THERAPEUTIC ADVANCES in
Urology

Review

Artificial intelligence and radiomics in evaluation of kidney lesions: a comprehensive literature review

Matteo Ferro, Felice Crocetto, Biagio Barone, Francesco del Giudice, Martina Maggi, Giuseppe Lucarelli, Gian Maria Busetto, Riccardo Autorino, Michele Marchioni, Francesco Cantiello, Fabio Crocerossa, Stefano Luzzago, Mattia Piccinelli, Francesco Alessandro Mistretta, Marco Tozzi, Luigi Schips, Ugo Giovanni Falagario, Alessandro Veccia, Mihai Dorin Vartolomei, Gennaro Musi, Ottavio de Cobelli, Emanuele Montanari and Octavian Sabin Tătaru

***Abstract*:** Radiomics and artificial intelligence (AI) may increase the differentiation of benign from malignant kidney lesions, differentiation of angiomyolipoma (AML) from renal cell carcinoma (RCC), differentiation of oncocytoma from RCC, differentiation of different subtypes of RCC, to predict Fuhrman grade, to predict gene mutation through molecular biomarkers and to predict treatment response in metastatic RCC undergoing immunotherapy. Neural networks analyze imaging data. Statistical, geometrical, textural features derived are giving quantitative data of contour, internal heterogeneity and gray zone features of lesions. A comprehensive literature review was performed, until July 2022. Studies investigating the diagnostic value of radiomics in differentiation of renal lesions, grade prediction, gene alterations, molecular biomarkers and ongoing clinical trials have been analyzed. The application of AI and radiomics could lead to improved sensitivity, specificity, accuracy in detecting and differentiating between renal lesions. Standardization of scanner protocols will improve preoperative differentiation between benign, low-risk cancers and clinically significant renal cancers and holds the premises to enhance the diagnostic ability of imaging tools to characterize renal lesions.

Keywords: artificial intelligence, machine learning, radiomics, renal cancer, imaging

Received: 11 October 2022; revised manuscript accepted: 4 March 2023.

Ther Adv Urol

2023, Vol. 15: 1–26

DOI: 10.1177/17562872231164803

Correspondence to:
Matteo Ferro
Department of Urology, IEO – European Institute of Oncology, IRCCS – Istituto di Ricovero e Cura a Carattere Scientifico, via Ripamonti 435, Milan 20141, Italy

Università degli Studi di Milano, Milan, Italy
matteo.ferro@ieo.it

Felice Crocetto
Biagio Barone
Department of Neurosciences and Reproductive Sciences and Odontostomatology, University of Naples Federico II, Naples, Italy

Francesco del Giudice
Martina Maggi
Department of Maternal Infant and Urologic Sciences, Policlinico Umberto I Hospital, Sapienza University of Rome, Rome, Italy

Giuseppe Lucarelli
Urology, Andrology and Kidney Transplantation Unit, Department of Emergency and Organ Transplantation, University of Bari, Bari, Italy

Gian Maria Busetto
Ugo Giovanni Falagario
Department of Urology and Organ Transplantation, University of Foggia, Foggia, Italy

Riccardo Autorino
Division of Urology, VCU Health, Richmond, VA, USA

Michele Marchioni
Department of Medical, Oral and Biotechnological Sciences, Urology Unit, SS Annunziata Hospital, G. d'Annunzio University of Chieti, Chieti, Italy

Department of Urology, ASL Abruzzo 2, Chieti, Italy

Francesco Cantiello
Fabio Crocerossa
Department of Urology, Magna Graecia University of Catanzaro, Catanzaro, Italy

Introduction

Renal cell carcinoma (RCC) is placed sixth in the diagnosed type of cancer in men and is at the 10th place in women.[1,2] In recent decades, the growing availability of non-invasive advanced radiological techniques to investigate non-specific abdominal and musculoskeletal pain has led to a steady increase of incidental kidney lesions, which are more frequently small, asymptomatic, and clinically localized; indeed, up to 50% of all diagnosed renal lesions are considered small renal masses (SRMs; i.e. ⩽4 cm in diameter).[3–5] Moreover, up to 30% are benign at final histology (i.e. after radical or partial nephrectomy),[6–8] while a non-negligible proportion of these lesions have slow growth rate. In these cases, any treatment might result in an overtreatment.[9]

The current diagnostic work-up with standard imaging tools (ultrasound, computed tomography (CT), and magnetic imaging resonance (MRI)[10] is still hampered by a suboptimal ability to correctly distinguish RCC from all benign lesions in the pre-operative setting.[11–13] Tissue sampling by renal lesion biopsy has shown high diagnostic accuracy for RCC; however, it is an invasive procedure and difficult to perform in some lesion localization.[10,14] Renal lesion biopsy has a high non-diagnostic rate (approximately 15% and erroneous diagnoses (approximately

Stefano Luzzago
Marco Tozzi
Department of Urology, IEO – European Institute of Oncology, IRCCS – Istituto di Ricovero e Cura a Carattere Scientifico, Milan, Italy

Università degli Studi di Milano, Milan, Italy

Mattia Piccinelli
Cancer Prognostics and Health Outcomes Unit, Division of Urology, University of Montréal Health Center, Montréal, QC, Canada

Department of Urology, IEO – European Institute of Oncology, IRCCS – Istituto di Ricovero e Cura a Carattere Scientifico, Milan, Italy

Francesco Alessandro Mistretta
Gennaro Musi
Ottavio de Cobelli
Department of Urology, IEO – European Institute of Oncology, IRCCS – Istituto di Ricovero e Cura a Carattere Scientifico, Milan, Italy

Department of Oncology and Hematology-Oncology, Università degli Studi di Milano, Milan, Italy

Luigi Schips
Department of Medical, Oral and Biotechnological Sciences, Urology Unit, SS Annunziata Hospital, G. d'Annunzio University of Chieti, Chieti, Italy

Alessandro Veccia
Urology Unit, Azienda Ospedaliera Universitaria Integrata Verona, University of Verona, Verona, Italy

Mihai Dorin Vartolomei
Department of Cell and Molecular Biology, George Emil Palade University of Medicine, Pharmacy, Science and Technology of Târgu Mures, Târgu Mures, Romania

Department of Urology, Medical University of Vienna, Vienna, Austria

Emanuele Montanari
Department of Urology, Foundation IRCCS Ca' Granda – Ospedale Maggiore Policlinico, Department of Clinical Sciences and Community Health, University of Milan, Milan, Italy

Octavian Sabin Tăaru
Institution Organizing University Doctoral Studies (IOSUD), George Emil Palade University of Medicine, Pharmacy, Science and Technology of Târgu Mures, Târgu Mures, Romania

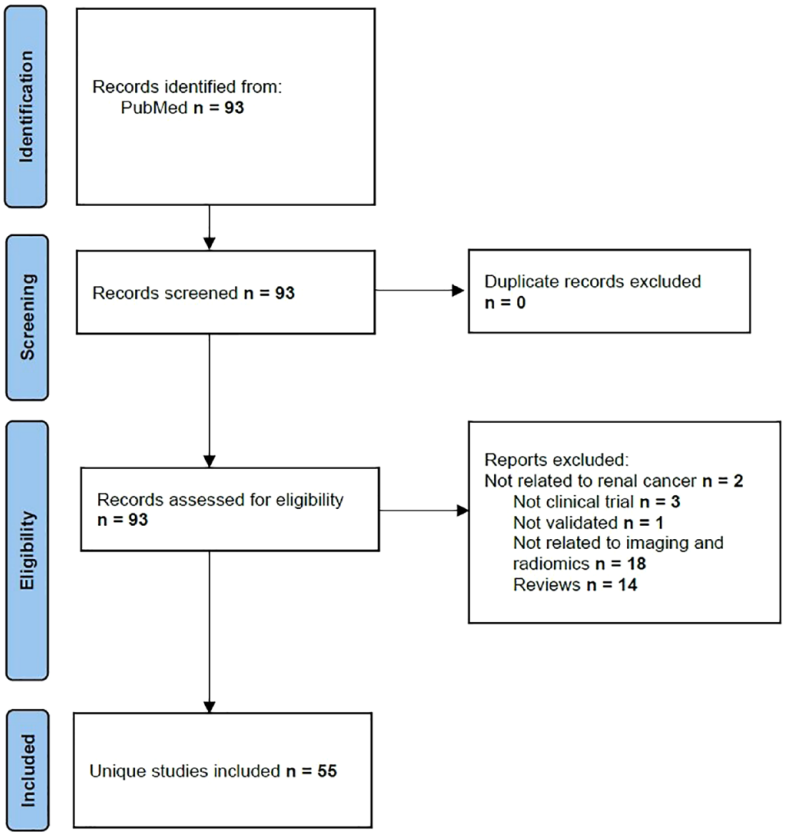

Figure 1. PRISMA methodology for identification of radiomics and renal lesions studies.

10%). This is somehow related to tumor heterogeneity.[15–17] Computer-aided diagnosis (CAD) using artificial intelligence (AI) and its subset machine learning (ML; and state-of-the-art ML approach deep learning (DL)) are a new area of interest in medical research.[18] Radiomics aims to help clinicians to improve the work-up and treatment of several oncological diseases, by mixing the qualitative features quantitative data obtained through imaging tools.[19–23] Radiomics combines and analyses the mix of these features and data from radiographic digital images,[23] develop descriptive and predictive models, combining image features and phenotypes with gene and protein signatures.[23,24] Specifically for kidney lesions, radiomics, by capturing subtle features that might escape human identification, has shown to enhance the diagnostic, prognostic, and predictive power of conventional radiological techniques.[6,10,14,25,26] Aim of this comprehensive literature review is to provide a holistic overview of the actual role of radiomics in kidney lesions assessment and to identify the future opportunities in preoperative diagnosis of renal lesions.

Materials and methods

PubMed–Medline database and clinicaltrials.gov were used to identify relevant, original studies from the last 7 years (until July 2022) on the topic. We had identified 93 research articles, of which 49 were published in the last 7 years. Earlier six studies has also been included in the analysis to reflect the data from leading studies. Data from seven ongoing clinical trials were also evaluated. In Figure 1, we have summarized the search methodology performed using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.[27]

Terms used for search were: kidney neoplasms, renal cell carcinoma, AI, ML, DL, and radiomics, evaluation, differentiation, characterization, and assessment. Inclusion criteria were: (1) studies applying radiomics for the study of renal masses, (2) articles written in English language, and (3) studies with a well-documented methodology to allow replication. Articles not related to kidney tumors and non-original articles (review publications, editorials, and replies to comments) were

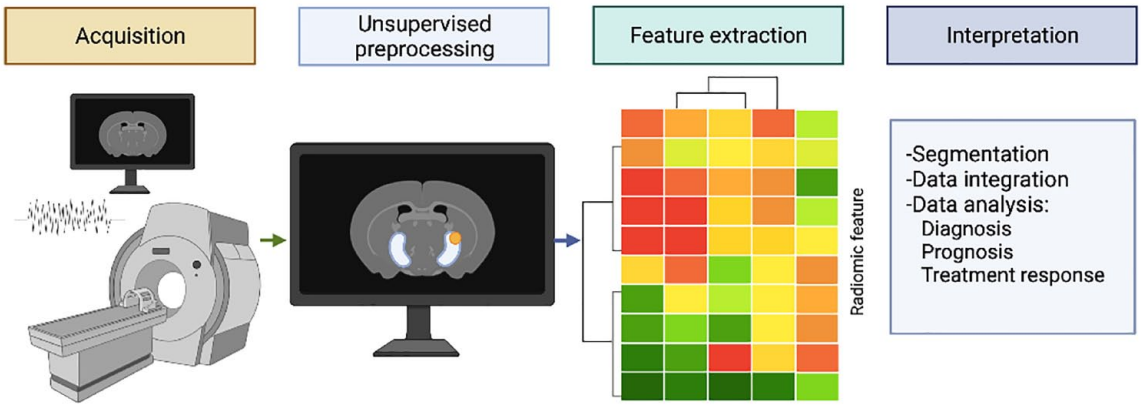

Figure 2. Radiomics flow analysis.

excluded from the analysis. Following the PRISMA guidelines, we have evaluated the identified studies. Published articles were analyzed if population of patients were with kidney tumors; intervention was the evaluation with CT, MRI, positron emission tomography (PET-CT) and contrast-enhanced ultrasound (CEUS)-based radiomics. The comparator was the radiologists' subjective image assessment. The aim was to analyze the role of radiomics and to assess the clinical aspects of radiomics in kidney tumors. Data were extracted after careful evaluation of the full text of the articles for this analysis. The data used in our review have been identified from each research: author, clinical outcomes and gold standard, the radiological modality, the prospective or retrospective design, the patients involved, radiomics method and results.

Results

The search of online databases yielded results on the topic on radiomics and kidney cancer that could be classified in studies that searched the differentiation of benign and cancerous tissue, of angiomyolipomas (AMLs) from RCC, of oncocytoma from RCC, between different subtypes of RCC. Also the prediction of Fuhrman grade, response to therapy, and prediction of gene mutations of molecular biomarkers have been identified. AI offers the unique opportunity to handle the huge volume of data being created by radiomics features extraction from kidney imaging and to combine these with clinical and pathological variables to provide even more accurate prediction of the outcome discussed in previous paragraphs. Figure 2 briefly depicts the process of radiomics research.

