European Journal of Radiology 121 (2019) 108738

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Research article

Clear cell renal cell carcinoma: Machine learning-based computed tomography radiomics analysis for the prediction of WHO/ISUP grade

Jun Shu[a], Didi Wen[a], Yibin Xi[a], Yuwei Xia[b], Zhengting Cai[b], Wanni Xu[c,d], Xiaoli Meng[e], Bao Liu[a], Hong Yin[a,*]

[a] Department of Radiology, Xijing Hospital, Fourth Military Medical University, Changle West Road 127#, Xi'an City, 710032, People's Republic of China
[b] Huiying Medical Technology Co., Ltd. Room C103, B2, Dongsheng Science and Technology Park, HaiDian District, Beijing City, 100192, People's Republic of China
[c] Department of Pathology, Xijing Hospital, Fourth Military Medical University, Changle West Road 127#, Xi'an City, 710032, People's Republic of China
[d] Deng Road 97#, Xi'an City, 710077, People's Republic of China
[e] Department of Radiology, Xi'an XD Group Hospital, Shaanxi University of Chinese Medicine, Feng Deng Road 97#, Xi'an City, 710077, People's Republic of China

ARTICLE INFO

Keywords:
Clear cell renal cell carcinoma
Machine learning
Computed tomography
WHO/ISUP grade

ABSTRACT

Purpose: To evaluate the performance of machine learning (ML)-based computed tomography (CT) radiomics analysis for discriminating between low grade (WHO/ISUP I-II) and high grade (WHO/ISUP III-IV) clear cell renal cell carcinomas (ccRCCs).

Methods: A total of 164 low grade and 107 high grade ccRCCs were retrospectively analyzed in this study. Radiomic features were extracted from corticomedullary phase (CMP) and nephrographic phase (NP) CT images. Intraclass correlation coefficient (ICC) was calculated to quantify the feature's reproducibility. The training and validation cohort consisted of 163 and 108 cases. Least absolute shrinkage and selection operator (LASSO) regression method was used for feature selection. The machine learning (ML) classifiers were k-NearestNeighbor (KNN), Logistic Regression (LR), multilayer perceptron (MLP), Random Forest (RF), and support vector machine (SVM). The performance of classifiers was mainly evaluated and compared by certain metrics.

Results: Seven CMP features (ICC range, 0.990-0.999) and seven NP features (ICC range, 0.931-0.999) were selected. The accuracy of CMP, NP and the combination of CMP and NP ranged from 82.2%–85.9 %, 82.8%–94.5 % and 86.5%–90.8 % in the training cohort, and 90.7%–95.4%, 77.8%–79.6 % and 91.7%–93.5 % in the validation cohort. The AUC of CMP, NP and the combination of CMP and NP ranged from 0.901 to 0.938, 0.912 to 0.976, 0.948 to 0.968 in the training cohort, and 0.957 to 0.974, 0.856 to 0.875, 0.960 to 0.978 in the validation cohort.

Conclusions: ML-based CT radiomics analysis can be used to predict the WHO/ISUP grade of ccRCCs preoperatively.

1. Introduction

Renal cell carcinoma (RCC) accounts for approximately 90 % of renal tumors and nearly 80 % of them are clear cell renal cell carcinomas (ccRCCs) [1,2]. Despite of the fact that most RCCs can be cured by surgery [3,4], preoperative assessment of the aggressiveness of RCCs can help to select the optimal intervention method. In elderly and/or morbid patients with small renal masses and limited life expectancy, radiofrequency ablation and cryoablation may be offered [5]. For selected low risk small renal mass (SRM), active surveillance (AS) can be a safe option [6].

Many grading systems have been proposed for RCC to define its aggressiveness in the past, in which the Fuhrman grading system was the most widely used one [7]. However, this system requires to assess nuclear size, nuclear shape, and nucleolar prominence simultaneously and therefore some cases might be confused when two or even three of the parameters contradict with each other. Moreover, Fuhrman grading system has a poor inter-observer reproducibility and failure to discriminate outcome adequately [8,9]. Previous study has reported that grading system based on nucleolar prominence alone showed a stronger association with patient outcome compared with those relying on Fuhrman grade for ccRCCs and papillary RCCs (pRCCs) [10]. In 2012, the International Society of Urological Pathology (ISUP) introduced a novel grading system for ccRCCs and pRCCs [11], and the system was incorporated in the latest World Health Organization (WHO) renal tumor classification, being designated as WHO/ISUP grading system

* Corresponding author.
E-mail address: yinhong@fmmu.deu.cn (H. Yin).

https://doi.org/10.1016/j.ejrad.2019.108738
Received 6 May 2019; Received in revised form 24 September 2019; Accepted 27 October 2019

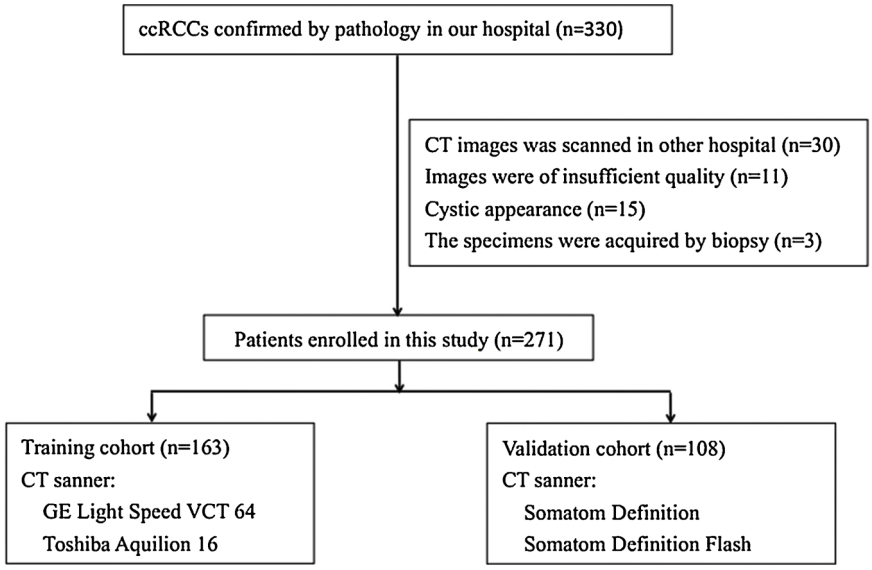

Fig. 1. Flow chart of the patient recruitment process.
Note: ccRCCs, clear cell renal cell carcinomas.

