Retrospective Cohort Study

INTERNATIONAL JOURNAL OF SURGERY

OPEN

Incremental value of automatically segmented perirenal adipose tissue for pathological grading of clear cell renal cell carcinoma: a multicenter cohort study

Shichao Li, MD[a], Ziling Zhou, MD, PhD[a], Mengmeng Gao, MD[a], Zhouyan Liao, MD[a], Kangwen He, MD[a], Weinuo Qu, MD[a], Jiali Li, MD, PhD[a], Ihab R Kamel, MD, PhD[d], Qian Chu, MD, PhD[b], Qingpeng Zhang, PhD[c], Zhen Li, MD, PhD[a,*]

Objectives: Accurate preoperative prediction of the pathological grade of clear cell renal cell carcinoma (ccRCC) is crucial for optimal treatment planning and patient outcomes. This study aims to develop and validate a deep-learning (DL) algorithm to automatically segment renal tumours, kidneys, and perirenal adipose tissue (PRAT) from computed tomography (CT) images and extract radiomics features to predict the pathological grade of ccRCC.

Methods: In this cross-ethnic retrospective study, a total of 614 patients were divided into a training set (383 patients from the local hospital), an internal validation set (88 patients from the local hospital), and an external validation set (143 patients from the public dataset). A two-dimensional TransUNet-based DL model combined with the train-while-annotation method was trained for automatic volumetric segmentation of renal tumours, kidneys, and visceral adipose tissue (VAT) on images from two groups of datasets. PRAT was extracted using a dilation algorithm by calculating voxels of VAT surrounding the kidneys. Radiomics features were subsequently extracted from three regions of interest of CT images, adopting multiple filtering strategies. The least absolute shrinkage and selection operator (LASSO) regression was used for feature selection, and the support vector machine (SVM) for developing the pathological grading model. Ensemble learning was used for imbalanced data classification. Performance evaluation included the Dice coefficient for segmentation and metrics such as accuracy and area under curve (AUC) for classification. The WHO/International Society of Urological Pathology (ISUP) grading models were finally interpreted and visualized using the SHapley Additive exPlanations (SHAP) method.

Results: For automatic segmentation, the mean Dice coefficient achieved 0.836 for renal tumours and 0.967 for VAT on the internal validation dataset. For WHO/ISUP grading, a model built with features of PRAT achieved a moderate AUC of 0.711 (95% CI, 0.604–0.802) in the internal validation set, coupled with a sensitivity of 0.400 and a specificity of 0.781. While model built with combination features of the renal tumour, kidney, and PRAT showed an AUC of 0.814 (95% CI, 0.717–0.889) in the internal validation set, with a sensitivity of 0.800 and a specificity of 0.753, significantly higher than the model built with features solely from tumour lesion (0.760; 95% CI, 0.657–0.845), with a sensitivity of 0.533 and a specificity of 0.767.

Conclusion: Automated segmentation of kidneys and visceral adipose tissue (VAT) through TransUNet combined with a conventional image morphology processing algorithm offers a standardized approach to extract PRAT with high reproducibility. The radiomics features of PRAT and tumour lesions, along with machine learning, accurately predict the pathological grade of ccRCC and reveal the incremental significance of PRAT in this prediction.

Keywords: automated segmentation, clear cell renal cell carcinoma, machine learning, radiomics analysis, WHO/ISUP grade

[a]Department of Radiology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, [b]Department of Oncology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, [c]Department of Pharmacology and Pharmacy, LKS Faculty of Medicine, and the Musketeers Foundation Institute of Data Science, University of Hong Kong, Hong Kong, China and [d]Department of Radiology, University of Colorado Denver Anschutz Medical Campus, Aurora, CO, USA

S.L. and Z.Z. contributed equally to this work

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

**Corresponding author. Address: Department of Radiology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology,1095 Jiefang Avenue, Wuhan 430030, Hubei, China. Office Tel.: +860 278 366 3738; fax:+862 783 662 314. E-mail: zhenli@hust.edu.cn (Z. Li).*

International Journal of Surgery (2024) 110:4221–4230

Received 17 December 2023; Accepted 4 March 2024

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.lww.com/international-journal-of-surgery.

Published online 3 April 2024

http://dx.doi.org/10.1097/JS9.0000000000001358

Introduction

Clear cell renal cell carcinoma (ccRCC) represents the most prevalent subtype of kidney cancer, accounting for ~80% of all renal cell carcinoma cases[1,2]. However, a considerable proportion of these tumours are indolent, and opting for partial or radical nephrectomy can lead to a medical burden and impairment of renal function, making active surveillance a more suitable option[3]. Accurate prediction of the WHO/International Society of Urological Pathology (ISUP) grade of ccRCC is of paramount clinical significance as it is closely associated with tumour aggressiveness and patient outcomes[4]. High-grade ccRCC, in contrast to low-grade cases, is characterized by increased malignancy, a higher risk of recurrence and metastasis, and a less favourable prognosis[5]. This distinction becomes particularly crucial for patients with multiple comorbidities, where prior knowledge of pathological grading informs critical clinical interventions[6].

Computed tomography (CT) has emerged as an indispensable imaging modality within the diagnosis and treatment of ccRCC patients. Beyond its capability to assess the kidneys and renal tumours, CT offers a wealth of body composition data[7]. By extracting quantitative radiomic features from a specific volume of interest (VOI), it becomes possible to capture subtle voxel intensity variations and reveal concealed tumour attributes that are imperceptible to the naked eye. Furthermore, the utilization of multiple image filters highlights various aspects of image information, thereby enriching the spectrum of feature selection and enabling the comprehensive exploration of imaging data[8].

Perirenal adipose tissue (PRAT), a part of visceral fat separated from the peritoneum by the renal fascia, plays an active role in metabolism and adipokine secretion[9]. However, its recognition and unified segmentation manually are challenging, time-consuming, and labour-intensive. This limitation obstructs establishing a unified definition range and investigating perirenal adipose tissue correspondingly. In the specific task of tumour segmentation training, annotating volumetric image slices at the pixel level is labour-intensive[10,11]. Currently, various deep-learning segmentation algorithms has solved the labelling problem to a large extent. Medical Imaging Computing and Computer Assisted Intervention (MICCAI) Society developed the KiTS19 (Kidney Tumour Segmentation 2019) Grand Challenge, achieving a state-of-the-art Dice score of 0.974 for kidney and 0.851 for renal tumours on their test group based on 3D U-Net architecture[12]. The already annotated data in KiTS19 provides a good foundation for developing deep-learning segmentation models of renal tumours and kidneys. Additionally, a convolutional neural network based on U-net architecture can segment visceral adipose tissue (VAT) accurately with a Dice score of 0.97[13]. However, there are currently no relevant studies reporting on the extraction of PRAT based on high-precision segmentation of kidney and VAT, although it is theoretically feasible.

