Research article

CT-based radiomic model predicts high grade of clear cell renal cell carcinoma

Jiule Ding, Zhaoyu Xing, Zhenxing Jiang, Jie Chen, Liang Pan, Jianguo Qiu, Wei Xing

ARTICLE INFO

Keywords:
Tomography
Image texture
Radiomics
Image feature
CT

ABSTRACT

Purpose: To compare the predictive models that can incorporate a set of CT image features for preoperatively differentiating the high grade (Fuhrman III–IV) from low grade (Fuhrman I–II) clear cell renal cell carcinoma (ccRCC).

Material and methods: One hundred and fourteen patients with ccRCC treated with a partial or radical nephrectomy were enrolled in the training cohort. The six non-texture features, including Pseudocapsule, Round mass, maximal tumor diameter (Diametermax), intratumoral artery (Arterytumor), enhancement value of the tumor (TEV) and relative TEV (rTEV), were assessed for each tumor. The texture features were extracted from the CT images of the section with the largest area of renal mass at both corticomedullary and nephrographic phases. The least absolute shrinkage and selection operator (LASSO) was used to screen the most valuable texture features to calculate a texture score (Texture-score) for each patient. A logistic regression model was used in the training cohort to discriminate the high from low grade ccRCC at nephrectomy. The predictors would include all non-texture features in Model 1, all non-texture features and Texture-score in Model 2, and Texture-score in Model 3. The performance of the predictive models were tested and compared in an independent validation cohort composed of 92 cases with ccRCC.

Results: Inter-rater agreement was good for each non-texture feature and Texture-score (the concordance correlation coefficient or Kappa coefficient > 0.70). The Texture-score was calculated via a linear combination of the 4 selected texture features. The three models shown good discrimination of the high from low grade ccRCC in the training cohort and the area under receiver operating characteristic curve (AUC) was 0.826 in Mode 1, 0.878 in Model 2 and 0.843 in Model 3, and a significant different AUC was found between Model 1 and Model 2.

Application of the predictive models in the validation cohort still gave a discrimination (AUC > 0.670), and the Texture-score based models with or without the non-texture features (Model 2 and 3) showed a better discrimination of the high from low grade ccRCC (P < 0.05).

Conclusion: This study presented the Texture-score based models can facilitate the preoperative discrimination of the high from low grade ccRCC.

1. Introduction

Renal cell carcinoma (RCC) is the seventh most common cancer. It leads to the cancer-related death of 140,000 patients globally every year [1]. The clear cell RCC (ccRCC) accounts for 75%–87% cases with RCC [2]. Partial nephrectomy is established as the preferred treatment method for ccRCC; however radical therapy has been proposed for the patients with high-risk tumors [3]. This proposal has spurred interest in the development of radiologic and biopsy-based diagnostic techniques that can identify the high risk ccRCC [3]. Currently Fuhrman grading, a grading system widely used for RCC, is one of the most powerful prognostic factors in the patients with RCC [4,5]. In the analysis of metastatic potential in the kidney carcinoma data, patients with Fuhrman III or IV ccRCC have a four-fold greater risk for metastasis than those with Fuhrman I or II ccRCC [5]. In the cohort of 4063 patients with RCC derived from 8 international centers, the 10-year cancer specific survival rate was 81.0%, 56.6%, 30.1% and 18.8% in the Fuhrman I–IV subgroups, respectively [2]. Other studies further confirmed that tumor grade was an independent predictor for patients with ccRCC [6]. Thus, estimating tumor grade before treatment substantially...
aids in the high-risk assessment.

Currently percutaneous renal biopsy is the preferred method for the pretreatment examination of the pathological characteristics within renal mass. Although the technique has an overall 97% accuracy for the diagnosis of malignant mass, the concordance rate of Fuhrman grade was less promising, from 31.3% to 97% \cite{3,7,8}. To assist in the individual estimation of Fuhrman grade, the predictive models were developed using clinical variables such as patient age, gender, symptom and tumor size; however, the predictive accuracy of these models is close to that of flipping a coin (50%) \cite{9,10}.