Table 1
CT scanning and reconstruction parameters.

CT scanner	CT 64	CT 16	CT 64	CT 64
Scanner model	Light speed VCT	Aquilion	Definition	Definition Flash
Manufacturer	General Electric	Toshiba	SIEMENS	SIEMENS
Tube voltage (Kv)	120	120	120	120
Tube current (mAs)	Smart mAs 200-380	350mAs, smart mAs	CARE Dose4D 350mAs	CARE Dose4D 350mAs
Collimation (mm)	64 × 0.625	16 × 0.5	64 × 0.6	64 × 0.6
Kernel	standard	FC30	B30f	B30f
Slice thickness (mm)	5	5	5	5
Field of view (mm^2)	350 × 350	350 × 350	350 × 350	350 × 350
Matrix	512 × 512	512 × 512	512 × 512	512 × 512

[12]. Based upon increasing nucleolar prominence and the presence of extreme nuclear pleomorphism and/or tumor giant cells and/or sarcomatoid and/or rhabdoid differentiation, this system grades ccRCCs and pRCCs from level 1–4. The WHO/ISUP grading system has an improved inter-observer reproducibility and prognostic ability compared to Fuhrman grading system [13].

Because imaging methods have a high diagnostic accuracy to characterize renal malignancy, the needle biopsy is not always necessary before surgery [4]. Moreover, due to the heterogeneity of ccRCCs, biopsy has some discordances with resection sample for both Fuhrman and WHO/ISUP grading systems [14,15].

Radiomics, which performs the high-throughput extraction and selection of large number of quantitative features from digital images to construct their relationship to the underlying pathophysiology, has grown exponentially in the field of medical imaging [16], and proved to be an effective method to predict the Fuhrman grade of ccRCCs [17–22]. To the best of our knowledge, there is only one published research that has investigated the correlation between radiomics signature and WHO/ISUP grade of ccRCCs so far [22], however, the study was not convincing enough and had the promotion space, which just used single-slice-based two-dimensional (2D) radiomics with only one unadjusted-parameter classifier constructed by SVM, and without true independent validation cohort obtained using different vendors.

Therefore, the purpose of this study is to investigate the ability of three-dimensional (3D) CT-based radiomics analysis by different machine learning (ML) classifiers to differentiate low and high WHO/ISUP grade of ccRCCs.

2. Materials and methods

2.1. Patients

The institutional review board of our hospital approved this retrospective study and the informed patient consent was waived. Consecutive patients were collected from January 2015 to July 2017. The inclusion and exclusion criteria are presented in Fig. 1. Finally, 271 patients with 271 ccRCCs were included in this study. The training cohort consisted of 163 cases scanned in two different scanners, and the validation cohort consisted of 108 cases in other two. There is a need to mention that the study shared a portion of data (103 patients scanned in GE LightSpeed VCT) with our previous report [19].

2.2. CT technique and WHO/ISUP grade assessment

The images were obtained by one of the four different CT scanners. Patients were injected with non-ionic intravenous contrast material according to their weight (1.0 mL/kg, with a minimum of 60 mL and a maximum of 100 mL) at the rate of 3.5 mL/s via the antecubital vein through a power injector. Corticomedullary phase (CMP) and nephrographic phase (NP) began 25 − 28 s and 65 − 70 s after contrast injection, respectively. The scanning and reconstruction parameters of the four CT scanners are shown in Table 1.

Histopathological evaluation was performed with hematoxylin and eosin staining, along with immunohistochemistry necessarily. A pathologist with 10 years of experience specializing in renal pathology re-examined all the specimens according to the WHO/ISUP grade criteria [12]. WHO/ISUP III and IV were grouped as high grade, and WHO/