An increasing number of studies highlight the role of VAT and PRAT in predicting the pathological features and prognosis of ccRCC[14–16]. Nonetheless, there is currently no automatic segmentation algorithm for PRAT, and the potential contributions of PRAT in WHO/ISUP ccRCC grading remain unexplored.

In this study, we proposed a comprehensive approach integrating TranUNet-assisted segmentation, multi-filter feature extraction, and ensemble learning techniques for ccRCC grading. The primary aim is to investigate the feasibility and reproducibility of this methodology for automated WHO/ISUP grading of ccRCC while concurrently exploring the influence of perirenal fat on the ccRCC grading process. To culminate, the SHapley Additive exPlanations (SHAP) method is employed for interpreting and visualizing the WHO/ISUP grading model.

HIGHLIGHTS

- The combination of radiomics features from the renal tumour, kidney, and perirenal adipose tissue (PRAT) demonstrated an improved ability to predict ccRCC grade compared to relying solely on tumour features [area under curve (AUC): 0.814 vs. 0.760] in a multicenter study of 614 patients.
- We employ TransUNet-based segmentation models for the segmentation of the kidneys, tumours, and visceral fat, resulting in a groundbreaking and remarkably consistent acquisition of PRAT.

Materials and methods

Study patients

This study included two retrospective datasets from our local hospital and one open-source dataset, namely the KiTS19 dataset[12,17]. The study received approval from the institutional review board of our institution, and the requirement for informed consent was waived. The training and internal validation dataset consisted of consecutive patients with ccRCC who underwent surgery in our hospital from January 2019 to June 2022. Nephrographic-phase CT images were downloaded from the Picture Archiving and Communication System. The detailed patient recruitment process is shown in Fig. 1 and Appendix E1 (online), Supplemental Digital Content 1, http://links.lww.com/JS9/C212. There were 471 patients included in the institutional dataset. Twenty percent of the institutional dataset ($n = 88$) was set aside to form the internal validation set. The external validation set, adhering to identical criteria, excluded 67 non-ccRCC patients and finally included 143 patients from the KiTS19

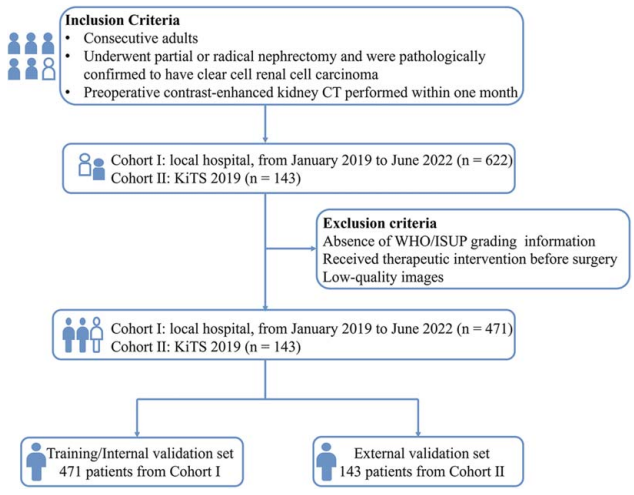

Figure 1. Flowchart of the study. CT, computed tomography; ISUP, International Society of Urological Pathology.

dataset. Finally, a total of 614 patients with pathologically confirmed ccRCC were included.

Pathological data were obtained from the postoperative pathological report in the hospital information system. Pathological diagnosis and grading were conducted by experienced senior pathologists in accordance with the 2016 WHO classification criteria[18]. WHO/ISUP grades I and II were assigned to the low-grade group, and WHO/ISUP grades III and IV to the high-grade group.

CT examination

In the local institutional dataset, patients underwent contrast-enhanced kidney CT within one month before surgery. All CT examinations were performed using one of the following scanners: Aquilion ONE, Toshiba Medical Systems; Discovery 750 HD, GE Healthcare; and Somatom Definition AS+, Siemens Healthineers. The specific scanning protocol is described in Appendix E2 (online), Supplemental Digital Content 1, http://links.lww.com/JS9/C212. The nephrographic phase was selected for subsequent analysis. Images of the KiTS19 dataset are from the late-arterial phase of preoperative contrast-enhanced abdominal CT.

Image preprocessing and segmentation

We employed a train-while-annotation strategy[10] to complete the segmentation of the three types of volume of interest (VOIs) (tumour, kidney, and visceral adipose tissue) on CT images from all included patients. Details of the train-while-annotation strategy regarding the annotation of the three types of VOIs and automatic segmentation protocols can be found in Appendix E3 (online), Supplemental Digital Content 1, http://links.lww.com/JS9/C212. The segmentation model is built based on annotated data and the TransUNet architecture. The code of the architecture is publicly available (https://github.com/Beckschen/TransUNet).

The final segmentation model was constructed using the training set data, and it was applied to predict VOIs for both the internal and external validation set. During the annotation process, cysts were included as part of the kidney. By establishing the coordinates of the median sagittal plane and employing the largest connected domain to determine whether tumour labels were present on the left and right sides of the plane, the tumour and the affected kidney are further extracted, and the labels of healthy kidney are removed. Based on VAT segmentation, we used the expansion-erosion algorithm to extract a range of 6 voxels around the kidney (defined as PRAT) and eliminate the VOI outside this range to define our target perirenal fat area. The workflow of our study is illustrated in Fig. 2.

Evaluation of the performance of the automatic segmentation model

We developed an automatic segmentation model using the training set data and assessed its performance on the internal