Radiomics, mineable data via high-throughput extraction of quantitative features followed by subsequent data analysis for decision support, has attracted increased attention \cite{11}. The radiomic models that combine a panel of biomarkers (such as texture features), rather than individual analyses, are currently the most promising approach to improve prediction of the high grade ccRCC. Therefore, the aim in this study is to develop and compare the radiomic models that incorporate CT image features for the individual prediction of the high grade ccRCC.

2. Material and methods

2.1. Patients

We retrospectively collected data on the patients who had been admitted for the resection of renal masses from 1 Jan 2011 to 31 Dec 2016, who received a surgical treatment and the histologic analysis of the resected whole tumors. Patient inclusion/exclusion details are presented in Fig. 1, and there are no metal items and motion artifacts on CT images. There was a renal mass per patient. The training cohort consisted of 114 cases scanned at 3 different CT scanners. And the validation cohort consisted of 92 cases scanned at another scanner (CT320).

2.2. CT image acquisition

All subjects had undergone a default abdominal CT scan using one of the four multi-detector row CT (MDCT) systems (Table 1) with the scanning and reconstruction parameters used in daily clinical practice.

All subjects had also undergone a triphasic CT scan, including i) a routine unenhanced CT scan, ii) a corticomedullary phase contrast-enhanced scan starting 30 s–35 s after the beginning of contrast material injection, and iii) a nephrographic phase contrast-enhanced scan performed after 60–70 s after intravenous injection of 60 ml–100 ml of iodinated contrast material (Ultravist 350, Bayer Schering Pharma, Berlin, Germany) at a rate of 2.5 ml/s–3.5 ml/s with a pump injector (Ulrich CT Plus 150, Ulrich Medical, Ulm, Germany).

CT images were retrieved from the picture archiving and communication system (Sesan, China).

2.3. Assessment of renal mass on CT images

For each renal lesion, the CT images of the section with the largest cross-sectional area of renal mass at both corticomedullary and nephrographic phases were chosen for further analysis. Both radiologists (Reader 1: J.Q. and Reader 2: J.C.) with more than 15 years of experience in abdominal CT interpretation, being blinded to all clinical, pathologic and imaging findings, conducted the image analysis. The assessment of renal mass on CT images conducted by Reader 1 was used for grading ccRCC. The assessment of CT images of renal mass scanned at CT16 conducted by Reader 2 was used for inter-rater agreement assessment of the serial image features.

Firstly, on the CT images of the selected section, both radiologists assessed whether each renal mass exhibited the following non-texture features: Pseudocapsule, Round mass, maximal tumor diameter of < 4 cm (\text{Diameter}_{\text{max}}), Intratumoral artery (\text{Artery}_{\text{tumor}}), for some, the image features of Pseudocapsule and \text{Diameter}_{\text{max}} were assessed at the nephrographic phase, and for others at the corticomedullary phase. The typical image features of Pseudocapsule, Round mass and \text{Artery}_{\text{tumor}} are presented in Fig. 2. Renal masses can be easily detected on the nephrographic phase, being present as a hypoattenuation mass compared with the enhanced renal parenchyma, and the \text{Diameter}_{\text{max}} of renal mass was measured and classified as ‘presence’ when it is < 4 cm.