Differentiation of benign from cancerous kidney tumors

Radiomics in renal malignancy aimed to improve the accuracy in distinguishing malignant *versus* benign histology and in case of renal malignancy, of different subtypes, to provide the best and tailored management. The possibility to obtain an abundance of quantitative features such as histograms, textures, and shapes extracted *via* high-throughput data from CT and MRI has surely influenced the recent research toward this direction.[6,28]

One of the first studies to assess CT radiomic features and texture analysis in renal tumors was performed by Yu *et al.*[29] on a total of 119 patients. Utilizing histogram-based features of skewness and kurtosis, the authors reported an area under the curve (AUC) of 0.91 and 0.93, respectively, in differentiating renal cancer from oncocytoma (ONC) with AUC of 0.92 in differentiating ONC from other tumors.[14,29] Analogously, Coy *et al.*[30] reported, for 200 patients with 200 unique masses, an AUC of 0.85 of CAD in discriminating malignant *versus* benign lesions. Despite the limitations of only comparing clear cell (cc) RCC and ONC and using two different contrast agents, the study uncovered the high potential of radiomics and ML in analyzing several texture-analysis features to distinguish malignant from benign kidney lesions.[31] A further study by Erdim *et al.*[32] aimed to investigate the possibility to identify benign from cancerous masses through CT and ML texture analysis on a total of 79 patients with 84 renal masses (21 benign and 63 malignant). With a total number of 198 features for unenhanced CT and 244 for contrast-enhanced CT, ML based on random forest (RF) algorithm

radiomics, yielded an accuracy in distinguishing renal cancer from other benign renal masses 90.5% with an AUC of 0.915, which, eliminating collinear features, increased to 91.7% and 0.916, respectively.[32] Zhou *et al.*[33] reported, in 192 patients with renal cancer analyzed *via* the InceptionV3 model, a DL radiomics model, an AUC of 0.97 for region of interest (ROI) data set and 0.93 for rectangular box region (RBR) data set. In a larger cohort of 290 renal lesions, Sun *et al.* selected 57 features to provide a classification model of RCCs *versus* other subtypes *versus* angiomyolipomas and ONCs. The resulting radiomic ML model yielded an AUC of 0.93–0.94 in differentiating RCCs from fat–poor benign renal lesions.[34,35] Uhlig *et al.* proposed a radiomic approach using CT features to discriminate malignant and benign clinical T1 renal masses, involving 94 patients for a total of 76 malignant lesions and 18 benign lesions. The best ML algorithm was the RF, which achieved the highest AUC, with 0.83 when compared with radiologists' assessment (AUC = 0.68). Interestingly, this result was obtained with 18 different CT scanners, confirming the role of radiomics in limiting the inter-observer and inter-machine variability and providing good results even in a pragmatic scenario.[36] In another study by Nassiri *et al.*,[37] performed on 684 patients with renal masses confirmed at CT imaging, two radiomics predictive models, the REAL AdaBoost and the RF, reported the best predictive performance with an AUC of 0.84 and 0.77, respectively, to distinguish benign from cancerous lesions overall and for SRMs when coupled with clinical variables. Finally, a large retrospective study by Yap *et al.*[38] involving 735 patients, showed, for different radiomic features (shape-only models, texture-only models, and combined models) increasing AUC from 0.67 to 0.75 in distinguishing malignant from benign renal tumors.

Similar studies have reported the efficacy radiomics using MRI in discriminating benign from malignant masses, utilizing the differences in contrast enhancement, heterogeneity, presence of cystic components, and signal intensity at T1, T2, and apparent diffusion coefficient (ADC) map signal. Xi *et al.* showed, analyzing 1162 renal lesions, a variable AUC for different clinical radiomics features that ranged from 0.52 to 0.76 of the ensemble DL model. Compared with expert radiologists, this predictive model reported higher accuracy, sensitivity, and specificity with, respectively, 0.70 *versus* 0.60, 0.92 *versus* 0.80, and 0.41 *versus* 0.35.[39] A smaller study by Said *et al.* involving 125 patients, reported, among significant qualitative and quantitative radiomic features, an AUC that ranged between 0.62 and 0.90, taken singularly. The related ML model reported instead, on validation sets, an AUC of 0.73 to differentiate RCC from benign lesions.[40] More recently, in a 2022 study performed by Xu *et al.*[41] on 217 patients, retrospectively analyzed, three DL models, created with ResNet-18 model, were evaluated using RF based on T2 weighted-imaging (T2WI) alone, diffusion-weighted imaging (DWI) alone, and an overlapping of the two image data sets to differentiate benign from malignant renal masses, yielding an AUC of 0.906 for T2WI, 0.846 for DWI and 0.925 for the combined model. Analogously, Massa'a *et al.*[42] investigated whether a high number of ML algorithms in 160 patients retrospectively analyzed, reporting the best results for the support vector machine (SVM) trained on T2WI (AUC = 0.79). Similar results were obtained for T1WI 4-min delayed features. Interestingly, the combination of radiomics features in this study did not raise the performance of the ML models. A summary of current studies that aimed to differentiate normal tissue from cancerous tissue is incorporated in Table 1.

Differentiation of AML from RCC

AML accounts for 40–55% of resected benign renal tumors and is characterized by the identification at the imaging of macroscopic fat surrounding and 'stuffing' the lesion.[43] Despite this peculiar characteristic, some AMLs could contain low intra-tumor fat that could not be detected by imaging.[44] The accurate characterization of those masses is, therefore, crucial due to the benign course and favorable prognosis of AML.[45] Quantitative approaches have been developed with the aim of detecting minor alterations in AML compared with RCC. One of the first studies to utilize a radiomics approach for this differentiation was performed by Feng *et al.*, which evaluated the capabilities of 42 CT-extracted features in a limited sample of 58 patients. In particular, 16 features were characterized by a significant intergroup difference in correctly diagnosing AML. The best features were selected *via* the SVM recursive feature elimination that reached an AUC of 0.939.[46] A similar study involved 95 patients for a total of 171 histopathological results from a single institution.[47] Features were obtained from three phases, which included

Table 1. Radiomics in studies differentiating benign from malignant kidney lesions.

Author	Clinical outcomes and gold standard	Imaging modality	Study design No. of patients	Imaging method	Results
Coy *et al.*[30]	- Discriminating benign lesions from cancerous lesions - Histology (surgery)	CT	Retrospective *n* = 200	Semi-automated method using in-house-developed software (U.S. FDA 510 K)[30]	Characterization of ccRCC, chRCC, papRCC, oncocytoma and AML had AUC of 0.850, 0.959, 0.792, and 0.825
Yu *et al.*[29]	- Discriminating benign lesions from cancerous lesions - Histology (surgery)	CT	Retrospective *n* = 119	Semi-automated method with manual segmentation. MATLAB software to perform texture analysis. SVM for classifying different tumor types	Histogram features skewness and kurtosis had the best discriminatory results (AUC = 0.91 and 0.93, respectively). ML AUC = 0.91–0.92
Zhou *et al.*[33]	- Discriminating benign lesions from cancerous lesions - Histology (surgery)	CT	Retrospective *n* = 192	Semi-automated method with manual segmentation. Preprocessing with Inception V3 software, pretrained on ImageNet and CNN models	Model trained on slice data set reported the worst result, with an ACC of 0.69. ROI data set reported ACC of 0.97, while RBR had an ACC of 0.93
Erdim *et al.*[32]	- Discriminating benign lesions from cancerous lesions - Histology (surgery)	CT	Retrospective *n* = 79	Semi-automated method with manual segmentation. Feature extraction *via* MaZda software	RF algorithm has been identified as having good prognostic potential with ACC of 0.905 and AUC of 0.905
Uhlig *et al.*[36]	- Differentiation of chrRCC and oncocytoma - Histology (surgery)	CT	Retrospective *n* = 94	Semi-automated method with manual segmentation. Feature selection with recursive feature elimination to build ML algorithms (RF) modeled to predict malignancy of specific renal mass	AUC of RF was 0.83, better than expert radiologists (AUC = 0.68)
Sun *et al.*[34]	- Differentiation of chrRCC and oncocytoma - Histology (surgery)	CT	Retrospective *n* = 290	Semi-automated method with semi-automated segmentation *via* Python v3.6.1 software. Radiomics was performed on volume of interest (VOI)	SVM model achieved a SENS ranging from 73–90% and a specificity ranging from 89–91.7% in distinguishing malignant from benign lesions
Yap *et al.*[38]	- Differentiation of chrRCC and oncocytoma - Histology (surgery)	CT	Retrospective *n* = 735	Semi-automated method with manual segmentation. 3D models used decision tree analysis (RF and REAL AdaBoost)	Median AUCs 0.68–0.75 achieved by combined models
Nassiri *et al.*[37]	- Differentiation of chrRCC and oncocytoma - Histology (surgery)	CT	Prospective *n* = 684	Semi-automated method with manual segmentation. VOI and decision tree analysis model (RF and REAL AdaBoost) has been used	Prognostic model achieved an AUC of 0.84
Said *et al.*[40]	- Differentiation of chrRCC and oncocytoma - Histology (surgery)	MRI	Retrospective *n* = 125	Semi-automated method with manual segmentation. OsiriX software. Radiomics analysis was performed by MRI physicist utilizing MATLAB	ML models (RF) with best results obtained an AUC of 0.73 in differentiate benign *versus* malignant lesions
Xi *et al.*[39]	- Differentiation of chrRCC and oncocytoma - Histology (surgery)	MRI	Retrospective *n* = 1162	Semi-automated method with manual segmentation	Ensemble DL model reported the highest test ACC, SENS, and SPEC, also when compared with the radiomics model
Massa'a *et al.*[42]	- Differentiation of chrRCC and oncocytoma - Histology (surgery)	MRI	Retrospective *n* = 160	Semi-automated method with semi-automated segmentation. HealthMyne software	Best algorithm (SVM) had ACC of 0.80 and an AUC of 0.79
Xu *et al.*[41]	- Differentiation of chrRCC and oncocytoma - Histology (surgery)	MRI	Retrospective *n* = 217	Semi-automated method with manual segmentation. ROIs were manually outlined. DL used ResNet-18 architecture and radiomics models used RF	Best performance of DL model (combination of T2WI and DWI) AUC 0.925

3D, three-dimensional; ACC, accuracy; AML, angiomyolipoma; AUC, area under the curve; ccRCC, clear cell RCC; chRCC, chromophobe RCC; CNN, convolutional neural network; CT, computed tomography; DL, deep learning; DWI, diffusion-weighted imaging; ML, machine learning; MRI, magnetic resonance imaging; papRCC, papillary RCC; RBR, rectangular box region; RCC, renal cell carcinoma; RF, random forest; ROI, region of interest; SENS, sensitivity; SPEC, specificity; SVM, support vector machine; T2WI, T2-weighted image; VOI, volume of interest.

the pre-contrast, the corticomedullary, and the nephrographic exposures.[47] Successively, an SVM algorithm has been developed to find the best classifiers for distinguishing AML and RCC. The best ML classifier reported an AUC of 0.96 for this purpose, significantly higher than that for differentiating AML from non-RCC.[47] With the increasing capabilities of radiomics in discriminating AML from RCC, other studies enriched the experience with this novel technology. Yang *et al.*,[48] for example, extracted 774 radiomics features from CT pictures to obtain the most discriminative model, reporting an AUC of 0.917, with a sensitivity of 0.66 and a specificity of 0.1, while Ma *et al.*[49] similarly constructed four single radiomics analysis logistic classifiers (which included five to seven features each one), to verify the proper diagnosis on 84 patients (22 with AML and 62 with RCC), reporting AUCs from 0.839 to 0.950. Analogous results have been obtained by Nie *et al.* developing a radiomics nomogram for the preoperative discrimination of AML from RCC. The built radiomics signature obtained promising results in the training data set (reaching an AUC of 0.879), which were confirmed in the validation set (AUC = 0.846) and four in the radiomics nomogram (AUC = 0.896–0.949). Also in this case, the number of features extracted were culled off from a massive number of CT-based features (over 2800).[50] Interestingly, in a retrospective study on 163 patients (118 RCC and 45 AML), digital picture features extracted from the unenhanced phase and fed into an ML model were similarly able to accurately discriminate between AML and RCC, reaching an AUC of 0.90.[51] Another radiomics approach considered the role of tumor and mini-peritumor features to differentiate AML from RCC in a study performed by Ma *et al.* on 230 patients, for a total of 58 AML and 172 RCC. With 396 radiomics features extracted, the best results were obtained for the nephrographic phase with an AUC of 0.726, followed by the corticomedullary phase (AUC = –0.694). This approach overcame the problem related to the accurate delineation of tumoral volume of interest (VOI).[52] The same authors provided, in addition, a radiomics CT nomogram for discriminating AML from RCC,[53] built using selected features reaching an AUC, for this nomogram, of 0.968.[54] More recently, Han *et al.* performed a retrospective research in 58 patients with AML and 140 with RCC, pathologically confirmed, to evaluate the prognostic value of CT radiomics in distinguishing AML from RCC. Five classifiers were used, for a total of 1029 features. The corticomedullary phase and nephrographic phase achieved an adequate performance (AUC = 0.767 and 0.783, respectively).[55] Similarly, Kim *et al.*[56] assessed the predictive role of CT radiomics in 28 AML and 56 RCC, reporting an AUC of 0.89, close to those of experienced radiologists (AUC = 0.78; Table 2).

For MRI, similar results were obtained. Razik *et al.*,[57] performed an MRI analysis to distinguish AML, RCC, and ONC, reporting in an MRI-based radiomics, an AUC > 0.8, with best performing parameter based on the mean of positive pixels (MPP) on DWI (AUC of 0.891). Jian *et al.*[53] instead, evaluated the combined use of MRI radiomics plus urinary creatinine for this purpose in a preliminary study, reporting the best AUC for the T2WI model (0.874), which increased up to 0.919 when combined with urinary creatinine, proposing the addition of other variables to radiomics approach to improve the diagnostic capabilities. Matsumoto *et al.*[58] demonstrated instead that the ADC map was enough in differentiating AML from RCC *via* a radiomics MRI-based approach, reporting good AUC (0.87) in the validation group.

Differentiation between oncocytoma and RCC

Renal ONC is usually a benign solid kidney neoplasm, which accounts for about 3–7% of all renal tumors.[59] Despite its benign natural history and excellent prognosis, ONC is usually treated with surgical resection due to the imaging available tools to properly distinguish it from RCC.[5] Indeed, due to the substantial overlap in imaging findings, differentiation of chromophobe RCC (chRCC) and clear cell RCC (ccRCC) subtypes through imaging modalities has traditionally been challenging.[60,61] Percutaneous biopsy represents a method to differentiate ONC from RCC with the risks of false-negative results, as well as the difficulty to consistently discriminate between RCC and ONC on pathology.[5,17] As a result, especially given the importance of distinguishing these two entities with such a difference in prognosis and tumor behavior, and to potentially avoid many unnecessary surgeries for benign lesions, a reliable non-invasive method that could properly differentiate ONC from RCC before surgery would be of particular clinical value. Several radiomic approaches have been proposed to this aim, showing promising results. Table 3 summarizes

Table 2. Radiomics in studies differentiating angiomyolipoma (AML) from RCC.