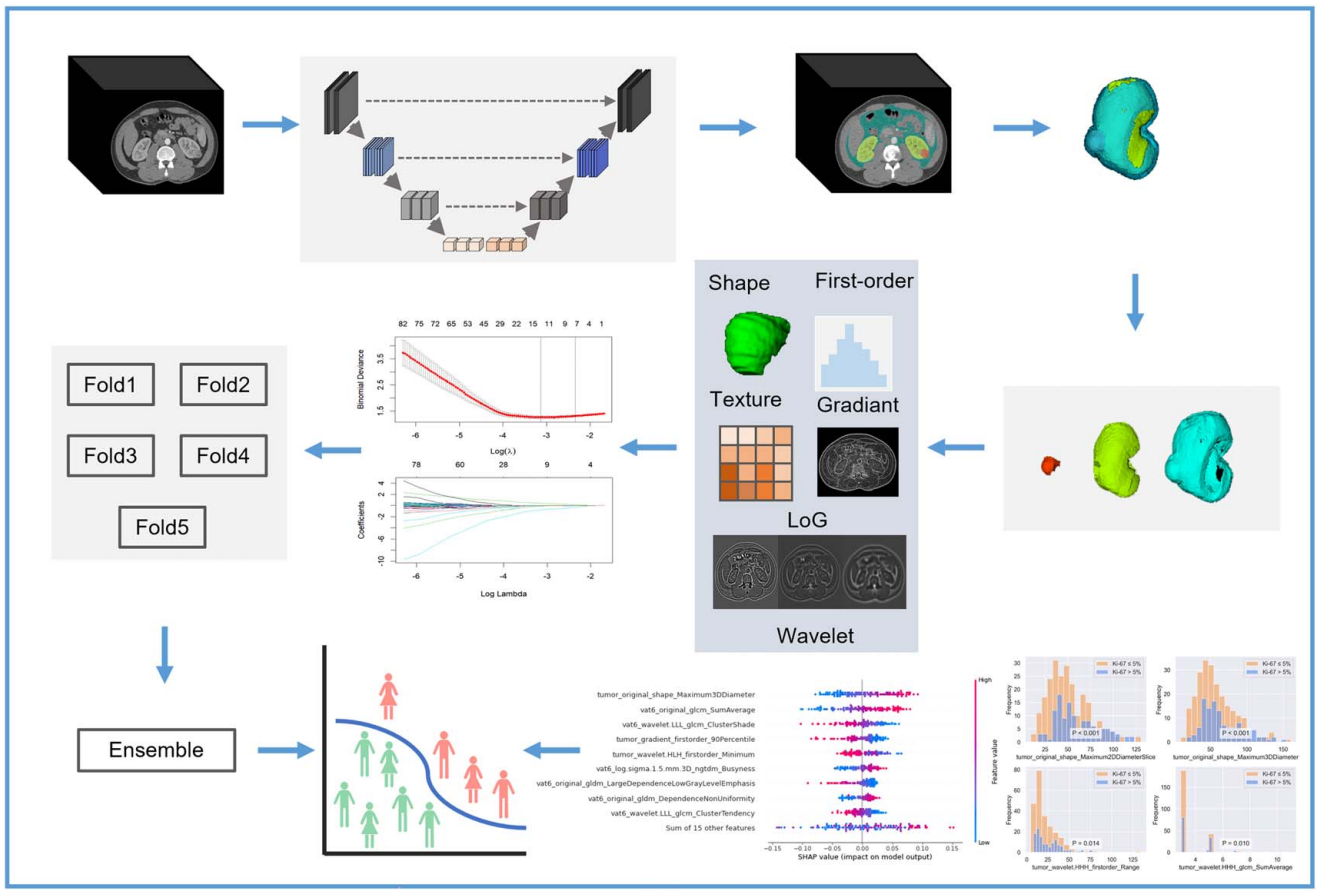

Figure 2. Workflow of the study.

validation set. The workflow of the automatic segmentation of the tumour, kidney, and PRAT were presented in Figure E1, Supplemental Digital Content 2, http://links.lww.com/JS9/C213 and detailed in Appendix E4 (online), Supplemental Digital Content 1, http://links.lww.com/JS9/C212. To evaluate the model's performance, we computed the Dice similarity coefficients (Dice), the Hausdorff distance (95%) (HD95), and volumetric similarity on the internal validation set.

The Dice quantifies the similarity between two segmentation results, and the formulas is as follows:

$$Dice = \frac{2TP}{2TP + FP + FN}$$

Where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positive, and FN is the number of false negatives.

The HD is used to evaluate boundary or volume differences between segmentation results, and the formulas is as follows:

$$HD = \max\left(\max_{a\in S(P)}\min_{b\in S(G)}\|a-b\|, \max_{b\in S(P)}\min_{a\in S(G)}\|b-a\|\right)$$

To mitigate the impact of outliers on the distance calculations, , the 95th percentile of the HD is used. HD95 involves sorting all the distances between pairs of points from smallest to largest and taking the value at the 95th percentile position. This means that it excludes the largest 5% of distance values, thereby providing a robust measure that is less sensitive to outliers.

Volume similarity is used to measure the similarity of two segmented volumes. The calculation is based on the difference between the two absolute volume values divided by the sum of the two compared volumes. The formula is defined as follows:

$$Volume \quad similarity = 1 - \frac{|FN - FP|}{|2TP + FP + FN|}$$

Radiomics workflow

Our radiomics workflow included (a) feature extraction, (b) feature selection, and (c) construction of the WHO/ISUP grading model. The radiomics approach is outlined in Appendix E5 (online), Supplemental Digital Content 1, http://links.lww.com/JS9/C212. Variance analysis, *t*-test, and LASSO regression were used for feature selection. Finally, we developed three different radiomics models for comparison: a kidney model, a tumour model, and a combination model (including all 1223×3 features from the kidney, tumour, and PRAT).

We randomly divided the local hospital images into a training set and a validation set with a ratio of 8:2. To address the problem of imbalanced classification in the training dataset, we adopted a random down-sampling scheme and ensemble techniques[19]. Determining the down-sampling fold based on the multiple of the group with a large sample size relative to the group with a small sample size. Established radiomics models were ensembled from all folds and tested on the same internal validation and the KiTS19 datasets set.

Statistical analysis

Dice, HD95, and volumetric similarity were used to characterize the performance of the segmentation model. All data analysis was performed using R Language (version 4.1.3) and Python (version 3.9.12). Numeric data were presented as mean ± standard deviation. Continuous variables were compared using the student *t*-test, Mann–Whitney U test, or Kruskal–Wallis H test, and categorical variables were compared using the χ^2 test. The differentiation performance of established models was evaluated by plotting a receiver operating characteristic curve, and the area under the receiver operating characteristic curve (AUC) analysis was calculated. DeLong's test on receiver operating characteristic (ROC) was performed to compare performance among different grading models. This study has been reported in line with the STROCSS criteria[20], Supplemental Digital Content 3, http://links.lww.com/JS9/C214.

Results

Patient characteristics

Among the 471 patients at our local hospital, 383 patients [mean age, 55.9 years ± 10.8 (SD)] were randomly assigned to the training set, and 88 [mean age, 53.5 years ± 11.0 (SD)] patients were assigned to the internal validation set. The external validation set, sourced from the KiST19 dataset, comprised a total of 143 patients [mean age, 59.2 years ± 12.5 (SD)]. The clinicopathologic characteristics of all patients are detailed in Table 1. In the local hospital dataset, 387 patients had Ki-67 index data reported in their pathological results.

Table 1

Baseline characteristics of patients with clear cell renal cell carcinoma.