Secondly, both radiologists measured and calculated another two non-texture image features of the enhancement value of the tumor (TEV) and relative TEV (rTEV). The selected CT images at the unenhanced and corticomedullary phases were used in the enhancement assessment of renal mass because the perfect performance of enhance features at the corticomedullary phase indicated in a previous study \cite{14}. The attenuation values of three separate ROIs within the renal lesions were measured. To avoid the necrosis in the ROI, each ROI (25 mm²–100 mm²) was placed on the most avidly enhancing parts of the tumor at corticomedullary phase but excluded blood vessels and calcification. The average attenuation value of the three ROIs was calculated to represent the tumor attenuation value in the unenhanced and corticomedullary phases, respectively. At the same time, the attenuation value of the renal cortex was also measured as a reference to indicate the iodine load. The TEV and the enhancement value of renal cortex (CEV) were calculated as the difference between the attenuation values in the corticomedullary phases and in the unenhanced phase, respectively, and were expressed as follows: \text{TEV} = \text{TAV}_{\text{corticomedullary phase}} - \text{TAV}_{\text{unenhanced phase}} and \text{CEV} = \text{CAV}_{\text{corticomedullary phase}} - \text{CAV}_{\text{unenhanced phase}}, where TAV and CAV represent the attenuation values of the tumor and of renal cortex respectively. The relative TEV (rTEV) is defined as the ratio of TEV and CEV.

Thirdly, on the CT images of the selected section of renal mass at the corticomedullary and nephrographic phases, texture analysis was conducted using the open-source software Imaging Biomarker Explorer (IBEX) \cite{15}. Regions of interest (ROIs) were initially delineated around the outline of the renal mass. Then, 184 texture features were extracted from each ROI using the categories of histogram, gray-level co-occurrence matrix (GLCM) and two-level run length matrix (GLRLM). Both GLCM and GLRLM were used to generate the second-order texture features.
features by comparing neighboring pixels for similarity and dissimilarity. The comparison was defined as a matrix $P[k, m]$ to indicate the relative frequency, with the intensity of two pixels ($i$ and $j$) at a distance of $\delta = (1, 4$ and $7)$ and in the direction $0^\circ$ for GLCM, and in the direction $90^\circ$ for GLRLM. The four of 184 texture features are as following: RunLengthNonuniformity, Contrast, GrayLevelNonuniformity and 0.025Quantile.

### 2.4. Pathological evaluation

Whole-tumor specimens were placed in formalin solution and carried to the pathology laboratory. Histopathological evaluation was performed with hematoxylin and eosin staining, and sometimes completed with immunohistochemistry when determined to be necessary by the pathologist. Tumor histological findings were classified according to the WHO 2004 system. Pathological analysis of the Fuhrman grading system was performed, as indicated in a previous study [16]. Fuhrman III and IV were grouped as high grade, and Fuhrman I and II as low grade, because the simplified grading schemes performed equally as well as the conventional Fuhrman grading system [17,18].

### 2.5. Statistical analysis

To compare the patient’s characteristics and the discontinuous variable between the high and low grade ccRCC, a $\chi^2$ or Fisher’s exact test was used. Both TEV and rTEV between the high and low grade ccRCC were compared using a two-independent sample $t$-test. The least absolute shrinkage and selection operator (LASSO) was used to select the most valuable features from the texture features, then the texture score (Texture-score) was calculated for each patient via a linear combination of the selected features weighted by their corresponding coefficients.

Inter-rater agreement was assessed for the continuous variable using the concordance correlation coefficient (CCC) with 95% confidence interval (95%CI), and for the binomial variable using Kappa test. And the CCC or Kappa coefficient $>0.7$ implies a good inter-rater agreement.

To develop the models for the individual prediction of the high grade ccRCC, a logistic regression with multiple variables or one variable was performed on the basis of all non-texture features (Pseudocapsule, Round mass, Diametermax, Arterytumor, TEV and rTEV) (Model 1), of all non-texture features and Texture-score (Model 2), and of Texture-score (Model 3); backward step-wise selection was applied in the logistic regression with multiple variables as indicated [11]. Receiver operating characteristic curve was performed to compare the capacity of the predictive models for the discrimination of the high from low grade ccRCC in the training and validation cohorts, and the area under receiver operating characteristic curve (AUC) with 95% confidence interval (95% CI) was calculated.

