.
Table 2  Kappa values. Interobserver combinations.

<table>
<thead>
<tr>
<th>Pathologist</th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>P5</th>
<th>P6</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>0.263</td>
<td>0.489</td>
<td>0.2</td>
<td>0.524</td>
<td>0.531</td>
<td></td>
</tr>
<tr>
<td>P2</td>
<td>0.305</td>
<td>0.144</td>
<td>0.481</td>
<td>0.471</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P3</td>
<td>0.532</td>
<td>0.332</td>
<td>0.371</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P4</td>
<td>0.136</td>
<td>0.181</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P5</td>
<td>0.674</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Discussion

This paper reflects the interest aroused by the recently proposed classification for chromophobe renal cell carcinomas, preceded by extensive discussions in the literature.\(^6\) Given the poor existing prognostic correlation with the classic grading systems of renal cell carcinomas, we detected a discrete-moderate interobserver reproducibility index in the new gradation following the criteria of Paner et al. The main differences in the allocation of grade affect grades 1 and 2. Tumors with anaplasia or sarcomatoid change (grades 3) appear to be identified with a high degree of agreement. The experience does not seem to be a determining factor in the index of interobserver reproducibility, since the highest level of agreement was obtained between a novel observer and the most experienced pathologist. In contrast, the lowest reproducibility index has occurred between the 2 most experienced. Recently, the proposal of Paner et al. has prompted comments in the form of ‘Letter to the Editor’ (Bradshaw et al.).\(^7\) The publication noted the possibility of grouping the tumors with a dichotomous classification (high/low grade), with the argument of increasing reproducibility and stratifying the different prognoses of the neoplasms according to the 2 possible categories. They point out the greater familiarity and simplicity for the pathologist of the 2-grade systems. They also highlight the difficulty that, in their view, the distinction between Paner et al. grades 1 and 2 represents. This has been reflected in our study, since, as previously noted, the greatest discrepancies in our case analysis were rooted precisely in the new grades 1 and 2. Regarding these objections, Paner and Amin in the corresponding ‘Response letter’ have argued that, if tumors were stratified with a dichotomous system (low grade: Paner et al. grade 1, and high grade: grades 2 and 3), a ‘high grade’ would be created which, according to the data of their series, would place 2 groups of tumors with very different prognoses in the same category (36 and 70% with adverse prognosis in grades 2 and 3, respectively). We have identified, the same as Bradshaw et al., difficulties in distinguishing between grades 1 and 2 for certain tumors with ‘neighboring’ morphological features. Moreover, the observed reproducibility is not optimal. In conclusion, we consider reasonable both doubts about this and the need for a stratification according to the prognosis. In this regard, we appreciate the new classification suggested by Paner et al.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgements

We had the collaboration of M Fontanillo, Biostatistics Unit of the Fundación del Complejo Hospitalario Universitario de Vigo (FICHUVI) for the statistical treatment of the work. Dr. A Mateos, from the Pathological Anatomy Department of the Hospital de Vigo Xeral, has also collaborated in the preparation of this article.

We very much appreciate your support and make you participants of this study.

References