Characteristics	Training set (*n* = 383)	Internal validation set (*n* = 88)	External validation set (*n* = 143)
Age (year)	55.9 ± 10.8	53.5 ± 11.0	59.2 ± 12.5
Sex			
Male	270	67	93
Female	113	21	50
Side			
Right	167	38	76
Left	216	50	67
Size (cm)	4.7 ± 2.3	4.4 ± 1.9	4.7 ± 2.9
BMI (Kg/mm^2)	24.4 ± 3.3	24.4 ± 3.0	31.2 ± 6.3
T category			
T1	305	74	101
T2	24	1	6
T3	51	13	35
T4	3	0	1
WHO/ISUP grade			
I	95	21	18
II	221	52	79
III	52	9	36
IV	15	6	10

ISUP, International Society of Urological Pathology.

Automated segmentation accuracy

The segmentation performance of $model_{2.5D}$ and $model_{2D}$ was compared, and both demonstrated satisfactory segmentation accuracy. In the $model_{2.5D}$, the Dice, HD95, and volumetric similarity of the kidney are 0.744 [interquartile range (IQR), 0.697–0.944], 31.705 (IQR, 3.742–12.900), and 0.863 (IQR, 0.788–0.978), respectively. These values were significantly higher than those achieved by $model_{2D}$ ($P = 0.011$, $P = 0.022$, $P = 0.001$, respectively). Regarding tumour segmentation, the Dice for $model_{2.5D}$ (0.836) was slightly higher than that for $model_{2D}$ (0.818), although not significantly. The performance of kidney, tumour, and VAT segmentation is shown in Appendixes Table E1 (online), Supplemental Digital Content 1, http://links.lww.com/JS9/C212 and Figure E2 (online), Supplemental Digital Content 4, http://links.lww.com/JS9/C215.

Development and testing of radiomics-based machine learning model

After performing variance analysis and student's *t*-test on the training data, 526 features for tumour, 233 features for kidney, and 254 features for PRAT were retained as candidates for LASSO analysis. We established three models: one based on tumour radiomics, one based on PRAT radiomics, and one combining radiomics features from tumour, kidney, and PRAT. Subsequently, five SVM models were built using the 5-fold data splitting from the training set and applied to both the

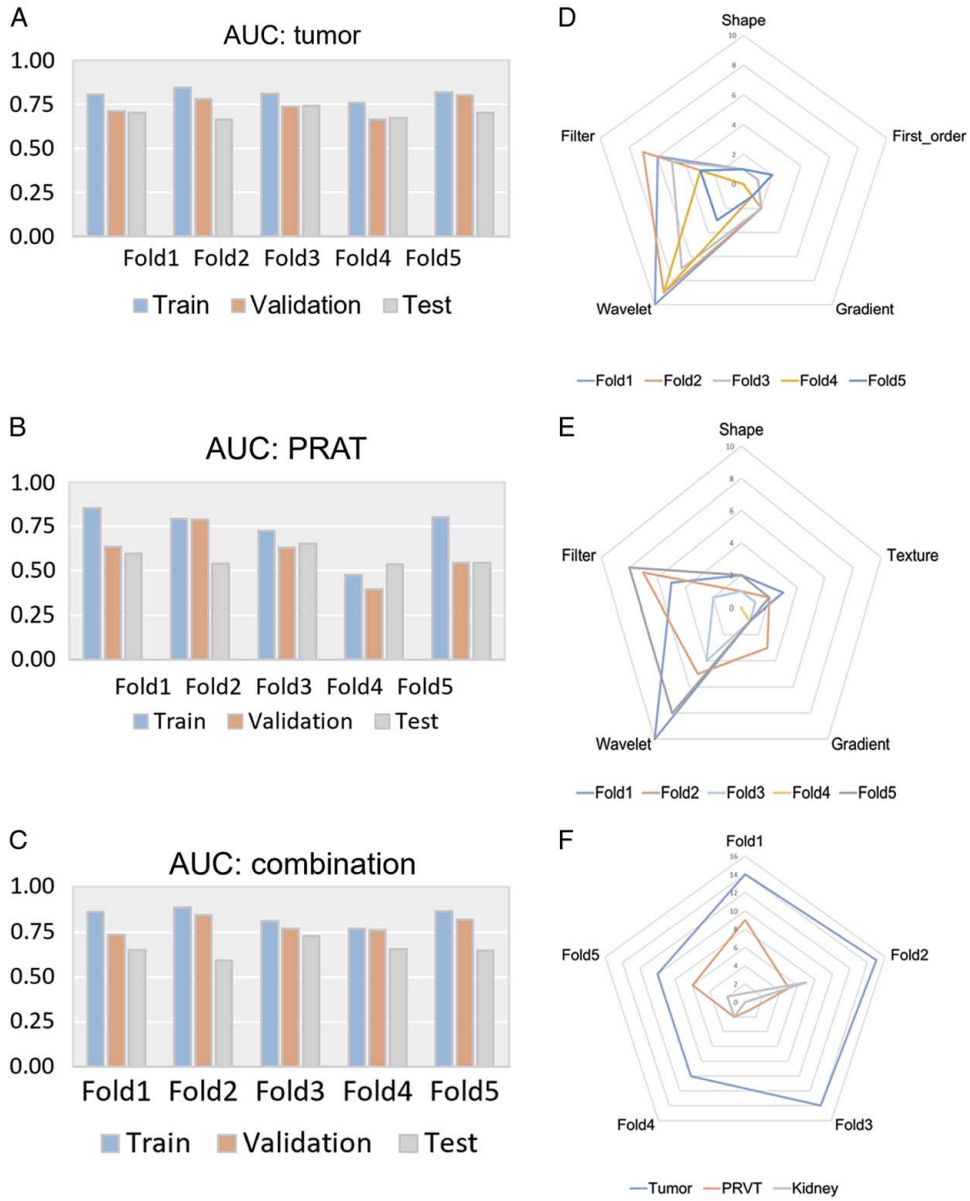

Figure 3. The analysis and results of multiple filtering feature extraction and performance of prediction models based on 5 down-sampled folds. (A–C) The AUC value for predicting the WHO/International Society of Urological Pathology (ISUP) grade of clear cell renal cell carcinoma based on tumour, perirenal adipose tissue, and the combination models. (D–F) Number of different types and filters of radiomics features ultimately incorporated into different radiomics models. AUC, area under curve; PRAT, perirenal adipose tissue.