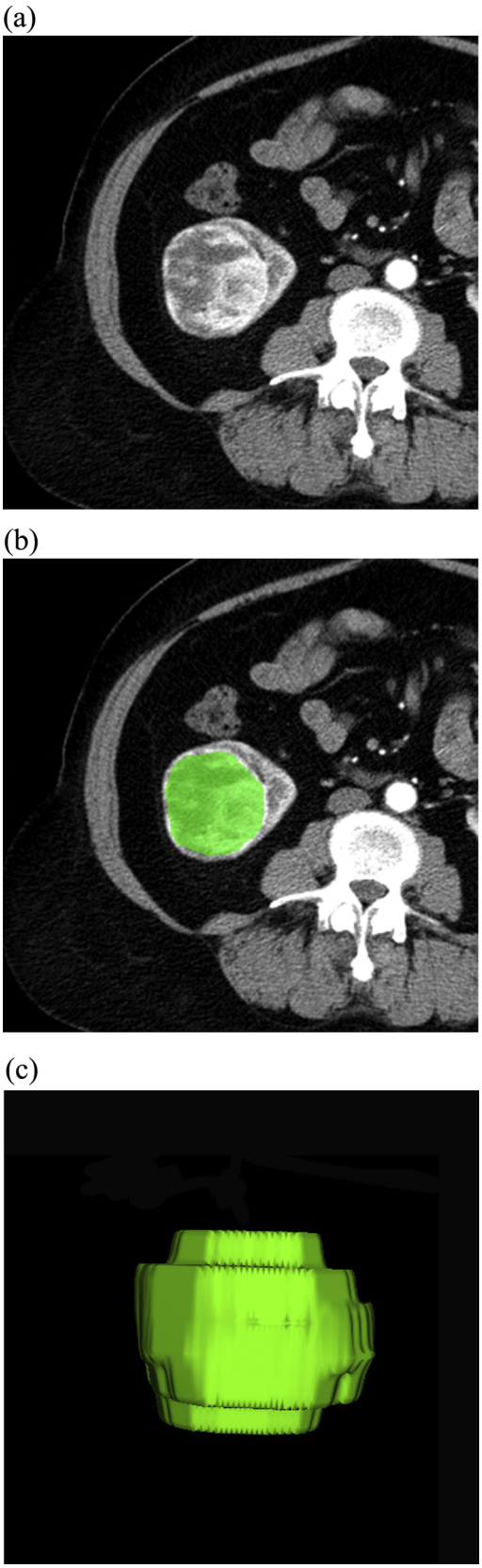

Fig. 2. An example of the manual segmentation in corticomedullary phase. A renal mass with heterogeneous enhancement in right kidney (a), manual segmentation on the same axial slice (b), three-dimensional (3D) volumetric reconstruction of the tumor (c).

ISUP I and II were grouped as low grade.

2.3. Tumor segmentation

To obtain the volume of interest (VOI) for further analysis, we uploaded the image data of DICOM format to Radcloud platform (version 3.1.0, http://radcloud.cn/, Huiying Medical Technology Co., Ltd, Beijing, China) which would automate anonymization, then delineated the outline of all contiguous slices of the entire tumor except for the first and last one to minimize the partial volume effect. Contouring was drawn slightly within the borders of the tumor masses on CMP and NP, including necrotic, cystic change and hemorrhagic areas (Fig. 2).

2.4. Image preprocessing and feature extraction

Image feature extraction was also implemented in the Radcloud platform. Before feature extraction, to eliminate the variance caused by different scanner acquisition, avoid anisotropic resolution and improve the reproducibility [23], original images of each patient were preprocessed: normalization by centering it at the mean with standard deviation, resample into the voxel size of $1 \times 1 \times 1 mm^3$ using B-Spline interpolation, and discretization of the gray level by a fixed bin width of 25 in the histogram.

Then, the radiomics feature pool including 1029 kinds of features was extracted from each VOI of each patient, which were classified as five subgroups of first order statistic, shape, gray level co-occurrence matrix (GLCM), gray level run length matrix (GLRLM) and gray level size zone matrix (GLSZM), based on the preprocessed images and those transformed by exponential, square, square root, logarithm and wavelet (wavelet-LHL、wavelet-LHH、wavelet-HLL、wavelet-LLH、wavelet-HLH、wavelet－HHH、wavelet－HHL、wavelet-LLL). Details of the radiomics features are accordance with our previous research [19].

2.5. Feature standardization and selection

Prior to the steps of feature selection, we use a common way to get all feature to have the same distribution by standardization: subtract the mean value and divided by the standard deviation.

Feature selection was performed in two steps to select the optimal grade-related features via the training cohort. Firstly, we needed to identify features that were predictive, stable and reproducible to construct generalizable classifiers. To do this, two radiologists independently drew VOIs in 60 randomly selected patients. Intra-class correlation coefficient (ICC) was calculated for each feature. Following the common strategy in previous studies [17,24–26], we considered a feature with ICC of 0.80 or greater as excellent reproducibility and retained it. Then, the least absolute shrinkage and selection operator (LASSO) regression method was used to decrease the high degree of redundancy and irrelevance. To select optimal feature subset in LASSO regression, we performed 5-fold cross-validation based on binomial deviance minimization criteria from the training cohort with the maximum iteration of 5000.

2.6. Classification

To compare the influences of different ML algorithms on the predictive performance of the model, the different algorithms including k-nearest neighbor (KNN), logistic regression (LR), multilayer perceptron (MLP), random forest (RF) and support vector machine (SVM), were applied to develop radiomics-based diagnostic classifiers using the selected CMP features, NP features and their combination in the training cohort. In the process of model building, every classifier was tuned and the hyperparameters were optimized to maximize the prediction performance. Details of optimal hyperparameters of all classifiers are shown in Supplementary S1.

Hyperparameter optimization comes at a cost. On the one hand, it

Table 2
Patient characteristics in the training and validation cohorts.

Patient characteristics	Training cohort (n = 163)	Validation cohorts (n = 108)	*p*
Age (yrs, mean ± SD)	57.2 ± 11.3	57.4 ± 10.5	0.788
Gender			
Male	98	74	0.160
Female	65	34	
WHO/ISUP grade			
Low grade	98	74	0.160
High grade	65	34	
Tumor size (cm, mean ± SD)	5.3 ± 1.9	5.2 ± 1.7	0.649

Note: yrs, years; SD, standard deviation; *p* from Chi-square test or Student's t test, as appropriate.

tunes the model to be adapted well to the underlying data from training cohort. On the other hand, the performance of the tuned model may be over-fit due to a favorable selection of hyperparameters. In order to estimate the generalization performance of a model, we divided the data into training cohort and validation cohort. The training cohort consisted of 163 cases scanned in CT1 and CT2, and the validation cohort consisted of 108 cases scanned in CT3 and CT4.