Author	Clinical outcomes and gold standard	Imaging modality	Study design No. of patients	Imaging method	Results
Feng *et al.*[46]	- Differentiation of AML from RCC - Histology (surgery)	CT	Retrospective *n* = 58	Semi-automated method with semi-automated segmentation *via* CT Kinetics software. Features were selected and classified *via* SVM	Optimal feature subset (using SVM and recursive feature elimination method), achieving an AUC of 0.955, SENS of 0.878 and SPEC of 0.1
Cui *et al.*[47]	- Differentiation of AML from RCC - Histology (surgery)	CT	Retrospective *n* = 171	Semi-automated method with semi-automated segmentation *via* ITK-SNAP v3.6.0 software. ROI was delineated manually. Texture analysis was performed *via* Python package PyRadiomics	ML classifier (SVM and SMOTE) had best performance to differentiate AML from RCC types and ccRCC. AUC = 0.96 and 0.97, respectively
Yang *et al.*[48]	- Differentiation of AML from RCC - Histology (surgery)	CT	Retrospective *n* = 60	Semi-automated method with manual segmentation. Feature extraction was performed *via* Spyder 3.2.8 software	Radiomics model (sRBFNN) yielded very good AUC (0.917), with an ACC, SENS, and SPEC of, respectively, 0.90, 0.66, and 1
Ma *et al.*[49]	- Differentiation of AML from RCC - Histology (surgery)	CT	Retrospective *n* = 84	Semi-automated method with manual segmentation. ROI was manually delineated and depicted *via* the software ITK-SNAP v3.4.0	Radiomics yielded an AUC of 0.988
Nie *et al.*[50]	- Differentiation of AML from RCC - Histology (surgery)	CT	Retrospective *n* = 99	Semi-automated method with manual segmentation. ROI was manually delineated and depicted *via* the software ITK-SNAP v3.8	Radiomics training signature (AUC = 0.879) Validation signature (AUC = 0.846)
Yang *et al.*[51]	- Differentiation of AML from RCC - Histology (surgery)	CT	Retrospective *n* = 163	Semi-automated method with manual segmentation. ROI was manually delineated and depicted *via* the software ITK-SNAP. Features were extracted using Python package Pyradiomics	SVM + *t*-score and SVM + relief, had the best performance (AUC = 0.90)
Ma *et al.*[52]	- Differentiation of AML from RCC - Histology (surgery)	CT	Retrospective *n* = 230	Semi-automated method with manual segmentation. VOI was depicted in ITK-SNAP software v3.4.0. Features were extracted *via* A.K. software, automatically	Nephrographic and corticomedullary phase's radiomics experienced best (AUC of 0.726 *versus* 0.694 in the training set and 0.767 *versus* 0.754 in the validation set)
Ma *et al.*[54]	- Differentiation of AML from RCC - Histology (surgery)	CT	Retrospective *n* = 139	Semi-automated method with manual segmentation. ROI was manually delineated and depicted *via* the software ITK-SNAP v3.4.0. 396 features were extracted using AK software	Radiomics model reported an AUC of 0.975 in the training set and 0.923 in the validation set
Han *et al.*[55]	- Differentiation of AML from RCC - Histology (surgery)	CT	Retrospective *n* = 198	Semi-automated method with manual segmentation. VOI was delineated manually	Corticomedullary and nephrographic phases achieved an adequate performance after using multilayer perceptron classifier (AUC of 0.85, SENS of 0.76, and SPEC of 0.78 for the corticomedullary phase and an AUC of 0.83, SENS of 0.79, and SPEC of 0.78 for the nephrographic phase)
Kim *et al.*[56]	- Differentiation of AML from RCC - Histology (surgery)	CT	Retrospective *n* = 84	Semi-automated method with semi-automated segmentation. Radiomics features were extracted using syngo.via Frontier software	Radiomics model reached an AUC of 0.89 resembling with that of radiologists (AUC = 0.78)
Razik *et al.*[57]	- Differentiation of AML from RCC - Histology (surgery)	MRI	Retrospective *n* = 42	Semi-automated method with manual segmentation. ROI was delineated manually. Features were extracted *via* TexRAD software	Several texture parameters reported AUC > 0.8. The best performing parameter was mean of positive pixels, with an AUC of 0.891 on DWI
Jian *et al.*[53]	- Differentiation of AML from RCC - Histology (surgery)	MRI	Retrospective *n* = 69	Semi-automated method with manual segmentation *via* ITK-SNAP software v3.6.0. ROI was manually delineated. Features were extracted *via* AK software	Radiomics model (AUC of 0.883) while Intravoxel incoherent motion-radiomics model (AUC = 0.874). Combined model (AUC = 0.919)
Matsumoto *et al.*[58]	- Differentiation of AML from RCC - Histology (surgery)	MRI	Retrospective *n* = 113	Semi-automated method with manual segmentation. Radiomics features were extracted *via* LIFEx software v4.00. VOI was manually delineated	Radiomics testing cohort (AUC = 0.90) Validation cohort (AUC = 0.87)

ACC, accuracy; AML, angiomyolipoma; AUC, area under the curve; ccRCC, clear cell RCC; chRCC, chromophobe RCC; CT, computed tomography; DWI, diffusion-weighted imaging; ML, machine learning; MRI, magnetic resonance imaging; papRCC, papillary RCC; RBR, rectangular box region; RCC, renal cell carcinoma; ROI, region of interest; SENS, sensitivity; SPEC, specificity; sRBFNN, sparse radial basis function neural network; SVM, support vector machine; T2WI, T2-weighted image; VOI, volume of interest.

Table 3. Radiomics in studies differentiating oncocytoma from RCC.

Author	Clinical outcomes and gold standard	Imaging modality	Study design No. of patients	Imaging method	Results
Baghdadi *et al.*[62]	- Differentiation of ccRCC from oncocytoma - Histology (surgery)	CT	Retrospective *n* = 192	Semi-automated method comprising two phases: (I) using DL network and manual tumor delineation; and (II) automated extraction of imaging features	SD Dice similarity score (0.66) for CNN model. PEER had ACC of 95% in tumor type classification (100% SENS and 89% SPEC) compared with the final pathology results[63]
Chen *et al.*[64]	- Differentiation of ccRCC from renal oncocytoma - Histology (surgery)	CT	Retrospective *n* = 94	CT whole lesions and region of interest evaluated	WL enhancement had poor results to differentiate between ccRCC and oncocytoma (AUC of 0.78 and 0.72, respectively), Combination model (AUC of 0.86)
Coy *et al.*[5]	- Differentiation of ccRCC from renal oncocytoma - Histology (not specified if surgery or biopsy)	CT	Retrospective *n* = 179	CT imaging features with neural network model	The best performance was obtained in the excretory phase (ACC = 74.4%, SENS = 85.8%, and positive predictive value = 80.1%)
Deng *et al.*[65]	- Discriminating RCC from noncancerous renal lesions - Histology (not specified if surgery or biopsy)	CT	Retrospective *n* = 501	An ROI was drawn on venous phase axial CT. Different texture analysis parameters were compared between cohorts	Differences in entropy were helpful in differentiation chrRCC from oncocytoma
Li *et al.*[66]	- Differentiation of chrRCC and oncocytoma - Histology (surgery)	CT	Retrospective *n* = 61	LASSO regression algorithm was used to analyze the CT image features. ROC curve and accuracy evaluation criteria	1029 features extracted. Diagnostic performance (AUC $>$ 0.85); SVM classifier had the best performance (AUC 0.96, SENS 0.99, SPEC 0.80, ACC 0.94)
Raman *et al.*[67]	- Differentiation of ccRCC, papRCC, oncocytomas and renal cysts - Histology (surgery)	CT	Retrospective *n* = 99	ROIs were drawn manually. A predictive model using quantitative parameters was constructed and externally validated	The RF model revealed 87% of ONC (SENS 89% and SPEC 99%). No AUC reported
Sasaguri *et al.*[68]	- Differentiation of oncocytoma *versus* RCC (papRCC and ccRCC and other subtypes) - Histology (surgery)	CT	Retrospective *n* = 166	CT tumor attenuation values and texture parameters used in-house (Matlab (MathWorks) software. logistic regression model used for differentiating types of renal lesions	AUC 0.91 for differentiating ccRCC and other subtype RCCs from papillary RCCs
Varghese *et al.*[35]	- Differentiation of malignant and benign renal masses (various subtypes) - Histology (surgery)	CT	Retrospective *n* = 174	WL were manually segmented and co-registered from CECT scans	Texture model had AUC of 0.87 ($p < 0.05$) for discriminating benign from cancerous kidney lesions
Varghese *et al.*[69]	- Differentiation of malignant and benign renal masses (various subtypes) - Histology (surgery)	CT	Retrospective *n* = 156	Manually segmentation of WL CT images (1) benign *versus* cancerous kidney lesions, (2) ONC *versus* ccRCC, and (3) ONC *versus* AML	ROC analysis (AUC curve $>$ 0.7, $p < 0.05$) between three groups
Yu *et al.*[29]	- Differentiation of ccRCC, papRCC, chrRCC and oncocytoma Histology (surgery)	CT	Retrospective *n* = 119	Manual segmentation of tumors. SVM method used for classification	ML applied to texture analysis to differentiate oncocytoma from other tumors (AUC of 0.86)
Hoang *et al.*[70]	- Differentiation of - benign and cancerous kidney lesions (oncocytoma *versus* ccRCC and papRCC) and RCC subtypes (ccRCC *versus* papRCC) - Histology (surgery)	MRI	Retrospective *n* = 41	Texture features from WL MRI slides	ONCs were distinguished from ccRCCs (SENS 67.3%, SPEC 88.9%, and ACC 79.3%), and from papRCC and ccRCCs (SENS 64.7%, SPEC 85.9%, and ACC 77.9%). No AUC reported
Paschall *et al.*[71]	- Differentiation of ccRCC *versus* papRCC and oncocytoma - Histology (not specified if surgery or biopsy)	MRI	Retrospective *n* = 55	WL measurements were performed. ROC curve analysis with optimal cutoff points was used to test the ability to the different groups	WL ADC values were very different between papRCC and oncocytoma. Best AUC = 95.8 for oncocytoma *versus* papRCC; SENS/SPEC 88.5% and 93.1% for oncocytoma *versus* papRCC, respectively

2D, two-dimensional; 3D, three-dimensional; ACC, accuracy; ADC, apparent diffusion coefficient; AML, angiomyolipoma; AUC, area under the curve; ccRCC, clear cell RCC; chRCC, chromophobe RCC; CECT, contrast-enhanced computed tomography; CNN, convolutional neural network; CT, computed tomography; DL, deep learning; DWI, diffusion-weighted imaging; IQR, inter-quartile range; LASSO, least absolute shrinkage and selection operator; ML, machine learning; MRI, magnetic resonance imaging; ONC, oncocytoma; papRCC, papillary RCC; PPV, positive predictive value; RCC, renal cell carcinoma; RF, random forest; ROC, receiver operator characteristics; ROI, region of interest; SD, standard deviation; SENS, sensitivity; SPEC, specificity; SVM, support vector machine; VOI, volume of interest; WL, whole lesion.

studies evaluating radiomics approaches for the differentiation of ONC from RCC.

Baghdadi *et al.*[62] aimed to design and evaluate a semi-automated method with the help of AI and image processing. To differentiate CD117-positive ONCs from chRCC, the authors used convolutional neural networks (CNNs) on CT obtained images. Tumor-to-cortex peak early-phase enhancement ratio (PEER) evaluation had an ACC of 95% in tumor type classification (100% SENS and 89% SPEC) compared with the histopathology results.

Chen *et al.*[64] evaluated the clinical utility of voxel parameters of whole lesion (WL) from CECT scans to differentiate ccRCC from kidney ONC. When compared with single ROI-based enhancement, WL enhancement did not perform well to distinguish ccRCC from ONC (AUC of 0.78 and 0.72, respectively). A combination with histogram parameters (AUC of 0.86) performed better. According to these results, authors concluded that the use of this method is probably not justified to be further studied to be clinically implemented.

Coy *et al.*[5] explored the performance of a DL lesion classifier, using a software library from Google (i.e. TensorFlow™ Inception), for the differentiation of ccRCC from ONC on CT images.[25] In this pilot study, 13 classification methods were tested, and the best performance was obtained using the excretory phase obtaining an accuracy of 74.4%, sensitivity of 85.8%, and positive predictive value (PPV) of 80.1%. The software used showed the potential of ML to discriminate the cancerous ccRCC from ONC. The experienced radiologists seemed to be able to properly classify an ONC better than the DL method.[25]

In their study on the distinction of RCC from benign tumors, Deng *et al.*[65] used histogram features to distinguish ONC from chRCC. Entropy seems to be a good prognosticator to differentiate ONC from chRCC.[25]

Li *et al.*[66] investigated the role of ML and CT image features to distinguish chRCC from ONC. In this analysis of 61 cases, five classifiers were trained to build a model. All radiomics models showed good diagnostic results (all AUC values >0.85), with SVM being the best (AUC 0.96, SENS 0.99, SPEC 0.80, and ACC 0.94), indicating that accurate preoperative distinction of ONC from chRCC might be eased by applying ML to CT imaging features.

Raman *et al.*[67] evaluated the possibility of differentiating common renal masses (i.e. ONC, ccRCC, cysts, and papillary RCC (papRCC)) using CT quantitative texture analysis and RF methods to construct a model. Analyzing CT scans from 99 patients, this approach demonstrated that 90% of oncocytomas and ccRCC could be identified with a sensitivity of 89% and a specificity of 99%, suggesting that data acquired from CT images can be used to accurately categorize different renal lesions, including oncocytomas.

Sasaguri *et al.*[68] searched to identify the role of biphasic CECT for the differentiation of SRM from RCC. The diagnostic performance of the proposed model achieved AUCs of 0.82, 0.95, and 0.84 for differentiating ONCs from ccRCCs.