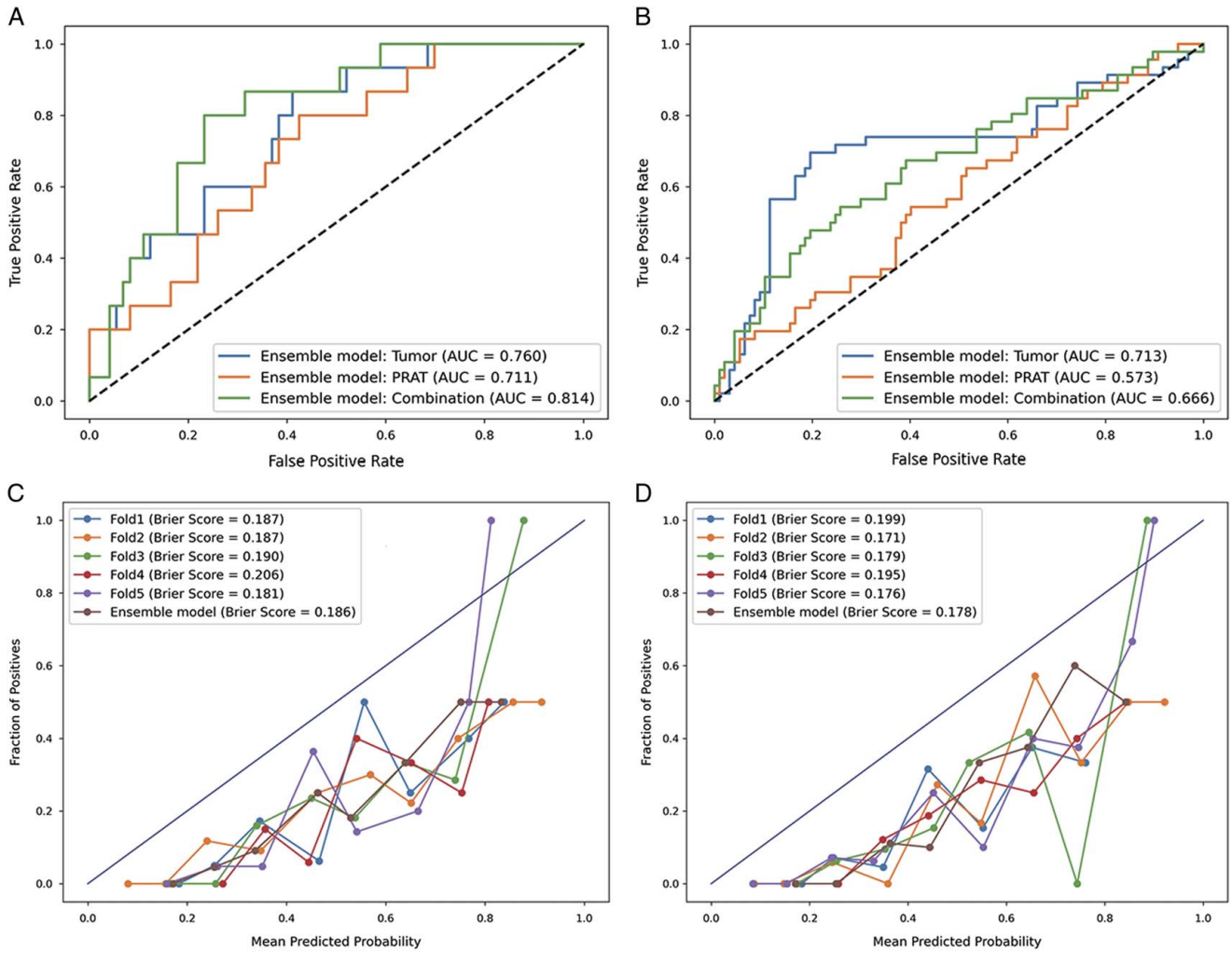

Figure 4. The receiver operating characteristic curves of different ensemble model on internal (A) and external (B) validation set. Calibration curve of the machine learning model based on tumour (C) and combined (D) radiomic features in the internal validation set. The Brier score loss of each model is shown in brackets. AUC, area under curve; PRAT, perirenal adipose tissue.

internal and external validation set for validation. Figure 3 illustrates the AUCs of the models based on each fold. Finally, we created ensemble learning models for tumours, PRAT, and combination by combining the results of the 5 SVM models. Calibration curves for the models are presented in Fig. 4. The Brier score loss was 0.186 for the ensemble tumour model and 0.178 for the ensemble combination model, indicating that the ensemble combination model slightly outperformed the ensemble tumour model.

WHO/ISUP grade prediction accuracy

The ensemble combination model showed the highest discrimination between the low and high WHO/ISUP grade, with an AUC of 0.814 (95% CI: 0.717–0.889) in the internal validation set. This was significantly higher than the AUC for the ensemble tumour model (AUC, 0.760; 95% CI, 0.657–0.845). For the internal validation set, the ensemble combination model achieved a sensitivity, specificity, and accuracy of 80.0%, 75.3%, and 76.1%, respectively, while for the external validation set, these values were 60.9%, 63.9%, and 63.9%, respectively. ROC curves for the pathological grade are shown in Fig. 4, and the performance of the prediction models is summarized in Table 2. DeLong's test results comparing AUCs are displayed in Table 3.

Table 2

Performance of the radiomics models for predicting the pathological grade.

Ensemble model	AUC (95% CI)	Sensitivity	Specificity	Accuracy
Internal validation set				
Tumour	0.760 (0.657–0.845)	0.533	0.767	0.727
PRAT	0.711 (0.604–0.802)	0.400	0.781	0.716
Combination	0.814 (0.717–0.889)	0.800	0.753	0.761
External test set				
Tumour	0.713 (0.631–0.785)	0.630	0.825	0.762
PRAT	0.573 (0.488–0.656)	0.674	0.423	0.503
Combination	0.666 (0.583–0.743)	0.609	0.639	0.639

Combination, the combination of radiomics features of tumour, kidney, and PRAT.
AUC, area under curve; PRAT, perirenal adipose tissue.

Table 3

DeLong's test *P* values for pairwise AUC comparisons, with the combination's AUC as a reference.

Ensemble model	AUC	*P*
Internal validation set		
Combination-tumour	0.814–0.760	0.034*
Combination-PRAT	0.814–0.711	0.050
External test set		
Combination-tumour	0.666–0.713	0.068
Combination-PRAT	0.666–0.573	0.014*

Combination, the combination of radiomics features of tumour, kidney, and PRAT.
AUC, area under curve; PRAT, perirenal adipose tissue.

Exploration of model interpretability

Figure 5 illustrates the SHapley Additive exPlanations (SHAP) values for all selected features in the radiomics-based machine learning models. The top contributing features to the combination model of the five folds were the maximum diameter exhibited by the original tumour shape in the 2D slice (tumor_original_shape_Maximum2DDiameterSlice), the maximum diameter manifested by the original tumour shape in the 3D representation (tumor_original_shape_Maximum3DDiameter), the range of values pertaining to the high-frequency sub-band (HHH) in the wavelet analysis (tumor_wavelet.HHH_firstorder_Range), and the sum average derived from the gray-level co-occurrence matrix within the HHH sub-band (tumor_wavelet.HHH_glcm_SumAverage). Furthermore, all these features exhibited significant differences between the group with Ki-67 less than or equal to 5% and group with Ki-67 greater than 5% ($P < 0.001$, $P < 0.001$, $P = 0.014$, and $P = 0.010$, respectively), as shown in Fig. 6.