The performance of the radiomics classifiers, measuring how well the model could distinguish patients between the low-grade or high-grade groups, was represented by receiver operating characteristic (ROC) curves with the associated areas under the ROC curve (AUC), accuracy, sensitivity and specificity.

All of feature standardization, selection and model building with the related statistics and graphing were done in the Anaconda3 platform (https://www.anaconda.com/) using the programming language of Python 3.6 (https://www.python.org/) and the packages of 'scikit-learn' (https://scikit-learn.org/) and 'matplotlib' (https://matplotlib.org/). Statistical analysis of clinical information was performed using SPSS software (version 20.0; SPSS, Inc., Chicago, IL, USA). The tests in our study were two-tailed and $p < 0.05$ was considered as statistical significance.

3. Results

3.1. Clinical characteristics

Of the 271 patients, there were 172 males and 99 females with a mean age of 57.3 ± 10.9 years and range from 23 to 85 years. Patient characteristics of the training and validation cohort are shown in Table 2, and showed no significant differences between the two cohorts. However, tumor size between low-grade and high-grade groups showed statistical difference (4.7 ± 1.5 cm vs 6.2 ± 2.0 cm, p < 0.001, t test).

3.2. The selection of radiomic features

Among 1029 radiomics features in each original feature set of CMP and NP, 102 and 144 features were excluded due to the ICCs lower than 0.80, respectively. After LASSO analysis, seven features were further selected both in CMP and NP with excellent ICCs, ranging from 0.990 to 0.999 and 0.931 to 0.999, respectively. The details of the selected features are shown in Table 3 (ICCs), Fig. 3 (the coefficients in the LASSO models) and Fig. 4 (heat maps between low-grade and high-grade groups).

3.3. The predictive performance of the ML classifiers

In each classifier, CMP model, NP model and the combined model were respectively constructed based on the selected CMP features, NP features, and their combination. The results of ML-based CT radiomics analysis to discriminate low and high WHO/ISUP grade ccRCCs in the training and validation cohorts are summarized in Tables 4 and 5. In general, all classifiers achieved a satisfying performance, especially in the combined model, where accuracy and AUC ranged from 91.7%–93.5 % and 0.960 to 0.978 in the validation cohort, compared to 86.5%–90.8 % and 0.948 to 0.968 in the training cohort.

4. Discussion

WHO/ISUP grade is a validated grading system for ccRCCs, with gradually worse outcome when the grade rises from 1 to 4 [13]. Therefore, preoperative prediction of WHO/ISUP grade is very important for clinical decision making. In this study, the ML classifiers based on CT radiomics features achieved satisfactory performance to discriminate between low and high WHO/ISUP grade of ccRCCs. The performance was nip and tuck among classifiers built by different ML algorithms, and generally more than 90 % of the ccRCCs were correctly classified.

Recent studies have reported the discrimination of Fuhrman grade by CT radiomics/texture analysis [17–21]. However, the potential value of CT radiomics analysis to predict WHO/ISUP grade of ccRCCs was only presented in one published research [22]. Compared to these studies, our study had some improvements and differences.

Firstly, the WHO/ISUP grading system had replaced Fuhrman grading system because of it's better reproducibility and prognostic ability [13]. Secondly, our study had an independent validation cohort in different CT scanners and therefore the results became more robust and convincing. Thirdly, the performance was improved dramatically, AUCs of classifiers using the features combining CMP an NP was more than 95.0. % in the validation cohort, compared with those from 76.6%–86.0 % in the previous Fuhrman grade studies [17–19]. Fourthly, we didn't evaluate the performance of the subjective findings, mainly because subjective findings were more easily affected by physician's experience, the inter-reader agreement was low or just fair

Table 3
ICCs of the selected features in CMP and NP images.

CMP Features	ICCs	NP Features	ICCs
First-order features		Firstorder features	
original_90Percentile	0.999	original_RootMeanSquared	0.998
wavelet-LLH_Kurtosis	0.988	squareroot_RootMeanSquared	0.999
wavelet-LLH_Uniformity	0.990	wavelet-HLL_Mean	0.931
Texture features		Texture features	
original_glrlm_Gray Level Non Uniformity Normalized	0.993	original_glcm_Idn	0.933
original_glrlm_Run Length Non Uniformity Normalized	0.997	squareroot_glszm_Large Area Emphasis	0.997
wavelet-LLH_glcm_Maximum Probability	0.992	wavelet-LLH_glcm_Energy	0.994
wavelet-LLL_glrlm_Run Length Non Uniformity Normalized	0.995	wavelet-LLL_glrlm_Short Run Emphasis	0.993

Note: ICC, Intra-class correlation coefficient; CMP, corticomedullary phase; NP, nephrographic phase; the representation of feature format: original/transformation_subgroup_name.