Varghese *et al.*[35] explored the accuracy of quantitative features obtained from CT scans. The lipid content of lesions and the cancerous kidney tissues were studied. According to their analysis–which comprised 31 texture metrics derived with 6 texture methods, the histogram analysis did not perform well (74% of differences could be identified). The addition of Fourier analysis improved the results of the combined model (AUC of 0.90). The combined model had AUC values of 0.87, 1.00, 0.91, and 0.94 ($p < 0.05$) for differentiating ONC from cancerous kidney lesions.

Varghese *et al.*[69] in a retrospective series of 156 patients, fast Fourier transform (FFT) has been found to be in ONC *versus* ccRCCs, in the excretory phase. The heterogeneity of gray zone texture seems to be high in ccRCC when compared with ONC.

Yu *et al.*[29] assessed how texture analysis of images obtained from CECT can discriminate RCC subtypes from ONC. In this case series of 119 patients, histogram feature can differentiate ONC from other tumors (AUC of 0.92), and the ML combined model did not perform better (AUC of 0.86).

Hoang *et al.*[70] using quantitative texture parameters extracted from MRI aimed to discriminate benign and cancerous SRM. In this cohort of 41 patients, 45 imaging features were extracted,

comprising 5 global (intensity histogram) and 40 texture features. Histogram of features could differentiate ONC from papRCC and ccRCC (accuracy 77.9%, sensitivity 64.7%, and specificity 85.9%). These suggest that analysis of features drawn from MRI can better characterize kidney tumors subtypes.

Paschall *et al.*[71] tested objective volumetric WL on ADC map to improve the conventional measurements using ROI to discriminate type I papRCC from ccRCC and ONC. In this case series of 55 patients, WL ADC values between papRCC and ONC were significantly different ($p < 0.001$); ROC AUC of 67.6 for ONC *versus* ccRCC and 95.8 for ONC *versus* papRCC (sensitivity of 100.0%, specificity of 10.3%, and sensitivity of 88.5%, specificity of 93.1%, for ONC *versus* ccRCC and *versus* papRCC, respectively), highlighting the utility of this objective methodology in providing information on lesion heterogeneity and reducing observer bias.

Differentiation of different subtypes of RCC

RCC involves three major subtypes (i.e. ccRCC, papRCC, and chRCC), which differ in spatial distribution of cellularity and vascularity at histopathology.[72] The ccRCC is an aggressive and lethal carcinoma (75% of all RCCs), and has a potential to metastasize, while the papRCC and chRCC subtypes are less common – accounting for about 10–15% and 5% of all RCCs, respectively – and show better survival rates.[73] As a consequence, RCC subtyping has clinically implications, and the use of molecular targeted drugs will improve the differentiation of RCC subtypes. The application of radiomics to RCC subtyping has been reported by several studies, and summarized in Table 4.

Kocak *et al.*[74] aimed to extend the validity of their results to externally validate, to allow replication of models and the possible generalization of algorithms. This used CT images and features in combination with ML algorithms. The best performance was achieved by the ANN classifier with adaptive boosting, showing an accuracy of 84.6% for differentiation of ccRCCs from other tumor types. The SVM classifier performed best (ACC = 69.2%) to differentiate ccRCC from papRCC and chRCC. The best performance was found to be related to the differentiation of papRCC from other RCCs, while they exhibited rather poor performance in differentiating ccRCC or chRCC from others.

Similarly, Han *et al.*[75] aimed to exploit reproducible and generalizable models to differentiate the ccRCC, papRCC, and chRCC using CT images along with an ML algorithm. In their cohort of 169 biopsy-proven RCC cases, the DL neural network achieved an AUC of 0.9 no matter the identified subtype (specifically, 0.93 for ccRCC, 0.91 for papRCC, and 0.87 for chRCC), showing promising performance in classification of RCC – although with worse performance for chRCC subtyping.

Li *et al.*[76] evaluated a CT radiomics model to differentiating ccRCC from other tumor subtypes and to evaluate radiogenomics potential combining the imaging features and the von Hippel-Lindau (VHL) mutation gene. Among 156 texture features extracted for each tumor, the eight most relevant from the corticomedullary phase were used to build the model, which had a good AUC (0.95; ACC of 92.9%); moreover, five out of eight had a strong association with VHL mutation gene.

Raman *et al.*[67] sought to assess the possibility of differentiating common renal masses (i.e. ONC, ccRCC, cysts, and papRCC) using CT texture analysis features embedded into a model. The model used was RF. This model correctly categorized ccRCCs in 91% of patients (SENS 91% and SPEC 97%), and papRCCs in 100% of patients (SENS 100% and SPEC 98%), suggesting CT texture analysis, in conjunction with RF modeling, might demonstrate a potential method to characterize renal masses.

Leng *et al.*[77] in their study exploring the effect of denoising heterogeneity scores. The aim was to distinguish AML from different subtypes of RCC. The authors found that, with regards to ccRCC and papRCC differentiation, the heterogeneity scores could discriminate these two subtypes, and that further reduction in noise improved AUC.

Yan *et al.*[78] investigated the diagnostic performance of texture analysis for the discrimination of AML with minimal fat, ccRCC, and papRCC on images obtained from CT scans. According to their analysis, excellent classification results in terms of discrimination between ccRCC and papRCC were obtained with nonlinear discriminant analysis (error of 0.0–9.3%), no matter

Table 4. Radiomics in studies differentiating different subtypes of RCC.

Author	Clinical outcomes and gold standard	Imaging modality	Study design No. of patients	Imaging method	Results
Kocak *et al.*[74]	- Differentiation of ccRCC, papRCC, and chRCC - Histology (either surgery or biopsy)	CT	Retrospective *n* = 68	Feature selection was done by three radiologists. Feature selection and model optimization has been performed. Using ANN and SVM as classifiers and a combination of three additional algorithms aimed to improve generalizability	For differentiating non-ccRCCs from ccRCCs, the best performance was the ANN classifier (ACC of 84.6%). The performance was poor for distinguishing ccRCC *versus* papRCC *versus* chRCC. Best SVM classifier with bagging algorithm (ACC of 69.2%)
Han *et al.*[75]	- Differentiation of ccRCC, papRCC, and chRCC - Histology (biopsy)	CT	Retrospective *n* = 169	Rectangular ROI was marked and cropped. A DL neural network has been used to identify subtypes of RCC	Deep DL method with the contouring given by radiologists for RCC subtype classification achieved an ACC 0.85, SENS 0.64–0.98, SPEC 0.83–0.93, and AUC 0.9
Li *et al.*[76]	- Differentiation of ccRCC from non-ccRCC - Histology (not specified if surgery or biopsy)	CT	Retrospective *n* = 170 (training cohort) *n* = 85 (validation cohort)	Two radiomics models were built. The radiogenomics association between selected features and VHL mutation has been analyzed. All models were independently validated	The model from obtained from corticomedullary images from CT had AUC of 0.95 (ACC of 92.9%)
Raman *et al.*[67]	- Differentiation of ccRCC, papRCC, oncocytomas and renal cysts - Histology (surgery)	CT	Retrospective *n* = 99	ROIs were delineated in different phases of CECT images. Heterogeneity has been further assessed. A predictive model using quantitative parameters was constructed and externally validated	Various renal masses (oncocytomas, ccRCC, cysts, and papRCC) were accurately classified. The RF algorithm better categorized ccRCCs in 91% of images (SENS 91% and SPEC 97%), and papRCCs in 100% of cases (SENS 100% and SPEC 98%)
Leng *et al.*[77]	- Differentiation of ccRCC and papRCC and AML - Histology (surgery)	CT	Retrospective *n* = 139	A largest possible ROI was manually drawn and SD, entropy, and uniformity were analyzed. Heterogeneity indices were further assessed with a denoising algorithm	Heterogeneity indices have the ability to differentiate ccRCC from papRCC. Best AUC (0.91) for the subjective score
Yan *et al.*[78]	- Differentiation of ccRCC and papRCC and AML - Histology (surgery or biopsy)	CT	Retrospective *n* = 50	Native and CECT images were analyzed and classified with texture analysis software (MaZda). Tumor attenuation values and enhancement degree was determined by an ROI	For the discrimination between ccRCC and papRCC, excellent classification results were obtained with nonlinear discriminant analysis; on comparison of the three scanning phases, better lesion classification was observed with corticomedullary and nephrographic phase's images
Hoang *et al.*[70]	- Differentiation of - benign and cancerous kidney lesions (ONC *versus* ccRCC and papRCC) and different RCC subtypes (ccRCC *versus* papRCC) - Histology (surgery)	MRI	Retrospective *n* = 41	The features obtained from native a contrast MRI images have been analyzed. Lasso regression used for false rate results	PapRCC was distinguished from ccRCC with an ACC of 77.9% (SENS 65.5% and SPEC 88.0%)
Li *et al.*[79]	- differentiation of ccRCC, papRCC, chrRCC, AML and oncocytoma - Histology (surgery)	MRI	Retrospective *n* = 92	ADC maps were constructed from FOV DWI images to identify the histogram parameters	ADC histogram parameters differentiated eight of 10 pairs of renal tumors
Paschall *et al.*[71]	- Differentiation of ccRCC *versus* papRCC and oncocytoma - Histology (not specified if surgery or biopsy)	MRI	Retrospective *n* = 55	WL measurements were performed and ADC map constructed from WL histogram	WL ADC features could discriminate papRCC from ONC. Best percentile ROC analysis demonstrated AUC of 95.2 (sensitivity of 84.5% and specificity of 93.1%)

2D, two-dimensional; 3D, three-dimensional; ACC, accuracy; ADC, apparent diffusion coefficient; AML, angiomyolipoma; ANN, artificial neural network; AUC, area under the curve; ccRCC, clear cell RCC; chRCC, chromophobe RCC; CECT, contrast-enhanced computed tomography; CT, computed tomography; DL, deep learning; DWI, diffusion-weighted imaging; FOV, field of view; IQR, inter-quartile range; LASSO, least absolute shrinkage and selection operator; ML, machine learning; MRI, magnetic resonance imaging; mRMR, minimum redundancy maximum relevance; ONC, oncocytoma; papRCC, papillary RCC; PPV, positive predictive value; RCC, renal cell carcinoma; RF, random forest; ROC, receiver operator characteristics; ROI, region of interest; SD, standard deviation; SENS, sensitivity; SPEC, specificity; SVM, support vector machine; VOI, volume of interest; WL, whole lesion.

which phase was used; on comparison with the three scanning phases, a trend toward better lesion classification was observed with corticomedullary and nephrographic phases images.

Hoang *et al.*[70] assessed if quantitative texture parameters obtained from MRI could potentially differentiate between common subtypes of RCC (i.e. ccRCC *versus* papRCC) in SRMs (i.e. <4 cm). Among the 45 imaging features extracted, textures helped to differentiate between subtypes of RCCs; papRCCs were differentiated from ccRCCs with an accuracy of 77.9% (sensitivity 65.5% and specificity 88%).

Li *et al.*[79] used volumetric histogram analysis from ADC maps to characterize SRMs. Combination of mean ADC and histogram values achieved the best AUC (0.851 with SENS of 80.0% and SPEC of 86.1%) and these results show that the volumetric analysis could potentially differentiate between certain types of kidney masses.

Paschall *et al.*[71] explored the volumetric WL ADC parameters could identify RCC. The differentiation was studied for type I papRCC from ccRCC and ONC. In their study, WL ADC could distinguish between papRCC and ccRCC ($p<0.001$), with AUC of 95.2 (SENS of 84.5% and SPEC of 93.1%).

Fuhrman grade prediction

Fuhrman grade is an important pathological risk factor impacting on patients' oncological outcomes, especially the risk of recurrence.[80] Indeed, despite this information might be achieved through renal mass biopsy, this procedure is invasive and not devoid of complications and suffers from several limitations.[17,81] Therefore, being able to preoperatively predict a renal mass nuclear grade differentiation, directly by imaging, could be of outmost importance to address the patient to the best treatment.

With the introduction of ML, a branch of AI-developing algorithm able to both learn and improve by analyzing data sets, one of the main imaging parameters adopted is texture analysis.[82,83] This post-processing technique, which can be applied either to CT or MRI, allows quantifying tumor heterogeneity assessing several parameters.

Shu *et al.* compared radiological features of different Fuhrman grade ccRCC and extracted 1029 radiomics features from corticomedullary and nephrographic CT scans. The authors found that 11 and 24 features correlated with Fuhrman grades. This investigational analysis confirmed that radiomics can assess preoperatively the Fuhrman grade of kidney lesions.[84] Again, a retrospective study on 290 patients with histologically confirmed 298 RCCs evaluated the levels of entropy and texture quantification within renal tumors at CT imaging. The authors found a significant increase of entropy value both in clear cell carcinoma and higher Fuhrman grade.[85]

Recent evidence from studies that looked at the role of ML also analyzed texture in MRI imaging as well. T2 and DWI windows are usually used and a study on 34 RCC masses demonstrated entropy at spatial scaling factors (SSF) on DWI, on corticomedullary phase and on nephrographic phase to be best parameters to assess RCC grading.[86] Accordingly, Stanzione *et al.*[87] developed five algorithms including different MRI features to predict tumor grading achieving accuracy greater than 90%.

Yin *et al.* developed and tested an ML model and images from CECT to predict Fuhrman grade of ccRCC. In 25 patients, SVMRadial, RF and Bayesian models had the best prognostic ability to predict Fuhrman grade of ccRCC using radiomics from CECT images.[88]

Finally, semantic segmentation is gaining popularity[89] and promising results have been achieved to differentiate different RCC subtypes.[90] Nevertheless, studies able to underline the role of semantic segmentation to differentiate nuclear grading are still based on the pathological sample.[91]

Prediction of gene mutation molecular biomarkers through radiomics

Radiogenomics integrates multi-scale genome data, with the help of refined CAD systems to develop imaging possibilities to assess the combination of imaging data with genome-related cellular data.[92–96]

Lee *et al.*[97] evaluated 58 kidney cancer patients (including 12 patients with metastatic disease) using a radiomics algorithm with images from CT

scans assessing progression of pT1 RCC. Combining radiomics parameters with gene expression data gathered from whole transcriptome sequencing (WTS). Four radiomic features, which included histogram features, gray-level co-occurrence matrix (GLCM), and the ratio of voxels from ROIs, were trained to prognose metastasis of patients. In addition, heterogeneous gene signatures correlated with these radiomics features were identified. These findings barred out the use of radiogenomics to highlight patients who could have an additional benefit from adjuvant therapy or metastases in pT1 RCC.