Discussion

In this study, we proposed a comprehensive and fully automatic pipeline for WHO/ISUP grade prediction based on CT scans. This pipeline leverages the TranUNet automated segmentation methods in conjunction with ensemble radiomics models. The segmentation model exhibited excellent performance in kidney, tumour, and VAT. The combination radiomics model proposed

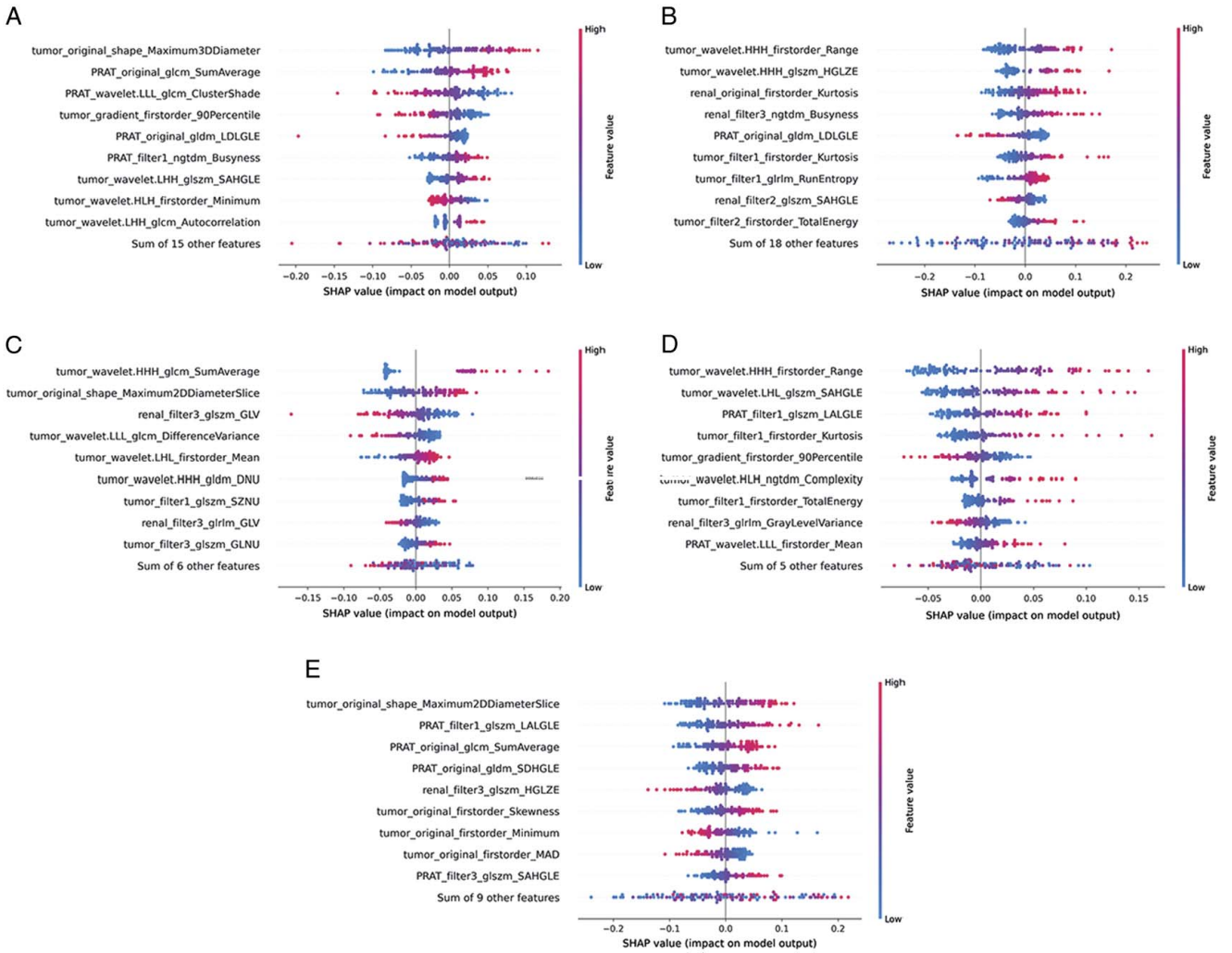

Figure 5. SHAP values of each feature in the combination model based on five-fold training data (A–E). The different colours (red and blue) represent different levels of effect on the output of the model. (filter 1, log.sigma.1.5.mm.3D; filter 2, log.sigma.3.0.mm.3D; filter 3, log.sigma.5.0.mm.3D).

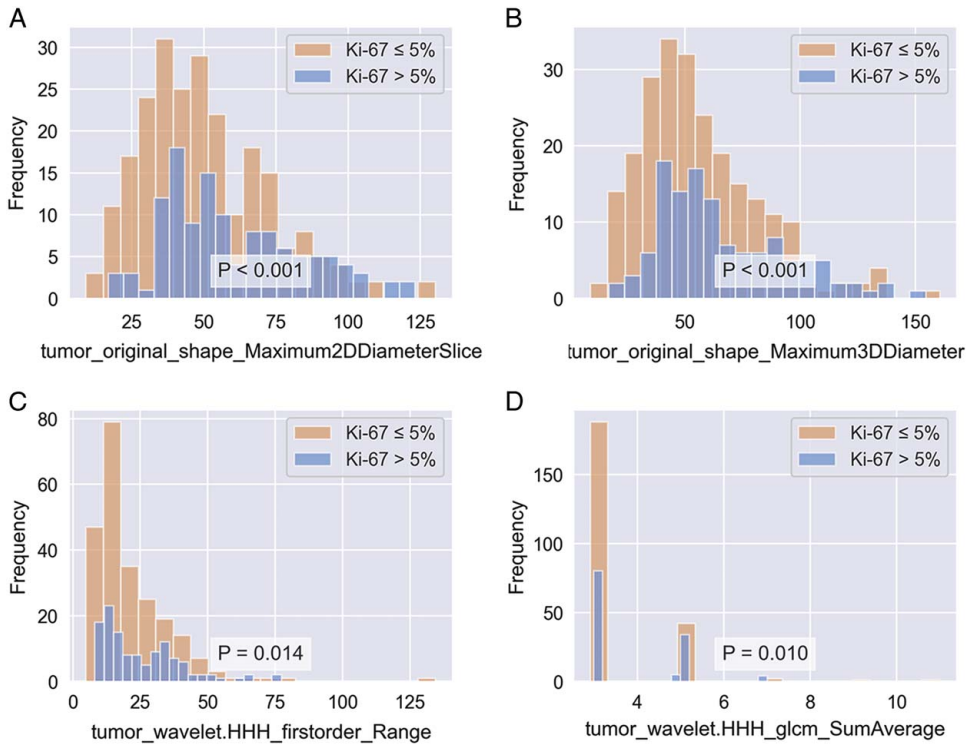

Figure 6. Distribution of radiomic features of patients grouped into different Ki-67 index groups. Tumor_original_shape_Maximum2DDiameterSlice (A), tumor_original_shape_Maximum3DDiameter (B), tumor_wavelet.HHH_firstorder_Range (C), and tumor_wavelet.HHH_glcm_SumAverage (D) were all significantly different between different Ki-67 index groups.

in this study, utilizing PRAT features, demonstrated significantly superior predictive performance compared to the traditional radiomics model that focused solely on the tumour.