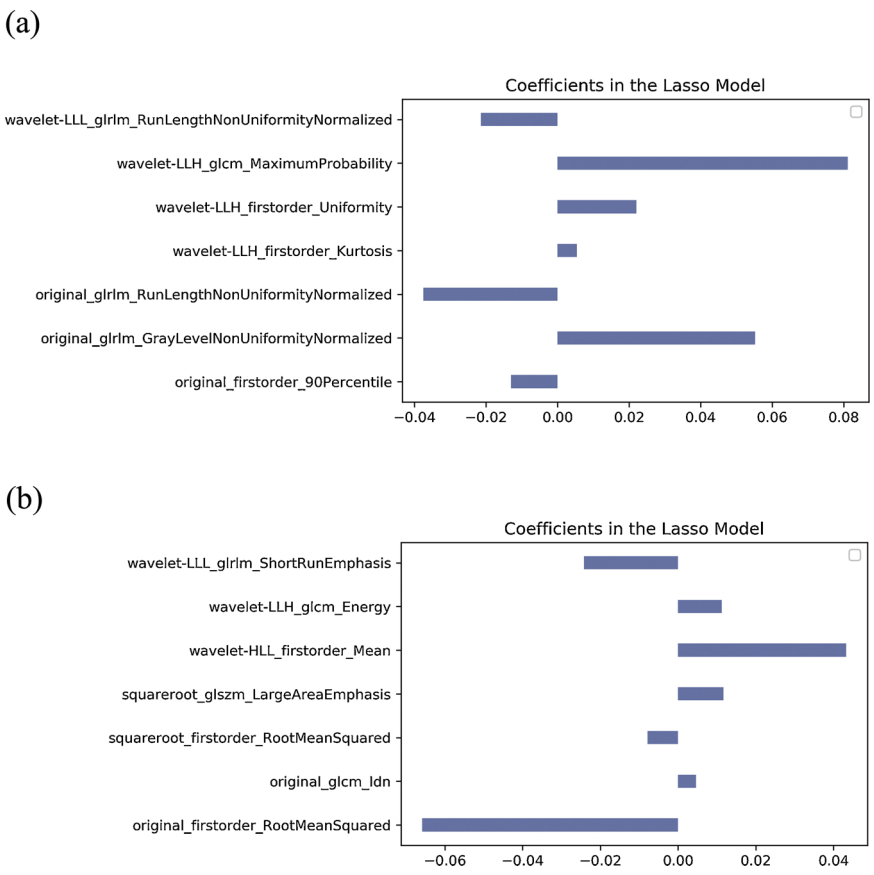

Fig. 3. The coefficients of selected features in the LASSO models of corticomedullary phase (a) and nephrographic phase (b). Bar chart lists the selected radiomics features on the y-axis with their respective coefficients in the LASSO model on the x-axis. Generally, the larger the absolute value of the coefficient is, the greater contribution to model the feature makes. Note: LASSO, least absolute shrinkage and selection operator.

[26,27]. Fifthly, Single-slice-based 2D CT texture analysis of renal masses is more sensitive to manual segmentation [28], and texture features extracted by the segmentation with margin shrinkage of 2 mm had the more excellent reproducibility than those of contour-focused segmentation [29], therefore the contouring of VOI based on 3D images was drawn slightly within the borders of the tumor in our study. Although low-grade ccRCCs had smaller size than high-grade, the wide overlap (4.7 ± 1.5 cm for low-grade vs 6.2 ± 2.0 cm for high-grade) makes this feature valueless in clinical practice.

The results of feature selection showed that three first-order features and four texture features both in CMP and NP were significantly associated with WHO/ISUP grade, and the diagnosis ability of texture features was superior to first-order features according to the coefficients in LASSO models. Texture features analysis had proved to be an effective method in discriminating malignant and benign renal masses and predicting their aggressiveness [17–19,30,31]. And we do have something of the selected features in common despite of various strategies of feature selection among these studies.

To be specific, lesion heterogeneity is a known feature of malignancy, increased heterogeneity along with increased aggressiveness [33]. GLRLM features counting consecutive voxels/pixels with the same gray-level value in any given direction, had two features related to the gray-level heterogeneity, namely run-length nonuniformity and gray-level nonuniformity [32]. The studies found there were differentiated GLRLM features distribution between each other of fat-poor renal angiomyolipoma, low and high Fuhrman grade ccRCCs and sarcomatoid RCC [17,18,26,27,30], and also contributed to differentiate high from low WHO/ISUP grade ccRCCs in our study. Furthermore, GLCM features as the most commonly used texture parameters in texture analysis, are constructed using the number, distance and angle of a combination of grey levels derived from CT images [34]. 'GLCM_MaximumProbability' of CMP, 'GLCM_Idn' and 'GLCM_Energy' of NP were finally selected by our experiments, which was in accordance with the finding of previous studies that GLCM features were strongly correlated with the aggressiveness of renal mass [31].

Our study has important practical implications. WHO/ISUP grade system is a well-recognized prognostic factor for ccRCCs and pRCCs, with superior performance than Fuhrman grade system [13]. If low WHO/ISUP grade of ccRCCs can be identified preoperatively, such patients will be the candidates for AS and vice versa, especially for those with small masses. Moreover, considering the high accuracy of 0.865 to 0.908 in training cohort and 0.917 to 0.935 in validation cohort we achieved, comparable with 88 % that the fine-needle aspiration (FNA) got [15], this non-invasive method of CT-based radiomics analysis can be an alternative way of FNA to predict the aggressiveness of ccRCCs.

There are some inevitable limitations in the research. Firstly, this was a retrospective study so the selection bias cannot be fully avoided, nevertheless, almost all of the current radiomics or texture studies are retrospective in nature. Secondly, external validation in different centers with more samples are needed to validate the robustness and generalization although we had independent validation cohort in different scanners, and the class imbalance in the validation set would influence the performance metrics to some extent. Thirdly, it was a bit cumbersome to draw three-dimensional VOI compared to two-dimensional ROI, however three-dimensional analysis appeared more representative of tumor heterogeneity [35], and we are confident there will be semi-automated software to recognize RCCs in CT images in the future. Fourthly, our study didn't include the non-contrast CT images, mainly because it is difficult to identify the border of some ccRCCs without enhancement based on our experience, though there was report that unenhanced CT texture analysis could predict the nuclear grade of ccRCCs [36].

In conclusion, ML-based CT radiomics classifiers can be a complementary tool to predict the WHO/ISUP grade of ccRCCs preoperatively, which might make contribution to personalized treatment.