The $\chi^2$ or Fisher’s exact test was performed using Prism 5 for Windows (Version 5.01, Serial Number: GPWS-384305-RAG-5235), and the regression models were built and validated using R version 3.4.0 (Foundation for Statistical Computing, Tsinghua University, China). The threshold for significance was a two-sided $P$-value (0.05).

### 3. Results

#### 3.1. Study participants

Patient characteristics in the training and validation cohorts were provided in Table 2, and no significant difference in patient characteristics was found between the high and low grade ccRCC.

#### 3.2. Inter-rater agreements of the non-texture features

Seventy-nine cases scanned in CT16 were included in the statistical analysis for the assessment of the inter-rater agreement of the non-
texture features. The image features and corresponding measurement values obtained by the two readers were listed for each patient in Table S1. The inter-rater agreement (Kappa coefficient) ranged from 0.722 to 0.945 for the image features of Diameter<sub>max</sub>, Pseudocapsule, Artery<sub>tumor</sub> and Round mass. The TEV was 106.963 ± 43.315 HU measured by Reader 1, and 107.155 ± 43.425 HU by Reader 2, and CCC was 0.978 with 95%CI (0.964, 0.985). The rTEV was 0.807 ± 0.297 and 0.821 ± 0.293 for Reader 1 and 2, respectively, and CCC was 0.960 with 95%CI (0.938, 0.974).

### 3.3. Texture feature selection

Of the 184 texture features, the 4 selected features were RunLengthNonuniformity at the corticomediullary phase, Contrast (β = 1), GrayLevelNonuniformity and 0.025Quantile at the nephrographic phase. They were chosen as the best valuable texture features for the discrimination of the high from low grade ccRCC in the training cohort (Fig. S1); their weighted coefficients, being present in a linear formula for the Texture-score calculation, were 0.056, 0.428, 0.462 and 0.004, respectively.

To assess the inter-rater agreement of Texture-score, Texture-score was calculated in the 79 cases scanned at CT16 (Table S1), and its CCC was 0.971 with 95% CI (0.955, 0.981).

### 3.4. Performance of individual feature for grading ccRCC in the training cohort

In the training cohort, a significant difference between the low and high grade ccRCC was found for the five features of Round mass, Diameter<sub>max</sub>, Artery<sub>tumor</sub>, rTEV and Texture-score with all \( P < 0.02 \) (Table 2). To discriminate the high from low grade ccRCC, the AUC was 0.757 (95%CI: 0.668, 0.833) for Round mass, 0.723 (95%CI: 0.632, 0.803) for Diameter<sub>max</sub>, 0.709 (95%CI: 0.617, 0.791) for Artery<sub>tumor</sub>, 0.638 (0.543, 0.726) for rTEV, and 0.843 (95%CI: 0.756, 0.920) for Texture-score. The AUC was bigger in Texture-score than Diameter<sub>max</sub> (\( Z = 4.806, P < 0.001 \)), Artery<sub>tumor</sub> (\( Z = 3.091, P = 0.002 \)) and rTEV (\( Z = 2.961, P = 0.003 \)). No significant AUC difference was found between Texture-score and Round mass (\( P > 0.05 \)), and between Diameter<sub>max</sub>, Artery<sub>tumor</sub> and rTEV (all \( P > 0.05 \)).

### 3.5. Development and validation of the predictive models

A logistic regression analysis was performed to establish the predictive models (Table 3). The AUC was different between Model 1 and 2 (\( Z = -2.149, P = 0.032 \)) for differentiating the high from low grade ccRCC in the training cohort. The AUC also was different between Model 1 and 2 (\( Z = -2.510, P = 0.012 \)), and between Model 1 and 3 (\( Z = -2.235, P = 0.025 \)) for grading ccRCC in the validation cohort.

### 4. Discussion

In this study, we sought to determine whether the combination of the non-texture and texture features could help grade ccRCC. We found that the model with texture features could stratify patients according to their risk assessment of having the simplified grade of ccRCC.