Our study uses and quantifies segmentation performance, which is rare but important in clinical studies. TransUNet, with strong innate self-attention mechanisms, was the main framework that we used for completing accurate annotation[21]. However, due to insufficient low-level details, the positioning capabilities of TransUNet may be limited. By implementing a two-pronged approach that involves the incremental expansion of annotated data through a train-while-annotation strategy and the application of a post-processing technique centred on the maximum connected domain, we successfully address the issues related to the imprecise positioning of TranUNet. As demonstrated by our results, $model_{2D}$ have achieved similar segmentation performance as $model_{2.5D}$ models that have integrated three-plane positioning information. The segmentation method, which merges predictive data from the automated segmentation model with manual corrections, substantially enhances the accuracy and repeatability of the final segmentation results that use for radiomics feature extraction.

Our study innovatively confirmed the contribution of radiomics features from PRAT to the pathological grading of ccRCC. VAT is considered to be an endocrine organ with abundant blood vessels, innervation, and inflammatory cells that secrete a variety of hormones and cytokines related to carcinogenesis and tumour progression and up-regulate genes involved in tumour cell invasion and metastasis[22–24]. Perinephric fat is VAT adjacent to the kidney, but it originates from adipocyte precursors and mature adipocytes and is active in metabolism and adipokine secretion, which may play a special role in kidney disease[25,26]. According to the American Joint Committee on Cancer (AJCC) staging manual (8th edition), perirenal fat invasion is an important basis for pT3a category, indicating a higher degree of malignancy and a worse prognosis. A related study showed that ccRCC induces the browning of adjacent perirenal fat, which promotes tumour growth, invasion and metastasis by secreting lactic acid[27]. Du *et al.*[28] showed that perirenal fat percentage was an independent predictor of postoperative mortality risk in ccRCC. However, current studies have focused on the thickness and area of perirenal fat, without exploring more advanced radiomics features. Due to the difficulty in identifying Gerota fascia on CT images, this study initially attempted to extract a reproducible method with a 6-pixel perirenal range and to study its role in the grading of ccRCC.

Radiomics analysis is a powerful tool for extracting a multitude of quantitative features from medical images. The features selected by LASSO regression are mainly focus on gradient, wavelet, and LoG filter features, which shows that multiple image filter feature extraction strategy enables us to capture subtle, yet pathological relevant, information from the imaging data. Based on the beeswarm diagram of the SHAP value of the five-fold data in the training set, it can be seen that the first-order feature (maximum diameter) of the tumour and the filter-based feature (first-order range) contribute the most to the pathological grading of ccRCC. This is consistent with the fact that diameter is an important predictor of tumour pathological grade; on the other hand, ccRCC with higher pathological grade are more prone to necrosis[29,30], resulting in a larger range of first-order features, which may be more prominent in the filtered image. Additionally,

we conducted further analysis to explore the performance of these features in patients with varying Ki-67 indices. Our findings revealed statistically significant differences among different Ki-67 groups, providing further validation of the viability of these imaging signatures in predicting tumour biological behaviour and demonstrating their robustness. This additional investigation strengthens the overall confidence in the effectiveness of these features as reliable indicators of tumour behaviour.

For the tumour prediction grading model, the prediction performance of the external validation set is slightly lower than the internal validation set, which may be related to the scanning scheme of CT in different data sets. The data of our centre are selected from the nephrographic phase, while the data of KiTS19 are from the late-arterial phase. On the other hand, this also reflects the good generalization performance of the model. In the prediction grading model of PRAT, the prediction performance of the external validation set was significantly lower than the internal validation set, which may be related to different races. The patients in the KiTS19 data set had a significantly higher BMI. Therefore, the predictive performance of the combination model is lower, too.

In addition to shedding light on the significance of perirenal fat segmentation for understanding tumour invasiveness, our study also paves the way for exploring the value of perirenal fat within different ranges. The complexity of perirenal fat's role in kidney disease and its potential impact on tumour biology warrant further investigation. Moreover, our research sets a new paradigm for studying the influence of peritumoral fat on the behaviour of other cancers, such as pancreatic and colorectal cancer. This work not only contributes to our understanding of renal cell carcinoma but also opens doors for more comprehensive research into the role of peritumoral fat in various cancer types, ultimately advancing our knowledge of tumour biology and clinical decision-making.

Our study acknowledges certain limitations. First, our data primarily consisted of data from our local hospital and an open-source dataset, which may introduce selection bias. Second, multicenter data with consistent scanning protocols are lacking, and external validation using different datasets is crucial to verify the robustness and applicability of our proposed method in real-life clinical settings. Third, there is sample imbalance, which can be unavoidable, but we effectively utilize all the data for model building through ensemble learning. Fourth, although an automatic segmentation model was used, manual annotation and correction by experienced radiologists are still required in the study. The experience of the radiologists may introduce bias to the results. Finally, this study initially defined a certain range of perinephric fat based on automatic segmentation but performed further exploration and comparison of the differences in adipose tissue characteristics surrounding the kidneys in different voxel ranges. Therefore, further large-sample multicenter studies are warranted to investigate the predictive performance of PRAT with different ranges.

In conclusion, our study proposes a novel method to predict the pathological grade of ccRCC using automated segmentation and radiomics analysis. The results showed that perirenal fat has incremental value for the prediction of ccRCC pathological grade. To advance our study, we recommend multicenter validation with a larger sample size to acquire high-level evidence for clinical applications.

Ethical approval

The study received approval from the institutional review board of our institution (ethics committee: Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology), and the requirement for informed consent was waived (TJ-IRB20230893).

Consent

The study received approval from the institutional review board of our institution, and the requirement for informed consent was waived.

Source of funding

This work is supported by the grants from National Natural Science Foundation of China (NSFC) 82371942, 82071889, 82202127, and 62131009.