Two retrospective studies assessed more granularly specific mutations and CT-based texture radiomics, such as BAP1 mutations. The first, reported an AUC of 0.77 highlight BAP1 mutation.[98] The second study assessed radiomics features extracted from CT scans of 65 ccRCC tumors, achieved a SENS, SPEC, and ACC of 90.4%, 78.8%, and 81% to predict BAP1 mutation (AUC = 0.89).[99] The gene encoding the protein polybromo-1 (PBRM1) mutation has been investigated with radiomics analysis with a good AUC of 0.925.[100] Previously, one trial studying ccRCC tumors found an AUC of 0.85 for VHL, PBRM1, and BAP1 genes.[101] Gene mutations BAP1 has been investigated in 78 tumors from The Cancer Genome Atlas, (AUC = 0.71 for prediction of BAP1 within the CT nephrogenic scan images.[102]

Studies investigating treatment response of renal masses using radiomics

Assessing the prediction of response to target therapies is paramount for clinical decision-making in metastatic RCC (mRCC) patients. Targeted therapies and immunotherapies have demonstrated a promising efficacy in mRCC, yet it remains challenging to delineate subgroups of responsive patient, despite the existence of several scores (e.g. International Metastatic RCC Database Consortium [IMDC] risk score) and biomarkers (e.g. PBRM1).[26,103,104] A deeper knowledge and evaluation of response/resistance status would lead to personalized algorithms, which can potentially avoid adverse events of unnecessary treatments – with a subsequent positive effect on patients' quality of life – optimize resources, save time (e.g. unresponsive patients to certain therapies could be switched earlier to other treatments) to improve survival outcomes.

Antunes *et al.*[105] attempted to test how well radiomics analysis perform on integrated positron emission tomography/magnetic resonance imaging (PET/MRI) in the assessment of metastatic RCC and response to Sunitinib. Their findings suggests that radiomics from PET/MRI have a potential to identify structural and functional modifications that can influence the response to tyrosine kinase inhibitor (TKI) therapy, thus identifying radiomics analysis as a modality for characterization and evaluation through PET/MRI.

Bharwani *et al.*[106] sought to explore whether TKI therapy (i.e. Sunitinib) can influence sequential gene changes in mRCC patients and its correlation with overall survival (OS), by prospectively assessing DWI and multiphase contrast-enhanced MRI as biomarkers of outcomes. In this case series of 20 patients, 47% of patients had a modification mean ADC following treatment, despite no correlation with outcome was found. Patients with higher baseline AUC (low; i.e. proportion of the tumor with ADC values $<$ 25th percentile of the ADC histogram) and greater-than median AUC-low increase, reported a reduced OS (hazard ratio (HR) = 3.67, 95% confidence interval (CI) = 1.2310.9; $p = 0.012$ and HR = 3.72, 95%CI = 0.98–14.21; $p = 0.038$, respectively), indicating that DWI-MRI can be a possible biomarker for OS.[106]

Boos *et al.*[107] evaluated median *versus* mean attenuation gathered from histograms of 19 RCC patients receiving Sunitinib or Sorafenib response assessed by CT scans. Authors found that distribution curves correlated themselves with outcomes (RECIST criteria were employed); lesions with –44 Hus, had a partial tumor response while those greater than –41 Hus, reported tumor progression.[35]

The study by Goh *et al.*[108] aimed to evaluate tumor textures on CT images correlation with PFS in 39 patients with mRCC – different subtypes – who received multiple TKIs (i.e. Sunitinib, Cedirinib, Pazopanib, or Regorafenib). By analyzing 87 metastatic lesions prior and after therapy, authors have found a lower texture entropy and higher uniformity after immunotherapy. Texture uniformity has been shown to be an independent predictor of progression ($p = 0.005$). According to these data, authors suggested that tumor heterogeneity could have the potential to

be assessed as predictive radiomics marker of response to therapy.

Haider *et al.*[109] performed a retrospective assessment of combined biomarkers and CT images to prediction survival in mccRCC patients receiving immunotherapy. Specifically, measurable lesions from different selection respected RECIST criteria, prior and after therapy. The evaluation of IMDC score plus texture parameters was also performed. Size normalized standard deviation (nSD) before and after therapy is a predictor of OS ($p = 0.01$ and 0.01); entropy is also a significant predictor of OS before and after therapy ($p = 0.02$ and $p = 0.04$). Same results were obtained for PFS at ($p = 0.01$ and $p = 0.003$) concluding that nSD correlates with OS and PFS.

Mains *et al.*[110] aimed to identify which DCE-CT functional parameters and models had the optimal correlations with OS and PFS in mRCC patients. According to these analyses, the best associations were spotted for blood volume, blood flow, and standardized perfusion values, calculated using deconvolution at baseline and during early treatment period ($p < 0.05$).

The study by Khene *et al.*[111] studied response to TKIs treatment on a model relying on k-nearest neighbor, RF, logistic regression, and SVM approaches with the worse ACC and AUC for RF and the best for logistic regression. Table 5 summarizes studies investigating radiomics and treatment response.

Table 5 summarizes studies investigating treatment response in mRCC patients.

Ongoing studies evaluating the radiomic value in kidney lesions

Radiomics has been heavily studied lately to assess its potential to individualize treatment. The correlative radio-genomics and theranostics could hold the promise to provide useful information for the detection, prediction, treatment, and how the kidney tumor reacts to oncological treatment.[23] The specific field of radiomics applied to kidney lesions for different purposes is currently moving at an unprecedented pace. Indeed, several ongoing trials are assessing new imaging techniques for RCC management – mainly in terms of characterization and staging – using different modalities and tracers. Table 6 summarizes

Table 5. Radiomics in studies investigating treatment response in metastatic RCC undergoing TKIs or T-cell immunotherapy.

Author	Treatment, ROI and end-point	Imaging modality	Study design No. of patients	Imaging method	Results
Antunes *et al.*[105]	- Sunitinib - Primary tumor - Modifications of radiomics features after TKI therapy	FLT-PET/MRI	Prospective $n = 2$	Image-related features before and after treatment using [18 F] FLT-PET, T2w, and DWI protocols, with DWI reporting an ADC map	The best radiomics yielded a modification of 63% within the RCC region and 17% in a distinct normal region
Bharwani *et al.*[106]	- Sunitinib - Primary tumor - Changes in histogram parameters and correlation with OS (changes prior and after treatment)	MRI	Retrospective analysis of prospective study $n = 20$	ADC maps and histograms have been assessed. Mean ADC and proportion of the tumor with ADC values $<$ 25th percentile of ADC histogram were recorded. ROI were manually delineated. Changes prior and after therapy in surviving patients have been compared for OS	Outcomes did not correlate to features. High baseline AUC low and greater median AUC low have been associated with poor OS ($p = 0.038$). OS had no correlation with MRI features
Boos *et al.*[107]	- TKI (Sunitinib $n = 18$, Sorafenib $n = 1$) - Measurable soft tissue lesion - Change in CT intensity distribution curves	CT	Retrospective $n = 19$	Histograms delineated from ROI. Shift was used to classify response of lesions to therapy and any modifications on scans, using the Choi, MASS, and RECIST criteria	Changes in histograms appeared in 58% of lesions, and a significant difference between mean and median lesion attenuation ($p < 0.001$). There has been an increased in changes of the accurate classification of tumors when Choi and Mass criteria were evaluated (63–68% and 74–79%)

(Continued)

Table 5. (Continued)

Author	Treatment, ROI and end-point	Imaging modality	Study design No. of patients	Imaging method	Results
Goh *et al.*[108]	- TKI (26 patients with sunitinib, 6 patients with cediranib and 4 patients with pazopanib, and 3 patients with regorafenib) - metastases. - Change in histogram parameters (entropy and uniformity) and correlation of texture parameters with PFS	CT	Retrospective $n=39$	A CAD software algorithm appreciated the changes in entropy and uniformity of metastasis. RECIST, Choi, and modified Choi criteria evaluated the response. The correlations of texture parameters and standard criteria with PFS have been assessed	Tumor entropy decreased and uniformity increased following TKI therapy. Kaplan–Meier curves of patients without disease progression reported better outcomes compared with standard response assessment ($p=0.008$ *versus* 0.267, $p=0.053$, and $p=0.042$ for RECIST, Choi, and modified Choi criteria, respectively). Texture uniformity was an independent predictor of time to progression ($p=0.005$)
Haider *et al.*[109]	- Sunitinib - Measurable lesion - Correlation of texture parameters with OS and PFS in ccRCC	CT	Retrospective $n=40$	Measurable lesions on CECT before and 2 months after therapy. TexRAD software (TexRAD Ltd, Cambridge, UK) has been employed to analyze textures. Cox regression model assessed changes in texture features and PFS/OS	Size-normalized SD (nSD) alone is good predictor of OS ($p=0.01$). Entropy modifications are a good predictor for OS ($p=0.02$ and $p=0.04$) and nSD can prognoses PFS ($p=0.01$ and $p=0.003$)
Mains *et al.*[110]	- Various treatments, not specified - Large artery - Association between OS and PFS with functional CT parameters	CT	Retrospective analysis of prospective study $n=69$	Scans performed at prior and after therapy (after 5 and 10 weeks). BVdeconv, BFdeconv, SPVdeconv, blood flow and standardized perfusion values (BFmax and SPVmax), were evaluated using the Patlak model (BVpatlak and PS)	The strongest association was found for BVdeconv, BVpatlak, and BFdeconv prior and after therapy ($p<0.05$). PS seemed to have opposite associations dependent on treatment. Inter-observer correlations were excellent ($r\geqslant 0.9$, $p<0.001$) with good agreement for BFdeconv, BFmax, SPVdeconv and SPVmax
Khene *et al.*[111]	- Nivolumab - Predict response to treatment	CT	Retrospective $n=48$	K-nearest neighbor, RF, logistic regression, and SVM algorithms have been used. Classification of patients: complete or partial response or stable disease and non-responders	95% of patients received nivolumab. 60.4% of patients were nivolumab responders. The ACC (0.71 till 0.91) and the AUC (0.67 till 0.92). RF reported the worse accuracy, while logistic regression the highest

ACC, accuracy; ADC, apparent diffusion coefficient; AUC, area under the curve; CAD, computer-aided diagnosis; CECT, contrast-enhanced computed tomography; CT, computed tomography; DWI, diffusion-weighted imaging; FLT-PET/MRI = F18 fluorothymidine-positron emission tomography/MRI; HR, hazard ratio; HU, Hounsfield unit; MASS, morphology, attenuation, size, and structure; MRI, magnetic resonance imaging; nSD, size-normalized SD; OR, odds ratio; OS, overall survival; PFS, progression-free survival; RCC, renal cell carcinoma; RECIST, response evaluation criteria in solid tumors; RF, random forest; ROI, region of interest; SD, standard deviation; SVM, support vector machine; T2w, T2-weighted; TKI, tyrosine kinase inhibitor; VOI, volume of interest; WL, whole lesion.

Table 6. Radiomics in ongoing trials investigating kidney lesions.

NCT number and topic	Status and no. of patients	Intervention	Outcome
NCT04271254 PET/MR characterization of RCC[112]	- Recruiting - 17	PET/MRI	Aims to establish the role of PET and MRI to define the molecular subtypes of ccRCC. The imaging features and the fraction of tumor cores will be analyzed[112]
NCT03996850 SPECT/CT for the characterization of renal masses[113]	- Recruiting - 100	99mTc-sestamibi SPECT/CT	To evaluate the role of technetium 99m sestamibi SPECT/CT on patients management
NCT04295174 KIDSTAGE- Staging of kidney cancer using dual time PET/CT and other biomarkers[114]	- Recruitment completed - 70	Fluorodeoxyglucose PET/CT	To evaluate fluorodeoxyglucose PET/CT in staging of secondary determination from renal neoplasia and to investigate the role of cfDNA for follow-up of renal cancer patients
NCT03470285 Multiparametric MRI for diagnosing small renal tumors (IRMK01)[115]	- Recruiting - 500	mpMRI	To assess diagnostic accuracy of mpMRI
NCT02526511 Perfusion magnetic resonance imaging in diagnosing patients with kidney tumors[116]	- Recruiting - 50	Perfusion MRI	The diagnostic ability of perfusion MRI to predict cancerous *versus* benign kidney confined lesions
NCT03821376 Correlation of renal mass pathologic grade and CEUS[117]	- Recruiting - 40	CEUS	To correlate CEUS findings with the Fuhrman grade of RCC

ccRCC, clear cell RCC; CEUS, contrast-enhanced ultrasound; cfDNA, cell-free DNA; CT, computed tomography; FDG, ^{18F} fluoro-D-glucose; mpMRI, multiparametric MRI; MRI, magnetic resonance imaging; PET, positron emission tomography; RCC, renal cell carcinoma; SPECT, single-photon emission CT.

main ongoing trials exploring radiomics and its ability to establish the stage and grade of RCC.