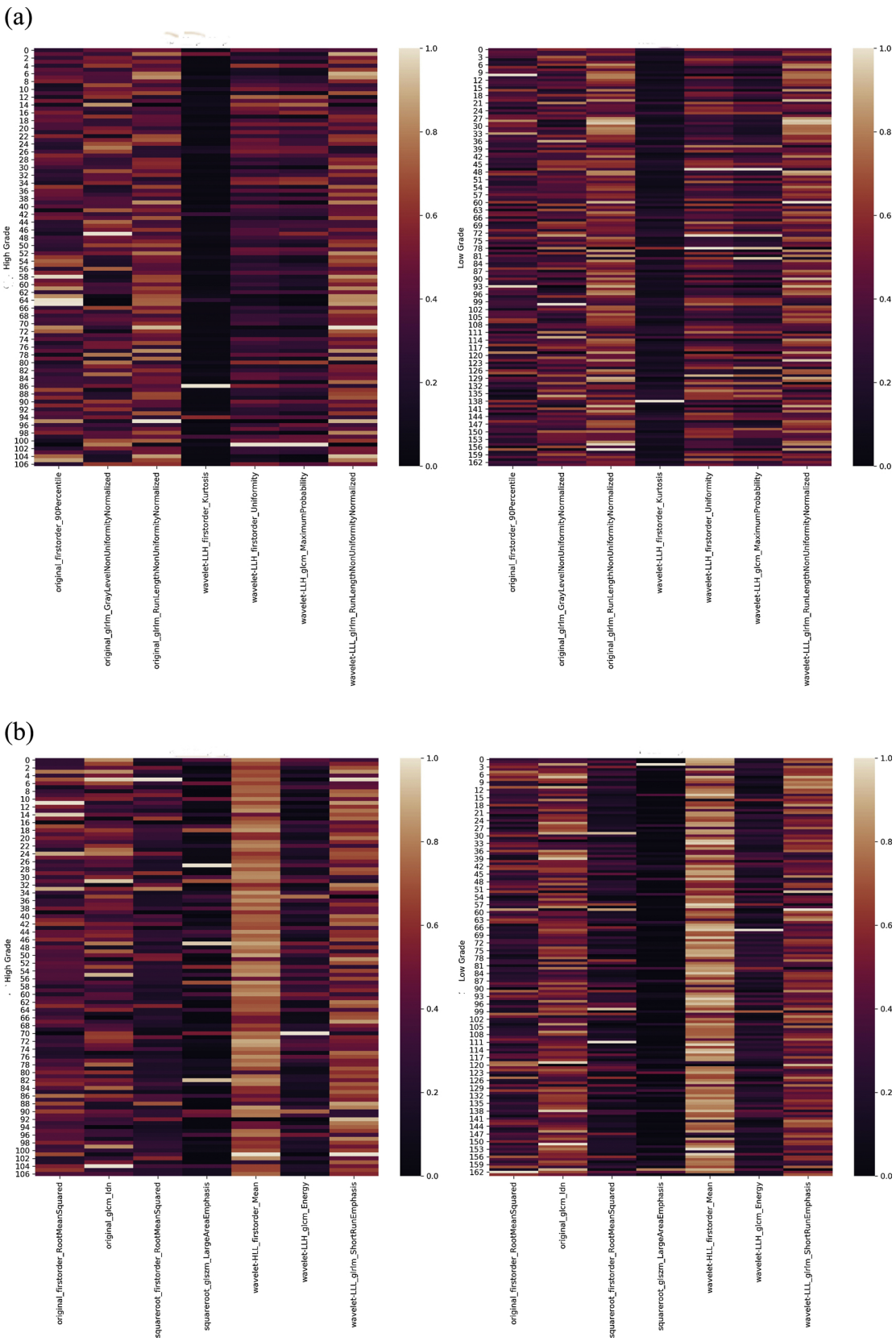

Fig. 4. The heat maps of selected standardized features of corticomedullary phase (a) and nephrographic phase (b) between low and high WHO/ISUP grade ccRCCs. Difference in colors and their shades indicates dissimilarity of the corresponding radiomics feature value. Note: WHO, World Health Organization; ISUP, International Society of Urological Pathology; ccRCCs, clear cell renal cell carcinomas.

Table 4
Diagnostic performance of machine learning-based CT radiomics classifiers to discriminate low and high WHO/ISUP grade ccRCCs in training cohort.

Classifiers	Accuracy (%)	Sensitivity	Specificity	AUC
KNN				
CMP	83.4 (77.9-87.4)	0.847	0.815	0.906 (0.867-0.940)
NP	84.0 (78.3-88.3)	0.847	0.831	0.917 (0.882-0.948)
CMP + NP	86.5 (79.5-92.2)	0.847	0.892	0.948 (0.923-0.970)
LR				
CMP	84.0 (77.5-89.1)	0.837	0.846	0.915 (0.875-0.951)
NP	83.4 (73.8-92.8)	0.786	0.908	0.912 (0.873-0.946)
CMP + NP	90.2 (81.8-97.7)	0.857	0.969	0.964 (0.943-0.981)
MLP				
CMP	82.8 (74.0-90.7)	0.796	0.877	0.901 (0.861-0.937)
NP	82.8 (71.3-96.1)	0.745	0.954	0.919 (0.881-0.951)
CMP + NP	90.8 (84.7-95.6)	0.888	0.938	0.966 (0.944-0.983)
RF				
CMP	85.9 (76.9-94.1)	0.816	0.923	0.938 (0.907-0.965)
NP	94.5 (92.4-95.9)	0.949	0.938	0.976 (0.958-0.990)
CMP + NP	89.6 (84.3-93.5)	0.888	0.908	0.968 (0.947-0.983)
SVM				
CMP	82.2 (73.7-89.7)	0.796	0.862	0.903 (0.861-0.940)
NP	85.9 (82.8-87.9)	0.888	0.815	0.923 (0.884-0.955)
CMP + NP	90.8 (82.9-97.7)	0.867	0.969	0.964 (0.941-0.982)

Note: 95 % confidence interval in parentheses; WHO, World Health Organization; ISUP, International Society of Urological Pathology; ccRCC, clear cell renal cell carcinoma; KNN, k-nearest neighbor; LR, logistic regression; MLP, multi-layer perceptron; RF, random forest; SVM, support vector machine; AUC, area under the curve; CMP, corticomedullary phase; NP, nephrographic phase.