Previous studies had indicated the value of radiological imaging in differentiating the high from low grade RCC. The CT imaging-based R.E.N.A.L. nephrometry score can identify the high grade RCC in an independent cohort with an AUC of 0.73 [19]. Magnetic resonance imaging (MRI) have reached a higher predictive accuracy in the pre-operative detection of the high grade RCC, and the AUC was 0.80 for the diffusion weighted imaging in vivo [20]. In the present study, each non-texture feature shown an AUC of < 0.757, being similar to the conclusion in previous studies [19-21].

As one of the non-texture features, rTEV exhibited the largest AUC (0.926) for the presence of the high grade ccRCC in a previous study [14], although its AUC was 0.638 in this study, which may be because of two key differences between the two studies. First, the cases with the high grade ccRCC accounted for 35.088% (40/114) in this study (training cohort) but 16.863% (43/255) in Zhu et al. study. Second, the ROI placement was different between the two studies. In Zhu et al. study, the ROI should include as large as possible the solid component of the tumor, but exclude the gross necrosis. However, a consensus about the detection of intratumoral necrosis on CT images was not reached so far, although the intratumoral necrosis can make the tumor ROI exclude the gross necrosis, but be most likely to include the micro-necrosis focus within the carcinoma. To avoid the micro-necrosis within the lesion as far as possible, Jonathan et al. curved the most avidly enhancing parts of enhancing lesions in the non-texture and texture features could help grade ccRCC.
Table 3

<table>
<thead>
<tr>
<th>Intercept or Variable</th>
<th>Model 1</th>
<th></th>
<th>Model 2</th>
<th></th>
<th>Model 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>β</td>
<td>OR (95% CI)</td>
<td>p</td>
<td>β</td>
<td>OR (95% CI)</td>
<td>p</td>
<td>β</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.253</td>
<td>0.553</td>
<td>2.799</td>
<td>0.003</td>
<td>0.374</td>
<td>0.215</td>
</tr>
<tr>
<td>Round mass</td>
<td>−2.412</td>
<td>0.121 (0.045, 0.338)</td>
<td>&lt; 0.001</td>
<td>−1.335</td>
<td>0.263 (0.084, 0.818)</td>
<td>0.021</td>
</tr>
<tr>
<td>Arterytumor</td>
<td>1.946</td>
<td>7.062 (2.218, 22.105)</td>
<td>&lt; 0.001</td>
<td>NA</td>
<td>−2.278</td>
<td>0.499 (0.267, 0.933)</td>
</tr>
<tr>
<td>Texture-score</td>
<td>NA</td>
<td>1.300</td>
<td>4.619 (1.806, 11.811)</td>
<td>0.001</td>
<td>1.704</td>
<td>7.432 (3.221, 17.148)</td>
</tr>
<tr>
<td>AUC (95% CI)</td>
<td>0.826 (0.748, 0.940)^*</td>
<td>0.878 (0.812, 0.945)^*</td>
<td>0.843 (0.765, 0.920)</td>
<td>0.771 (0.650, 0.892)</td>
<td>0.780 (0.666, 0.894)^*</td>
<td></td>
</tr>
<tr>
<td>Training Cohort</td>
<td>0.671 (0.543, 0.800)^*</td>
<td>0.680 (0.551, 0.821)^*</td>
<td>0.753 (0.613, 0.923)^*</td>
<td>0.751 (0.611, 0.922)^*</td>
<td>0.780 (0.666, 0.894)^*</td>
<td></td>
</tr>
<tr>
<td>Validation Cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: β is the regression coefficient; Arterytumor: intratumoral artery; AUC, area under receiver operating characteristic curve; CI, confidence interval; NA, not available; OR, odds ratio; rTEV, relative enhancement values of the tumor. ^* exhibited a significant difference of AUC value in each row between the two models (P < 0.05). Bold values signify p values are less than 0.05.