Comparison of AI algorithms used in radiomics studies

AI algorithms are defined as mathematical models of computers that allow the hardware to learn and work independently.[118] The ML algorithms are classified into four major categories: supervised, unsupervised, semi-supervised, and reinforcement learning.[119] Supervised learning uses labeled data and collection training data to accomplish a task.[120] Decision tree, random forest, SVM, naïve Bayes, linear regression, and logistic regression are most frequently used in medical application area.[119] Unsupervised learning performs clustering, which means they sort the unlabeled data points into pre-defined clusters. The data point has to belong to one cluster with no overlap. There can be more than one data point in any given cluster, but a data point cannot belong to more than one cluster, without human intervention.[119,121] K-means clustering and Gaussian mixture model are the most common employed learning methods that uses unlabeled data to develop models and to extract generative features, groupings in results, and exploratory purposes.[119–121] Both supervised and unsupervised algorithms, such as K-nearest neighbor, neural networks, and reinforcement learning algorithms are being used in the evaluation of medical images. In this study, the identified articles used supervised ML algorithms. Furthermore, we will summarize the results and comparison derived from these studies. Some studies did not use any ML algorithms to analyze the variables obtained.[64,71,77–79,105–108] SVMs are identified as the most used algorithms used for classification in the analyzed studies.[34,36,42,46,47,51,55,66,74] When trained, the learning process searches to differentiate between two data sets (for example, benign

from malignant kidney tumors or to differentiate between certain tumor types and some of them to predict gene mutation or response to therapy). The data used to train and learn SVMs are not entirely used for this purpose. Just the closest data between two support vectors, which makes it a linear algorithm, that is useful when the two data sets are clearly separated. When data are perturbed, the performance is reduced.[122,123] The latest developed AI algorithms are artificial neural networks (ANNs) that modulate the human brain functions.[124] ANNs were developed and used in several renal radiomics studies for the differentiation of chRCC and oncocytomas[66] and with adaptive boosting for differentiating non-cc-RCCs from cc-RCCs.[74] ANNs have artificial nodes, in layers that can have different transfer functions. The performance is good when data are of great quantity, as in real life and have the ability to identify and model complex relationship among data. The drawbacks are represented by the training that results in relative minimum and not absolute minimum of the error function and the overfitting of data because it lacks the ability to generalize data (it is relatively easy to observe their effect when the performance of training and test data sets starts to split in opposite directions).[125] Two studies used ANNs[66,74] achieving best results in AUC along with SVMs, and two used CNNs[62,75] having the advantage of not using hand-crafted features from experts. RFs algorithms combine predictive data of decision trees (forests) in one model. Each decision tree learns from data that are chosen randomly and calculates the average of the predictions as the final result.[126] RF has advantage of dealing with non-linear data and can reduce the variables space to emphasize the value of each feature.[127] It has been used in vast number of studies that analyzed radiomics and kidney tumors. RFs obtained the best AUC in the articles published on radiomics and kidney lesions.[32,36,38] Lasso regression (least absolute shrinkage and selection operator regression) is one of the algorithm models that modify and push regression coefficients to the nil value and improves the interpretability and can select important predictors for the studied pathology.[128,129] In the renal lesions, radiomics studies were used to reduce overfitting with good results.[49–53,55,70,111]

The advancements in the study of radiomics and ML algorithms are huge, but still approximately 30% of studies used traditional algorithms for comparison.[51,66,101,130–132] Therefore, direct comparisons have not been published in the literature so far. More data are required for a more precise evaluation and generalization of the best ML algorithms to be used in radiomics research.

Current limitations and the future of AI and radiomics

Generally, when researchers develop AI models that can be applied to the analysis of radiomic features extracted from images,[133] they must consider the robustness and standardization of the proposed model.[134] We have found that a rather high number of researches have been performed using non-robust features.[135] Current imaging methods are not standardized, whether we talk about CT or MRI. All technical specifications of using different CT phasing, contrast enhanced or not, and also the classifiers used to discriminate between different radiomics features applied to analyze the gray zone levels of the studied renal lesions (SVM, histograms, and others) are not validated independently. Currently, the lack of internal and external validation of the proposed models cannot lead to the generalizability of these methods.

The reason of dimension reduction in radiomics is a method to increase the modeling performance using the highly relevant, robust, or uncorrelated features. This can lead to data perturbations[136] and contamination[137] and add bias to validation purposes. ML classifiers can select features during model development[67] and limit the biases of validation techniques. DL uses end-to-end automatic pipelines and in radiomics, the robustness and selection of features is not properly understood. The currently used ML models did not make it clear as to which one of those are better to be used or what combination of imaging methods should be used (contrast enhanced or not). In depth, analysis of the most appropriate AI method to be applied (or combinations) should be studied further. MRI seems to have less ability to provide a good evidence for the use in radiomics. The DL methods provide good evidence that support further research for clinical use especially with its combination with genomics data (radiogenomics).[33] Only a few studies assessed the comparison between traditional biomarkers and radiomics in terms of clinical validity.[47,51,99,101,130,131] Radiomics is able to evaluate large data sets, but the manual predefinition of metrics limits its performance.

Discussion

The novelty that radiomics can provide more objective interpretation of images to limit the subjective image interpretation of radiologists with the aim of improving renal tumor diagnostic accuracy[14] is gathering more followers and this research is keen to provide deeper insights on the possible use and clinical practice translation. There is an important percentage of renal lesions (up to 30%)[9] that are benign at histopathology result for patients that underwent partial or radical nephrectomy and this is one of the reasons that radiomics is seen with the potential to improve the preoperative detection of benign tumors. AI and its subset ML is increasingly used in radiomics analysis for assessing renal lesions. Being a new area of research and development, there are still a lot of challenges that limit the implementation in current practice. When we look at the studies that analyzed the role of radiomics in differentiating benign from cancerous tissues, we see that all identified articles have retrospective design, and another limitation is due to the small number of images or patients included in studies, that also still need manual delineation and contouring of ROIs or VOIs, and having semi-automated method of analysis. Also using different scanners, protocols for obtaining images, AI software for texture, shapes, and volumetric or geometrical analysis of big data limits the possibility of properly defining gray zone features. AI through ML and DL may even go few steps further and may help to automize the acquisition and reporting of ultrasound, CT scans, and MR images. The first step in the process of reducing the workload and costs was kidneys volumetric analysis, which, for instance, can now be completely performed by DL algorithms both on ultrasound[138,139] and CT scans.[140,141] This can improve the accuracy of images pre-processing for subsequent radiomics features extraction.[141] Moreover, accurate volumetric segmentation of kidneys and tumor is pivotal when nephron sparing surgery is planned in patients with kidney cancer.[142]

Models trained and validated indicate a roughly big difference in AUCs obtained from these studies (from 0.64 to 0.97). These can lead to the conclusion that by using different AI algorithms and mathematical models for statistical analysis (most used: RF, SVM, logistic regression, and linear discriminant analysis) show that radiomics still is an area of intense research and has limitations in understanding the fundamentals of how AI and ML can improve the outcomes of studies. The comparison of the AI models with expert radiologist need to be further improved in prospective studies and we believe that future result will show the ability, especially of DL algorithms, to ease the clinical embrace of radiomics and AI. AI may be the key to provide fast, reliable reports from all radiologists, reducing inter-reader variability, and to obtain improved reading accuracy. Conclusions are indeed subjective due to the radiologist's interpretation and experience.[143,144] Toda *et al.*[145] demonstrated that DL algorithms in contrast-enhanced CT have high accuracy for the diagnoses of SRMs with both internal and external validation. Manual or semi-automated segmentation have been used in most of the studies (either on CT or MRI platforms); Kart *et al.*,[146] using national databases of whole-body MR imaging from United Kingdom and Germany, developed and trained an automated segmentation DL model for abdominal organs; and Zhao *et al.*[147] clinically assessed assisted compressed sensing technology in renal MRI imaging with an AI algorithm that can adjust scanner settings to improve image acquisition and automatically adjust images to patients' movements, and can allow ultra-fast MR imaging acquisition.

Good accuracies have been reported[5,33] and one author[39] implemented a DL method to study a variety of kidney lesions to mimic the real world and clinical practice setting. ccRCC tumors have a poorer prognosis than other types of renal tumors, such as papillary RCC or chRCC. Therefore, the preoperative differentiation of related tumors using ML-based radiomics interpretation could lead to differentiation of non-ccRCCs from ccRCCs,[74] and could achieve an accuracy of up to 89.9%.[34] Grading of tumors assessed by radiomics and AI studies can be affected by sarcomatoid features, but nonetheless studies developed to differentiate this aspect with poor accuracy of only 55%.[132] Due to recent advancements in genomics and radiomics, radiogenomics could improve the identification of aggressive tumors and hence establishing tailored treatment. The encountered genomic alterations in ccRCC could be heavily studied due to the rarity of mutations and the stability of sample and radiomics and AI models.[148] Studies showed good AUCs, specificity, sensitivity and accuracy, especially for RF algorithm but in the evaluation of biologic aggressiveness.[99,100] Many of the studies were not externally validated, AI systems will need further training and independent validation

to limit the risk of low generalizability.[149,150] DL provides good evidence that support further research for clinical use (ROI data set reported ACC of 0.97 while RBR had an accuracy of 0.93).[33] Despite the interesting results highlighted in this review, it is worth noting that due to some main issues – mostly related to the paucity of trials, the lack of homogeneity of data, especially with regard to ROI and end-points of these studies – evidence on the ability of radiomics in the evaluation of response to TKIs is premature for its integration into routine clinical practice at the present time. Summarizing the evidence, we can state that the imaging methods are not standardized, either CT or MRI, including all specifications of using different CT phasing, contrast enhanced or not, and also the classifiers used to discriminate between different radiomics features applied to analyze the gray zone levels of the studied renal lesions. The ML models will have to be better trained with vast amounts of data, and which is better to use will have to be determined by future studies. In depth analysis of the most appropriate AI method to be applied (or combinations) should be studied further. MRI seems to not have the ability to provide reliable evidence for the use of radiomics.

Conclusion

AI evidence so far indicates a strong association with improved sensitivity, specificity, accuracy in detecting and differentiating between renal lesions, and its algorithms that can adjust scanner settings to improve image acquisition (especially the gray zone levels) and standardization of scanner protocols between institutions will improve preoperative differentiation between benign, low-risk cancers and clinically significant renal cancers. Radiomics holds the premises to enhance the diagnostic ability of imaging tools to characterize renal lesions, but integration in clinical practice will have to be preceded by standardized radiomics models and methodology, and future prospective external validation of obtained data and their comparison with existing traditional, well-validated tools, will have to be performed prior to further integration in current practice.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Author contributions

Matteo Ferro: Conceptualization; Formal analysis; Methodology; Supervision; Validation; Visualization; Writing – original draft; Writing – review & editing.

Felice Crocetto: Methodology; Validation; Writing – original draft; Writing – review & editing.

Biagio Barone: Methodology; Supervision; Validation; Visualization; Writing – original draft; Writing – review & editing.

Francesco del Giudice: Methodology; Writing – original draft; Writing – review & editing.

Martina Maggi: Methodology; Validation; Visualization; Writing – original draft; Writing – review & editing.

Giuseppe Lucarelli: Supervision; Validation; Writing – review & editing.

Gian Maria Busetto: Supervision; Validation; Writing – review & editing.

Riccardo Autorino: Methodology; Supervision; Validation; Writing – review & editing.

Michele Marchioni: Methodology; Validation; Visualization; Writing – original draft; Writing – review & editing.

Francesco Cantiello: Validation; Writing – review & editing.

Fabio Crocerossa: Methodology; Validation; Writing – review & editing.

Stefano Luzzago: Supervision; Validation; Writing – review & editing.

Mattia Piccinelli: Validation; Writing – review & editing.

Francesco Alessandro Mistretta: Supervision; Validation; Writing – review & editing.

Marco Tozzi: Validation; Writing – review & editing.

Luigi Schips: Validation; Writing – review & editing.

Ugo Giovanni Falagario: Methodology; Validation; Visualization; Writing – original draft; Writing – review & editing.

Alessandro Veccia: Methodology; Validation; Writing – original draft; Writing – review & editing.

Mihai Dorin Vartolomei: Validation; Writing – review & editing.

Gennaro Musi: Supervision; Validation; Writing – review & editing.

Ottavio de Cobelli: Supervision; Validation; Writing – review & editing.

Emanuele Montanari: Supervision; Validation; Writing – original draft.

Octavian Sabin Tătaru: Conceptualization; Formal analysis; Methodology; Supervision; Validation; Visualization; Writing – original draft; Writing – review & editing.

Acknowledgements
None.

Funding
The authors received no financial support for the research, authorship, and/or publication of this article.

Competing interests
The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Availability of data and materials
Not applicable.

ORCID iDs
Matteo Ferro https://orcid.org/0000-0003-4687-7353

Biagio Barone https://orcid.org/0000-0003-4884-132X

Gian Maria Busetto https://orcid.org/0000-0002-7291-0316

Riccardo Autorino https://orcid.org/0000-0001-7045-7725

Fabio Crocerossa https://orcid.org/0000-0002-5114-2952

References

1. Siegel RL, Miller KD and Jemal A. Cancer statistics, 2019. *CA Cancer J Clin* 2019; 69: 7–34.
2. Liu B, Xiao Y, Li H, *et al*. Identification and verification of biomarker in clear cell renal cell carcinoma via bioinformatics and neural network model. *Biomed Res Int* 2020; 2020: 6954793.
3. Vartolomei L, Cotruş A, Stanciu C, *et al*. Quality of life and psychological distress among patients with small renal masses. *J Clin Med* 2022; 11: 3944.
4. Hollingsworth JM, Miller DC, Daignault S, *et al*. Rising incidence of small renal masses: a need to reassess treatment effect. *J Natl Cancer Inst* 2006; 98: 1331–1334.
5. Coy H, Hsieh K, Wu W, *et al*. Deep learning and radiomics: the utility of Google TensorFlowTM Inception in classifying clear cell renal cell carcinoma and oncocytoma on multiphasic CT. *Abdom Radiol N Y* 2019; 44: 2009–2020.
6. Ursprung S, Beer L, Bruining A, *et al*. Radiomics of computed tomography and magnetic resonance imaging in renal cell carcinoma – a systematic review and meta-analysis. *Eur Radiol* 2020; 30: 3558–3566.
7. Patel HD, Johnson MH, Pierorazio PM, *et al*. Diagnostic accuracy and risks of biopsy in the diagnosis of a renal mass suspicious for localized renal cell carcinoma: systematic review of the literature. *J Urol* 2016; 195: 1340–1347.
8. Di Lascio G, Sciarra A, Del Giudice F, *et al*. Which factors can influence post-operative renal function preservation after nephron-sparing surgery for kidney cancer: a critical review. *Cent European J Urol* 2022; 75: 14–27.
9. Sohlberg EM, Metzner TJ and Leppert JT. The harms of overdiagnosis and overtreatment in patients with small renal masses: a mini-review. *Eur Urol Focus* 2019; 5: 943–945.
10. Roussel E, Capitanio U, Kutikov A, *et al*. Novel imaging methods for renal mass characterization: a collaborative review. *Eur Urol* 2022; 81: 476–488.
11. Hindman N, Ngo L, Genega EM, *et al*. Angiomyolipoma with minimal fat: can it be differentiated from clear cell renal cell carcinoma by using standard MR techniques? *Radiology* 2012; 265: 468–477.
12. Kim JH, Li S, Khandwala Y, *et al*. National trends of preoperative imaging modalities before partial nephrectomy for renal masses in the U.S. from 2007-2015. *Can Urol Assoc J* 2019; 13: 89–94.
13. Sevcenco S, Heinz-Peer G, Ponhold L, *et al*. Utility and limitations of 3-Tesla diffusion-weighted magnetic resonance imaging for

differentiation of renal tumors. *Eur J Radiol* 2014; 83: 909–913.