Table 5
Diagnostic performance of machine learning-based CT radiomics classifiers to discriminate low and high WHO/ISUP grade ccRCCs in validation cohort.

Classifiers	Accuracy (%)	Sensitivity	Specificity	AUC
KNN				
CMP	90.7 (87.8-95.2)	0.894	0.929	0.959 (0.930-0.984)
NP	77.8 (66.7-88.9)	0.727	0.857	0.861 (0.803-0.914)
CMP + NP	92.6 (88.6-95.3)	0.924	0.929	0.974 (0.952-0.991)
LR				
CMP	90.7 (86.4-93.8)	0.909	0.905	0.961 (0.932-0.986)
NP	79.6 (70.8-86.7)	0.788	0.810	0.861 (0.804-0.916)
CMP + NP	91.7 (88.4-93.8)	0.924	0.905	0.960 (0.928-0.986)
MLP				
CMP	95.4 (92.9-100.0)	1.000	0.881	0.974 (0.947-0.994)
NP	77.8 (65.0-93.8)	0.682	0.909	0.869 (0.818-0.920)
CMP + NP	93.5 (90.7-95.3)	0.939	0.929	0.978 (0.957-0.995)
RF				
CMP	92.6 (88.6-95.3)	0.924	0.929	0.972 (0.951-0.990)
NP	79.6 (738-83.3)	0.833	0.738	0.875 (0.824-0.925)
CMP + NP	93.5 (889-96.8)	0.924	0.952	0.976 (0.957-0.993)
SVM				
CMP	90.7 (82.0-98.3)	0.864	0.976	0.957 (0.924-0.987)
NP	77.8 (68.0-86.2)	0.758	0.810	0.856 (0.798-0.911)
CMP + NP	92.6 (91.4-94.7)	0.970	0.857	0.971 (0.948-0.989)

Note: 95 % confidence interval in parentheses; WHO, World Health Organization; ISUP, International Society of Urological Pathology; ccRCC, clear cell renal cell carcinoma; KNN, k-nearest neighbor; LR, logistic regression; MLP, multi-layer perceptron; RF, random forest; SVM, support vector machine; AUC, area under the curve; CMP, corticomedullary phase; NP, nephrographic phase.

Declaration of Competing Interest

None.

Acknowledgement

This work was supported by the Discipline promotion project of Xijing hospital (No. XJZT18ML76).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ejrad.2019.108738.