relative small AUC of 0.70 [26]. Thus it is a challenge to distinguish nicely the simplified grade of ccRCC using the lesion size. To our knowledge, the great majority of malignant small renal masses (diameter of < 4 cm) have a good prognosis because they are low stage and low grade [27], with even merely active surveillance being an accepted course of action in the selected cases [28]. Therefore, the size of 4 cm in axial maximal tumor diameter is a very important threshold in the surgical and prognosis management of renal tumors. This is why a Diameter_{max} of renal mass was classified into < 4 cm or not in this study although the Diameter_{max} contributed little to improve the predictive model. The limited value of Diameter_{max} in discriminating the high from low grade ccRCC can be attributed to its limited capacity (AUC = 0.723) compared with the combination of other image features for grading ccRCC.

Morphologic feature is very important in the conventional image analysis. Irregular masses have commonly been considered a primary malignant carcinoma in the organs of lung, kidney and breast et al. [29–31] while round masses are seen as an indication of a benign tumor. Therefore the morphologic feature was also included to build a predictive model that can show a promise as a classification tool [32,33]. In this study, the morphologic feature (Round mass) was assessed for each renal mass, and the results showed that irregular mass was also an independent index for the prediction of the high grade ccRCC, which is consistent with that in previous study did by Zhu et al. [14] study.

There were many predictive models in the previous studies exhibiting an excellent performance for grading ccRCC, but the performances of the predictive models were not validated externally in an independent cohort. In this study, the radiomic model with non-texture features improved the discrimination capacity for grading ccRCC, exhibiting a higher AUC of 0.826 in the training cohort. However, the high discrimination capacity of Model 1 did not performed well in the validation cohort (AUC = 0.671). The Texture-score combined with or without non-texture features also improved the capacity of the prediction models for grading ccRCC, and also performed well in the validation cohort. Moreover, no significant difference was found between the Model 2 and 3, indicating the limited value of non-texture features in the predictive model. Therefore, the non-texture features would not be recommended to build the predictive models for grading ccRCC in the future.

The current study has several inherent limitations. First, it has a relatively small sample size in the training cohort. Second, this study investigated the clinical value of the radiomic model in differentiating the high from low grade ccRCC, but not in discriminating the malignant from benign masses which would receive more attention from clinicians. Third, a performance comparison of the regular cross-sectional with whole-mass-based model construction was not conducted in this study, however both 2D and 3D CT radiomic features have a certain prognostic ability in NSCLC and 2D features have exhibited a better performance [34].

In conclusion, we proposed a radiomic model with texture features that can be used to facilitate individual risk prediction for grading ccRCC. Going forward, large, multicenter and prospective studies are needed to validate the model as a clinical tool.

Author contributions

J.D. and Z.X. conceived and designed the project with supervision from W.X. Z.J. and L.P. participated in the data collection and processing. J.Q. and J.C. contributed to the image analysis. All authors edited the manuscript.

Competing financial interest

The authors declare no conflict of interest.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (81777198), the Jiangsu Provincial Medical Youth Talent Program, China (QNRC2016299), the General Program of Jiangsu Provincial Commission of Health and Family Planning, China (H20170003), and the Changzhou Medical Youth Talent Program, Jiangsu, China (QN201610). The authors gratefully acknowledge our colleagues who performed the multiphasic CT scans in the Third Affiliated Hospital of Soochow University.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ejrad.2018.04.013.

References


学霸图书馆
www.xuebalib.com

本文献由“学霸图书馆-文献云下载”收集自网络，仅供学习交流使用。

学霸图书馆（www.xuebalib.com）是一个“整合众多图书馆数据库资源，提供一站式文献检索和下载服务”的24小时在线不限IP图书馆。图书馆致力于便利、促进学习与科研，提供最强文献下载服务。

图书馆导航：
图书馆首页 文献云下载 图书馆入口 外文数据库大全 疑难文献辅助工具