14. Suarez-Ibarrola R, Basulto-Martinez M, Heinze A, *et al*. Radiomics applications in renal tumor assessment: a comprehensive review of the literature. *Cancers* 2020; 12: 1387.
15. Ball MW, Bezerra SM, Gorin MA, *et al*. Grade heterogeneity in small renal masses: potential implications for renal mass biopsy. *J Urol* 2015; 193: 36–40.
16. Chung BI, Leow JJ, Gelpi-Hammerschmidt F, *et al*. Racial disparities in postoperative complications after radical nephrectomy: a population-based analysis. *Urology* 2015; 85: 1411–1416.
17. Marconi L, Dabestani S, Lam TB, *et al*. Systematic review and meta-analysis of diagnostic accuracy of percutaneous renal tumour biopsy. *Eur Urol* 2016; 69: 660–673.
18. Chan H-P, Samala RK, Hadjiiski LM, *et al*. Deep learning in medical image analysis. *Adv Exp Med Biol* 2020; 1213: 3–21.
19. Kumar V, Gu Y, Basu S, *et al*. Radiomics: the process and the challenges. *Magn Reson Imaging* 2012; 30: 1234–1248.
20. Ferro M, de Cobelli O, Vartolomei MD, *et al*. Prostate cancer radiogenomics-from imaging to molecular characterization. *Int J Mol Sci* 2021; 22: 9971.
21. Busetto GM, Del Giudice F, Maggi M, *et al*. Prospective assessment of two-gene urinary test with multiparametric magnetic resonance imaging of the prostate for men undergoing primary prostate biopsy. *World J Urol* 2021; 39: 1869–1877.
22. Cerrato A, Bedia C, Capriotti AL, *et al*. Untargeted metabolomics of prostate cancer zwitterionic and positively charged compounds in urine. *Anal Chim Acta* 2021; 1158: 338381.
23. Ferro M, de Cobelli O, Musi G, *et al*. Radiomics in prostate cancer: an up-to-date review. *Ther Adv Urol* 2022; 14: 17562872221109020.
24. Lambin P, Rios-Velazquez E, Leijenaar R, *et al*. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer Oxf Engl* 2012; 48: 441–446.
25. Campi R, Stewart GD, Staehler M, *et al*. Novel liquid biomarkers and innovative imaging for kidney cancer diagnosis: what can be implemented in our practice today? A systematic review of the literature. *Eur Urol Oncol* 2021; 4: 22–41.
26. Rallis KS, Kleeman SO, Grant M, *et al*. Radiomics for renal cell carcinoma: predicting outcomes from immunotherapy and targeted therapies – a narrative review. *Eur Urol Focus* 2021; 7: 717–721.
27. Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009; 151: 264–269, W64.
28. Schieda N, Krishna S, Pedrosa I, *et al*. Active surveillance of renal masses: the role of radiology. *Radiology* 2022; 302: 11–24.
29. Yu H, Scalera J, Khalid M, *et al*. Texture analysis as a radiomic marker for differentiating renal tumors. *Abdom Radiol (NY)* 2017; 42: 2470–2478.
30. Coy H, Young JR, Douek ML, *et al*. Quantitative computer-aided diagnostic algorithm for automated detection of peak lesion attenuation in differentiating clear cell from papillary and chromophobe renal cell carcinoma, oncocytoma, and fat-poor angiomyolipoma on multiphasic multidetector computed tomography. *Abdom Radiol (NY)* 2017; 42: 1919–1928.
31. Kim JK, Park SY, Shon JH, *et al*. Angiomyolipoma with minimal fat: differentiation from renal cell carcinoma at biphasic helical CT. *Radiology* 2004; 230: 677–684.
32. Erdim C, Yardimci AH, Bektas CT, *et al*. Prediction of benign and malignant solid renal masses: machine learning-based CT texture analysis. *Acad Radiol* 2020; 27: 1422–1429.
33. Zhou L, Zhang Z, Chen YC, *et al*. A deep learning-based radiomics model for differentiating benign and malignant renal tumors. *Transl Oncol* 2019; 12: 292–300.
34. Sun XY, Feng QX, Xu X, *et al*. Radiologic-radiomic machine learning models for differentiation of benign and malignant solid renal masses: comparison with expert-level radiologists. *AJR Am J Roentgenol* 2020; 214: W44–W54.
35. Varghese BA, Chen F, Hwang DH, *et al*. Differentiation of predominantly solid enhancing lipid-poor renal cell masses by use of contrast-enhanced CT: evaluating the role of texture in tumor subtyping. *AJR Am J Roentgenol* 2018; 211: W288–W296.
36. Uhlig J, Biggemann L, Nietert MM, *et al*. Discriminating malignant and benign clinical T1 renal masses on computed tomography: a pragmatic radiomics and machine learning approach. *Medicine (Baltimore)* 2020; 99: 19725.

37. Nassiri N, Maas M, Cacciamani G, *et al.* A radiomic-based machine learning algorithm to reliably differentiate benign renal masses from renal cell carcinoma. *Eur Urol Focus* 2022; 8: 988–994.

38. Yap FY, Varghese BA, Cen SY, *et al.* Shape and texture-based radiomics signature on CT effectively discriminates benign from malignant renal masses. *Eur Radiol* 2021; 31: 1011–1021.

39. Xi IL, Zhao Y, Wang R, *et al.* Deep learning to distinguish benign from malignant renal lesions based on routine MR imaging. *Clin Cancer Res* 2020; 26: 1944–1952.

40. Said D, Hectors SJ, Wilck E, *et al.* Characterization of solid renal neoplasms using MRI-based quantitative radiomics features. *Abdom Radiol (NY)* 2020; 45: 2840–2850.

41. Xu Q, Zhu Q, Liu H, *et al.* Differentiating benign from malignant renal tumors using T2- and diffusion-weighted images: a comparison of deep learning and radiomics models versus assessment from radiologists. *J Magn Reson Imaging* 2022; 55: 1251–1259.

42. Massa'a RN, Stoeckl EM, Lubner MG, *et al.* Differentiation of benign from malignant solid renal lesions with MRI-based radiomics and machine learning. *Abdom Radiol (NY)* 2022; 47: 2896–2904.

43. Fittschen A, Wendlik I, Oeztuerk S, *et al.* Prevalence of sporadic renal angiomyolipoma: a retrospective analysis of 61,389 in- and out-patients. *Abdom Imaging* 2014; 39: 1009–1013.

44. Vos N and Oyen R. Renal angiomyolipoma: the good, the bad, and the ugly. *J Belg Soc Radiol* 2018; 102: 41.

45. Yang CW, Shen SH, Chang YH, *et al.* Are there useful CT features to differentiate renal cell carcinoma from lipid-poor renal angiomyolipoma? *AJR Am J Roentgenol* 2013; 201: 1017–1028.

46. Feng Z, Rong P, Cao P, *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* 2018; 28: 1625–1633.

47. Cui E-M, Lin F, Li Q, *et al.* Differentiation of renal angiomyolipoma without visible fat from renal cell carcinoma by machine learning based on whole-tumor computed tomography texture features. *Acta Radiol Stockh Swed* 2019; 60: 1543–1552.

48. Yang Y, Zhu J, Zhou Z, *et al.* An effective radiomics model for noninvasive discrimination of fat-poor angiomyolipoma from clear cell renal cell carcinoma. In: *2019 IEEE Symposium Series on Computational Intelligence (SSCI)*, Xiamen, China, 6–9 December 2019, pp.1551–1558. New York: IEEE.

49. Ma Y, Cao F, Xu X, *et al.* Can whole-tumor radiomics-based CT analysis better differentiate fat-poor angiomyolipoma from clear cell renal cell carcinoma: compared with conventional CT analysis? *Abdom Radiol (NY)* 2020; 45: 2500–2507.

50. Nie P, Yang G, Wang Z, *et al.* A CT-based radiomics nomogram for differentiation of renal angiomyolipoma without visible fat from homogeneous clear cell renal cell carcinoma. *Eur Radiol* 2020; 30: 1274–1284.

51. Yang R, Wu J, Sun L, *et al.* Radiomics of small renal masses on multiphasic CT: accuracy of machine learning–based classification models for the differentiation of renal cell carcinoma and angiomyolipoma without visible fat. *Eur Radiol* 2020; 30: 1254–1263.

52. Ma Y, Xu X, Pang P, *et al.* A CT-based tumoral and mini-peritumoral radiomics approach: differentiate fat-poor angiomyolipoma from clear cell renal cell carcinoma. *Cancer Manag Res* 2021; 13: 1417–1425.

53. Jian L, Liu Y, Xie Y, *et al.* MRI-based radiomics and urine creatinine for the differentiation of renal angiomyolipoma with minimal fat from renal cell carcinoma: a preliminary study. *Front Oncol* 2022; 12: 876664.

54. Ma Y, Ma W, Xu X, *et al.* A convention-radiomics CT nomogram for differentiating fat-poor angiomyolipoma from clear cell renal cell carcinoma. *Sci Rep* 2021; 11: 4644.

55. Han Z, Zhu Y, Xu J, *et al.* Predictive value of CT-based radiomics in distinguishing renal angiomyolipomas with minimal fat from other renal tumors. *Dis Markers* 2022; 2022: 9108129.

56. Kim TM, Ahn H, Lee HJ, *et al.* Differentiating renal epithelioid angiomyolipoma from clear cell carcinoma: using a radiomics model combined with CT imaging characteristics. *Abdom Radiol (NY)* 2022; 47: 2867–2880.

57. Razik A, Goyal A, Sharma R, *et al.* MR texture analysis in differentiating renal cell carcinoma from lipid-poor angiomyolipoma and oncocytoma. *Br J Radiol* 2020; 93: 20200569.

58. Matsumoto S, Arita Y, Yoshida S, *et al.* Utility of radiomics features of diffusion-weighted magnetic resonance imaging for differentiation of fat-poor angiomyolipoma from clear cell renal cell carcinoma: model development and

external validation. *Abdom Radiol (NY)* 2022; 47: 2178–2186.

59. Van der Kwast T and Perez-Ordoñez B. Renal oncocytoma, yet another tumour that does not fit in the dualistic benign/malignant paradigm? *J Clin Pathol* 2007; 60: 585–586.
60. Wu J, Zhu Q, Zhu W, *et al.* Comparative study of CT appearances in renal oncocytoma and chromophobe renal cell carcinoma. *Acta Radiol Stockh Swed* 2016; 57: 500–506.
61. Ishigami K, Jones AR, Dahmoush L, *et al.* Imaging spectrum of renal oncocytomas: a pictorial review with pathologic correlation. *Insights Imaging* 2015; 6: 53–64.
62. Baghdadi A, Aldhaam NA, Elsayed AS, *et al.* Automated differentiation of benign renal oncocytoma and chromophobe renal cell carcinoma on computed tomography using deep learning. *BJU Int* 2020; 125: 553–560.
63. Mühlbauer J, Egen L, Kowalewski K-F, *et al.* Radiomics in renal cell carcinoma – a systematic review and meta-analysis. *Cancers* 2021; 13: 1348.
64. Chen F, Gulati M, Hwang D, *et al.* Voxel-based whole-lesion enhancement parameters: a study of its clinical value in differentiating clear cell renal cell carcinoma from renal oncocytoma. *Abdom Radiol (NY)* 2017; 42: 552–560.
65. Deng Y, Soule E, Cui E, *et al.* Usefulness of CT texture analysis in differentiating benign and malignant renal tumours. *Clin Radiol* 2020; 75: 108–115.
66. Li Y, Huang X, Xia Y, *et al.* Value of radiomics in differential diagnosis of chromophobe renal cell carcinoma and renal oncocytoma. *Abdom Radiol (NY)* 2020; 45: 3193–3201.
67. Raman SP, Chen Y, Schroeder JL, *et al.* CT texture analysis of renal masses: pilot study using random forest classification for prediction of pathology. *Acad Radiol* 2014; 21: 1587–1596.
68. Sasaguri K, Takahashi N, Gomez-Cardona D, *et al.* Small (< 4 cm) Renal mass: differentiation of oncocytoma from renal cell carcinoma on biphasic contrast-enhanced CT. *AJR Am J Roentgenol* 2015; 205: 999–1007.
69. Varghese BA, Chen F, Hwang DH, *et al.* Differentiating solid, non-macroscopic fat containing, enhancing renal masses using fast Fourier transform analysis of multiphase CT. *Br J Radiol* 2018; 91: 20170789.
70. Hoang UN, Mojdeh Mirmomen S, Meirelles O, *et al.* Assessment of multiphasic contrast-enhanced MR textures in differentiating small renal mass subtypes. *Abdom Radiol (NY)* 2018; 43: 3400–3409.
71. Paschall AK, Mirmomen SM, Symons R, *et al.* Differentiating papillary type I RCC from clear cell RCC and oncocytoma: application of whole-lesion volumetric ADC measurement. *Abdom Radiol (NY)* 2018; 43: 2424–2430.
72. Gaing B, Sigmund EE, Huang WC, *et al.* Subtype differentiation of renal tumors using voxel-based histogram analysis of intravoxel incoherent motion parameters. *Invest Radiol* 2015; 50: 144–152.
73. Cheville JC, Lohse CM, Zincke H, *et al.* Comparisons of outcome and prognostic features among histologic subtypes of renal cell carcinoma. *Am J Surg Pathol* 2003; 27: 612–624.
74. 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. *Eur J Radiol* 2018; 107: 149–157.
75. Han S, Hwang SI and Lee HJ. The classification of renal cancer in 3-phase CT images using a deep learning method. *J Digit Imaging* 2019; 32: 638–643.
76. Li ZC, Zhai G, Zhang J, *et al.* Differentiation of clear cell and non-clear cell renal cell carcinomas by all-relevant radiomics features from multiphase CT: a VHL mutation perspective. *Eur Radiol* 2019; 29: 3996–4007.
77. Leng S, Takahashi N, Gomez Cardona D, *et al.* Subjective and objective heterogeneity scores for differentiating small renal masses using contrast-enhanced CT. *Abdom Radiol (NY)* 2017; 42: 1485–1492.
78. Yan L, Liu Z, Wang G, *et al.* Angiomyolipoma with minimal fat: differentiation from clear cell renal cell carcinoma and papillary renal cell carcinoma by texture analysis on CT images. *Acad Radiol* 2015; 22: 1115–1121.
79. Li A, Xing W, Li H, *et al.* Subtype differentiation of small (⩽ 4 cm) solid renal mass using volumetric histogram analysis of DWI at 3-T MRI. *AJR Am J Roentgenol* 2018; 211: 614–623.
80. Tabourin T, Pinar U, Parra J, *et al.* Impact of renal cell carcinoma histological variants on recurrence after partial nephrectomy: a multi-institutional, prospective study (UROCCR Study 82). *Ann Surg Oncol.* Epub ahead of print 3 July2022. DOI: 10.1245/s10434-022-12052-8.
81. Volpe A, Mattar K, Finelli A, *et al.* Contemporary results of percutaneous biopsy

of 100 small renal masses: a single center experience. *J Urol* 2008; 180: 2333–2337.