References

[1] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2016, CA Cancer J. Clin. 66 (2016) 7–30.
[2] H. Moch, T. Gasser, M.B. Amin, et al., Prognostic utility of the recently recommended histologic classifcation and revised TNM staging system of renal cell carcinoma: a Swiss experience with 588 tumors, Cancer 89 (2000) 604–614.
[3] B. Ljungberg, L. Albiges, Y. Abu-Ghanem, et al., European Association of Urology Guidelines on Renal Cell Carcinoma: The 2019 Update, Eur. Urol. 75 (5) (2019) 799–810.
[4] R.J. Motzer, E. Jonasch, N. Agarwal, et al., Kidney Cancer, version 2.2017, NCCN clinical practice guidelines in oncology, J. Compr. Canc. Netw. 15 (June (6)) (2017) 804–834.
[5] M.W. Vetterlein, T. Jindal, A. Becker, M. Regier, L.A. Kluth, D. Tilki, F.K. Chun, Small renal masses in the elderly: contemporary treatment approaches and comparative oncological outcomes of nonsurgical and surgical strategies, Investig. Clin. Urol. 57 (4) (2016) 231–239.
[6] M.A. Jewett, K. Mattar, J. Basiuk, et al., Active surveillance of small renal masses: progression patterns of early stage kidney cancer, Eur. Urol. 60 (July (1)) (2011) 39–44.
[7] S.A. Fuhrman, L.C. Lasky, C. Limas, Prognostic significance of morphologic parameters in renal cell carcinoma, Am. J. Surg. Pathol. 6 (October (7)) (1982) 655–663.
[8] B. Delahunt, Advances and controversies in grading and staging of renal cell carcinoma, Mod. Pathol. 22 (June (Suppl 2)) (2009) S24–36.
[9] J.I. López, Gleason and Fuhrman no longer make the grade, Histopathology 69 (2) (2016) 340–341.
[10] B. Delahunt, D. Sika-Paotonu, P.B. Bethwaite, et al., Grading of clear cell renal cell carcinoma should be based on nucleolar prominence, Am. J. Surg. Pathol. 135 (2011) 1134–1139.
[11] B. Delahunt, J.C. Cheville, G. Martignoni, et al., The International Society of Urological Pathology (ISUP) grading system for renal cell carcinoma and other prognostic parameters, Am. J. Surg. Pathol. 37 (10) (2013) 1490–1504.
[12] H. Moch, A.L. Cubilla, P.A. Humphrey, et al., The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs-Part A: Renal, Penile, and Testicular Tumours, Eur. Urol. 70 (1) (2016) 93–105.
[13] J. Dagher, B. Delahunt, N. Rioux-Leclercq, et al., Clear cell renal cell carcinoma: validation of World Health Organization/International Society of Urological Pathology grading, Histopathology 71 (6) (2017) 918–925.
[14] L. Marconi, S. Dabestani, T.B. Lam, et al., Systematic review and meta-analysis of diagnostic accuracy of percutaneous renal tumour biopsy, Eur. Urol. 69 (4) (2016) 660–673.
[15] C.M. Perrino, H.M. Cramer, S. Chen, et al., World Health Organization (WHO)/International Society of Urological Pathology (ISUP) grading in fine-needle aspiration biopsies of renal masses, Diagn. Cytopathol. 46 (11) (2018) 895–900.
[16] R.J. Gillies, P.E. Kinahan, Hricak H. Radiomics, Images are more than pictures, they are data, Radiology 278 (2) (2016) 563–577.
[17] C.T. Bektas, B. Kocak, A.H. Yardimci, et al., Clear cell renal cell carcinoma: machine learning-based quantitative computed tomography texture analysis for prediction of fuhrman nuclear grade, Eur. Radiol. 29 (3) (2019) 1153–1163.
[18] J. Ding, Z. Xing, Z. Jiang, et al., CT-based radiomic model predicts high grade of clear cell renal cell carcinoma, Eur. J. Radiol. 103 (June) (2018) 51–56.
[19] J. Shu, Y. Tang, J. Cui, et al., Clear cell renal cell carcinoma: CT-based radiomics features for the prediction of Fuhrman grade, Eur. J. Radiol. 109 (December) (2018) 8–12.
[20] Z. Feng, Q. Shen, Y. Li, Z. Hu, CT texture analysis: a potential tool for predicting the Fuhrman grade of clear-cell renal carcinoma, Cancer Imaging 19 (1) (2019) 6.
[21] B. Kocak, E.S. Durmaz, E. Ates, M.B. Ulusan, Radiogenomics in clear cell renal cell carcinoma: machine learning-based high-dimensional quantitative CT texture analysis in predicting PBRM1 mutation status, AJR Am. J. Roentgenol. 212 (3) (2019) W55–W63.
[22] X. Sun, L. Liu, K. Xu, et al., Prediction of ISUP grading of clear cell renal cell carcinoma using support vector machine model based on CT images, Bull. Sch. Med. Md 98 (April (14)) (2019) e15022.
[23] R.T.H. Leijenaar, G. Nalbantov, S. Carvalho, et al., The effect of SUV discretization in quantitative FDG-PET Radiomics: the need for standardized methodology in tumor texture analysis, Sci. Rep. 5 (2015) 11075.
[24] Kocak B, Yardimci AH, Bektas CT, et al. Textural differences between renal cell carcinoma subtypes: Machine learning-based quantitative computed tomography texture analysis with independent external validation.
[25] G. Bier, S. Bier, M.N. Bongers, et al., Value of computed tomography texture analysis for prediction of perioperative complications during laparoscopic partial nephrectomy in patients with renal cell carcinoma, PLoS One 13 (April (4)) (2018) e0195270.
[26] N. Schieda, R.E. Thornhill, M. Al-Subhi, et al., Diagnosis of sarcomatoid renal cell carcinoma with CT: evaluation by qualitative imaging features and texture analysis, AJR 204 (5) (2015) 1013–1023.
[27] S.Y. Choi, D.J. Sung, K.S. Yang, et al., Small (< 4 cm) clear cell renal cell carcinoma: correlation between CT findings and histologic grade, Abdom Radiol (NY) 41

(6) (2016) 1160–1169.
[28] B. Kocak, E.S. Durmaz, O.K. Kaya, et al., Reliability of Single-Slice-Based 2D CT Texture Analysis of Renal Masses: Influence of Intra- and Interobserver Manual Segmentation Variability on Radiomic Feature Reproducibility, AJR Am. J. Roentgenol. 7 (May) (2019) 1–7.
[29] B. Kocak, E. Ates, E.S. Durmaz, et al., Influence of segmentation margin on machine learning-based high-dimensional quantitative CT texture analysis: a reproducibility study on renal clear cell carcinomas, Eur. Radiol. (February (12)) (2019).
[30] Z. Feng, P. Rong, P. Cao, et al., Machine learning-based quantitative texture analysis of CT images of small renal masses: differentiation of angiomyolipoma without visible fat from renal cell carcinoma, Eur. Radiol. 28 (4) (2018) 1625–1633.
[31] T. Hodgdon, M.D. McInnes, N. Schieda, et al., Can Quantitative CT Texture Analysis be Used to Differentiate Fat-poor Renal Angiomyolipoma from Renal Cell Carcinoma on Unenhanced CT Images? Radiology 276 (3) (2015) 787–796.
[32] M. Galloway, Texture analysis using gray level run lengths, Comput Graph Image Proc. 4 (2) (1975) 172–179.
[33] B. Ganeshan, E. Panayiotou, K. Burnand, et al., Tumour heterogeneity in non-small cell lung carcinoma assessed by CT texture analysis: a potential marker of survival, Eur. Radiol. 22 (4) (2012) 796–802.
[34] G. Lee, H.Y. Lee, H. Park, et al., Radiomics and its emerging role in lung cancer research, imaging biomarkers and clinical management: state of the art, Eur. J. Radiol. 86 (January) (2017) 297–307.
[35] F. Ng, R. Kozarski, B. Ganeshan, V. Goh, Assessment of tumor heterogeneity by CT texture analysis: can the largest cross-sectional area be used as an alternative to whole tumor analysis? Eur. J. Radiol. 82 (2) (2013) 342–348.
[36] B. Kocak, E.S. Durmaz, E. Ates, et al., Unenhanced CT texture analysis of clear cell renal cell carcinomas: a machine learning-based study for predicting histopathologic nuclear grade, AJR Am. J. Roentgenol. 11 (April) (2019) W1–W8.