Author contribution

Z.L., S.L., and Z.Z. conceived the study and designed major experiments. S.L., M.G., Z.L., K.H., and W.Q. collected the data and made the clinical diagnosis. Z.Z. and S.L. performed the automatic segmentation and machine learning. Z.Z. and S.L. performed analysis and interpretation. Z.Z. and S.L. wrote the manuscript. J.L., Q.C., Z.L., I.K., and Q.Z. reviewed and revised the manuscript.

Conflicts of interest disclosure

The authors declare that they have no competing interests.

Research registration unique identifying number (UIN

The study received approval from the institutional review board of our institution (TJIRB20230893). We have registered it on the https://www.clinicaltrials.gov/ and the Unique Identifying Number (UIN) is NCT06167863.

Guarantor

The corresponding author Dr Zhen Li is the Guarantor of this work.

Data statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Provenance and peer review

Not commissioned, externally peer-reviewed.

References

[1] Znaor A, Lortet-Tieulent J, Laversanne M, *et al*. International variations and trends in renal cell carcinoma incidence and mortality. Eur Urol 2015;67:519–30.

[2] Bektas CT, Kocak B, Yardimci AH, *et al*. Clear cell renal cell carcinoma: machine learning-based quantitative computed tomography texture analysis for prediction of Fuhrman nuclear grade. Eur Radiol 2019;29: 1153–63.
[3] McIntosh AG, Ristau BT, Ruth K, *et al*. Active surveillance for localized renal masses: tumor growth, delayed intervention rates, and > 5-yr clinical outcomes. Eur Urol 2018;74:157–64.
[4] Dagher J, Delahunt B, Rioux-Leclercq N, *et al*. Clear cell renal cell carcinoma: validation of World Health Organization/International Society of Urological Pathology grading. Histopathology 2017;71:918–25.
[5] Delahunt B, Eble JN, Egevad L, *et al*. Grading of renal cell carcinoma. Histopathology 2019;74:4–17.
[6] Stewart GD, Klatte T, Cosmai L, *et al*. The multispeciality approach to the management of localised kidney cancer. Lancet 2022;400:523–34.
[7] Yang H, Wu K, Liu H, *et al*. An automated surgical decision-making framework for partial or radical nephrectomy based on 3D-CT multi-level anatomical features in renal cell carcinoma. Eur Radiol 2023;33:7532–41.
[8] Niu J, Tan Q, Zou X, *et al*. Accurate prediction of glioma grades from radiomics using a multi-filter and multi-objective-based method. Math Biosci Eng 2023;20:2890–907.
[9] Liu BX, Sun W, Kong XQ. Perirenal fat: a unique fat pad and potential target for cardiovascular disease. Angiology 2019;70:584–93.
[10] Wang S, Li C, Wang R, *et al*. Annotation-efficient deep learning for automatic medical image segmentation. Nat Commun 2021;12:5915.
[11] Vorontsov E, Molchanov P, Gazda M, *et al*. Towards annotation-efficient segmentation via image-to-image translation. Med Image Anal 2022;82:102624.
[12] Heller N, Isensee F, Maier-Hein KH, *et al*. The state of the art in kidney and kidney tumor segmentation in contrast-enhanced CT imaging: Results of the KiTS19 challenge. Med Image Anal 2021;67:101821.
[13] Weston AD, Korfiatis P, Kline TL, *et al*. Automated abdominal segmentation of CT scans for body composition analysis using deep learning. Radiology 2019;290:669–79.
[14] Nguyen GK, Mellnick VM, Yim AK-Y, *et al*. Synergy of sex differences in visceral fat measured with CT and tumor metabolism helps predict overall survival in patients with renal cell carcinoma. Radiology 2018; 287:884–92.
[15] Du Y, Yang W, Liu H, *et al*. Perirenal fat as a new independent prognostic factor in patients with surgically treated clear cell renal cell carcinoma. Clin Genitourin Cancer 2022;20:e75–80.
[16] Li S, Liao Z, He K, *et al*. Association of sex-specific abdominal adipose tissue with WHO/ISUP grade in clear cell renal cell carcinoma. Insights Imaging 2023;14:194.
[17] Heller N. KiTS19 Challenge Data: 300 kidney tumor cases with clinical context, ct semantic segmentations, and surgical outcomes. arXiv preprint 2019;arXiv:190400445.
[18] Moch H, Cubilla AL, Humphrey PA, *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 2016;70:93–105.
[19] Salunkhe UR, Mali SN. Classifier Ensemble Design for Imbalanced Data Classification: A Hybrid Approach. Procedia Computer Science 2016;85: 725–32
[20] Mathew G, Agha R, Albrecht J, *et al*. STROCSS Group. STROCSS 2021: Strengthening the reporting of cohort, cross-sectional and case-control studies in surgery. Int J Surg 2021;96:106165.
[21] J. Chen YL, Yu Q, Luo X, *et al*. TransUNet: transformers make strong encoders for medical image segmentation. ArXiv 2021;2102: 04306.
[22] Chau YY, Bandiera R, Serrels A, *et al*. Visceral and subcutaneous fat have different origins and evidence supports a mesothelial source. Nat Cell Biol 2014;16:367–75.
[23] Sanchez-Gurmaches J, Guertin DA. Adipocytes arise from multiple lineages that are heterogeneously and dynamically distributed. Nat Commun 2014;5:4099.
[24] Doyle SL, Donohoe CL, Lysaght J, *et al*. Visceral obesity, metabolic syndrome, insulin resistance and cancer. Proc Nutr Soc 2012;71:181–9.
[25] Ibrahim MM. Subcutaneous and visceral adipose tissue: structural and functional differences. Obes Rev 2010;11:11–8.
[26] Chen X, Mao Y, Hu J, *et al*. Perirenal fat thickness is significantly associated with the risk for development of chronic kidney disease in patients with diabetes. Diabetes 2021;70:2322–32.
[27] Wei G, Sun H, Dong K, *et al*. The thermogenic activity of adjacent adipocytes fuels the progression of ccRCC and compromises anti-tumor therapeutic efficacy. Cell Metab 2021;33:2021–39.e2028.
[28] Du Y, Yang W, Liu H, *et al*. Perirenal fat as a new independent prognostic factor in patients with surgically treated clear cell renal cell carcinoma. Clin Genitourin Cancer 2022;20:e75–80.
[29] Khor LY, Dhakal HP, Jia X, *et al*. Tumor necrosis adds prognostically significant information to grade in clear cell renal cell carcinoma: a study of 842 consecutive cases from a single institution. Am J Surg Pathol 2016; 40:1224–31.
[30] Cai Q, Christie A, Rajaram S, *et al*. Ontological analyses reveal clinically-significant clear cell renal cell carcinoma subtypes with convergent evolutionary trajectories into an aggressive type. EBioMedicine 2020;51: 102526.