82. Gillies RJ, Kinahan PE and Hricak H. Radiomics: images are more than pictures, they are data. *Radiology* 2016; 278: 563–577.

83. Nazari M, Shiri I, Hajianfar G, *et al.* Noninvasive Fuhrman grading of clear cell renal cell carcinoma using computed tomography radiomic features and machine learning. *Radiol Med* 2020; 125: 754–762.

84. Shu J, Tang Y, Cui J, *et al.* Clear cell renal cell carcinoma: CT-based radiomics features for the prediction of Fuhrman grade. *Eur J Radiol* 2018; 109: 8–12.

85. Deng Y, Soule E, Samuel A, *et al.* CT texture analysis in the differentiation of major renal cell carcinoma subtypes and correlation with Fuhrman grade. *Eur Radiol* 2019; 29: 6922–6929.

86. Goyal A, Razik A, Kandasamy D, *et al.* Role of MR texture analysis in histological subtyping and grading of renal cell carcinoma: a preliminary study. *Abdom Radiol (NY)* 2019; 44: 3336–3349.

87. Stanzione A, Ricciardi C, Cuocolo R, *et al.* MRI radiomics for the prediction of Fuhrman grade in clear cell renal cell carcinoma: a machine learning exploratory study. *J Digit Imaging* 2020; 33: 879–887.

88. Yin RH, Yang YC, Tang XQ, *et al.* Enhanced computed tomography radiomics-based machine-learning methods for predicting the Fuhrman grades of renal clear cell carcinoma. *J Xray Sci Technol* 2021; 29: 1149–1160.

89. Rickman J, Struyk G, Simpson B, *et al.* The growing role for semantic segmentation in urology. *Eur Urol Focus* 2021; 7: 692–695.

90. Sun MR, Ngo L, Genega EM, *et al.* Renal cell carcinoma: dynamic contrast-enhanced MR imaging for differentiation of tumor subtypes – correlation with pathologic findings. *Radiology* 2009; 250: 793–802.

91. Tian K, Rubadue CA, Lin DI, *et al.* Automated clear cell renal carcinoma grade classification with prognostic significance. *PLoS ONE* 2019; 14: 0222641.

92. Jansen RW, van Amstel P, Martens RM, *et al.* Non-invasive tumor genotyping using radiogenomic biomarkers, a systematic review and oncology-wide pathway analysis. *Oncotarget* 2018; 9: 20134–20155.

93. Mazurowski MA. Radiogenomics: what it is and why it is important. *J Am Coll Radiol* 2015; 12: 862–866.

94. Bai HX, Lee AM, Yang L, *et al.* Imaging genomics in cancer research: limitations and promises. *Br J Radiol* 2016; 89: 20151030.

95. Sala E, Mema E, Himoto Y, *et al.* Unravelling tumour heterogeneity using next-generation imaging: radiomics, radiogenomics, and habitat imaging. *Clin Radiol* 2017; 72: 3–10.

96. Alessandrino F, Krajewski KM and Shinagare AB. Update on radiogenomics of clear cell renal cell carcinoma. *Eur Urol Focus* 2016; 2: 572–573.

97. Lee HW, Cho H-H, Joung J-G, *et al.* Integrative radiogenomics approach for risk assessment of post-operative metastasis in pathological T1 renal cell carcinoma: a pilot retrospective cohort study. *Cancers* 2020; 12: 866.

98. Feng Z, Zhang L, Qi Z, *et al.* Identifying BAP1 mutations in clear-cell renal cell carcinoma by CT radiomics: preliminary findings. *Front Oncol* 2020; 10: 279.

99. Kocak B, Durmaz ES, Kaya OK, *et al.* Machine learning-based unenhanced CT texture analysis for predicting BAP1 mutation status of clear cell renal cell carcinomas. *Acta Radiol Stockh Swed* 2020; 61: 856–864.

100. Kocak B, Durmaz ES, Ates E, *et al.* 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* 2019; 212: W55–W63.

101. Chen X, Zhou Z, Hannan R, *et al.* Reliable gene mutation prediction in clear cell renal cell carcinoma through multi-classifier multi-objective radiogenomics model. *Phys Med Biol* 2018; 63: 215008.

102. Ghosh P, Tamboli P, Vikram R, *et al.* Imaging-genomic pipeline for identifying gene mutations using three-dimensional intra-tumor heterogeneity features. *J Med Imaging (Bellingham)* 2015; 2: 041009.

103. Powles T, Plimack ER, Soulières D, *et al.* Pembrolizumab plus axitinib versus sunitinib monotherapy as first-line treatment of advanced renal cell carcinoma (KEYNOTE-426): extended follow-up from a randomised, open-label, phase 3 trial. *Lancet Oncol* 2020; 21: 1563–1573.

104. Motzer RJ, Banchereau R, Hamidi H, *et al.* Molecular subsets in renal cancer determine outcome to checkpoint and angiogenesis blockade. *Cancer Cell* 2020; 38: 803–817.4.

105. Antunes J, Viswanath S, Rusu M, *et al.* Radiomics analysis on FLT-PET/MRI for

characterization of early treatment response in renal cell carcinoma: a proof-of-concept study. *Transl Oncol* 2016; 9: 155–162.

106. Bharwani N, Miquel ME, Powles T, *et al*. Diffusion-weighted and multiphase contrast-enhanced MRI as surrogate markers of response to neoadjuvant sunitinib in metastatic renal cell carcinoma. *Br J Cancer* 2014; 110: 616–624.

107. Boos J, Revah G, Brook OR, *et al*. CT intensity distribution curve (histogram) analysis of patients undergoing antiangiogenic therapy for metastatic renal cell carcinoma. *AJR Am J Roentgenol* 2017; 209: W85–W92.

108. Goh V, Ganeshan B, Nathan P, *et al*. Assessment of response to tyrosine kinase inhibitors in metastatic renal cell cancer: CT texture as a predictive biomarker. *Radiology* 2011; 261: 165–171.

109. Haider MA, Vosough A, Khalvati F, *et al*. CT texture analysis: a potential tool for prediction of survival in patients with metastatic clear cell carcinoma treated with sunitinib. *Cancer Imaging* 2017; 17: 4.

110. Mains JR, Donskov F, Pedersen EM, *et al*. Use of patient outcome endpoints to identify the best functional CT imaging parameters in metastatic renal cell carcinoma patients. *Br J Radiol* 2018; 91: 20160795.

111. Khene ZE, Mathieu R, Peyronnet B, *et al*. Radiomics can predict tumour response in patients treated with Nivolumab for a metastatic renal cell carcinoma: an artificial intelligence concept. *World J Urol* 2021; 39: 3707–3709.

112. UNC Lineberger Comprehensive Cancer Center. Quantitative PET-MRI imaging correlated with transcriptome analysis for noninvasive characterization of renal cell carcinomas. Clinical Trial Registration NCT04271254. https://clinicaltrials.gov/ct2/show/NCT04271254 (26 October 2021, accessed 28 August 2022).

113. Jonsson Comprehensive Cancer Center. SPECT/CT for the characterization of renal masses: impact on clinical decision making. Clinical Trial Registration NCT03996850. https://clinicaltrials.gov/ct2/show/NCT03996850 (21 July 2022, accessed 28 August 2022).

114. Geertsen L. KIDSTAGE- staging of kidney cancer using dual time PET/CT and other biomarkers. Clinical Trial Registration NCT04295174. https://clinicaltrials.gov/ct2/show/NCT04295174 (3 March 2020, accessed 28 August 2022).

115. University Hospital Bordeaux. Diagnostic value of multiparametric MR imaging of small solid renal tumors (IRMK01). Clinical Trial Registration NCT03470285. https://clinicaltrials.gov/ct2/show/NCT03470285 (16 August 2021, accessed 28 August 2022).

116. Pilot study of renal neoplasms with perfusion magnetic resonance imaging. Clinical Trial Registration NCT02526511. https://clinicaltrials.gov/ct2/show/NCT02526511 (14 September 2021, accessed 28 August 2022).

117. Swensson JK. Correlation of contrast enhanced ultrasound of renal masses with pathologic grade: a prospective comparison of quantitative and qualitative findings. Clinical Trial Registration NCT03821376. https://clinicaltrials.gov/ct2/show/NCT03821376 (12 July 2022, accessed 28 August 2022).

118. Moschovakis YN. What is an algorithm? In: Engquist B and Schmid W (eds) *Math unlimited—2001 beyond*. Cham: Springer, 2001, pp. 919–936.

119. Sarker IH. Machine learning: algorithms, real-world applications and research directions. *SN Comput Sci* 2021; 2: 160.

120. Sarker IH, Kayes ASM, Badsha S, *et al*. Cybersecurity data science: an overview from machine learning perspective. *J Big Data* 2020; 7: 1–29.

121. Han J, Pei J and Tong H. *Data mining: concepts and techniques*. Burlington, MA: Morgan Kaufmann, 2022.

122. Giulietti M, Cecati M, Sabanovic B, *et al*. The role of artificial intelligence in the diagnosis and prognosis of renal cell tumors. *Diagn Basel Switz* 2021; 11: 206.

123. Huang MW, Chen CW, Lin WC, *et al*. SVM and SVM ensembles in breast cancer prediction. *PLoS ONE* 2017; 12: 0161501.

124. Tătaru OS, Vartolomei MD, Rassweiler JJ, *et al*. Artificial intelligence and machine learning in prostate cancer patient management-current trends and future perspectives. *Diagn Basel Switz* 2021; 11: 354.

125. Supriya M and Deepa AJ. A novel approach for breast cancer prediction using optimized ANN classifier based on big data environment. *Health Care Manag Sci* 2020; 23: 414–426.

126. Wongvibulsin S, Wu KC and Zeger SL. Clinical risk prediction with random forests for survival, longitudinal, and multivariate (RF-SLAM) data analysis. *BMC Med Res Methodol* 2019; 20: 1.

127. Caruana R and Niculescu-Mizil A. *Proceedings of the 23rd international conference on Machine*

learning. New York: Association for Computing Machinery, 2006.

128. Raita Y, Goto T, Faridi MK, *et al*. Emergency department triage prediction of clinical outcomes using machine learning models. *Crit Care Lond Engl* 2019; 23: 64.

129. Shu J, Wen D, Xi Y, *et al*. Clear cell renal cell carcinoma: machine learning-based computed tomography radiomics analysis for the prediction of WHO/ISUP grade. *Eur J Radiol* 2019; 121: 108738.

130. Kunapuli G, Varghese BA, Ganapathy P, *et al*. A decision-support tool for renal mass classification. *J Digit Imaging* 2018; 31: 929–939.

131. Lee HS, Hong H, Jung DC, *et al*. Differentiation of fat-poor angiomyolipoma from clear cell renal cell carcinoma in contrast-enhanced MDCT images using quantitative feature classification. *Med Phys* 2017; 44: 3604–3614.

132. Schieda N, Thornhill RE, Al-Subhi M, *et al*. Diagnosis of sarcomatoid renal cell carcinoma with CT: evaluation by qualitative imaging features and texture analysis. *AJR Am J Roentgenol* 2015; 204: 1013–1023.

133. Le E. *Radiomics and machine learning in the prediction of cardiovascular disease*. Thesis, University of Cambridge, Cambridge, 2021.

134. Lambin P, Leijenaar RTH, Deist TM, *et al*. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* 2017; 14: 749–762.

135. Kocak B, Kaya OK, Erdim C, *et al*. Artificial intelligence in renal mass characterization: a systematic review of methodologic items related to modeling, performance evaluation, clinical utility, and transparency. *AJR Am J Roentgenol* 2020; 215: 1113–1122.

136. Haury AC, Gestraud P and Vert JP. The influence of feature selection methods on accuracy, stability and interpretability of molecular signatures. *PLoS ONE* 2011; 6: 28210.

137. Zwanenburg A and Löck S. Why validation of prognostic models matters? *Radiother Oncol* 2018; 127: 370–373.

138. Chen G, Dai Y, Zhang J, *et al*. MBANet: multi-branch aware network for kidney ultrasound images segmentation. *Comput Biol Med* 2022; 141: 105140.

139. Song SH, Han JH, Kim KS, *et al*. Deep-learning segmentation of ultrasound images for automated calculation of the hydronephrosis area to renal parenchyma ratio. *Investig Clin Urol* 2022; 63: 455–463.

140. Korfiatis P, Denic A, Edwards ME, *et al*. Automated segmentation of kidney cortex and medulla in CT images: a multisite evaluation study. *J Am Soc Nephrol* 2022; 33: 420–430.

141. Hsiao CH, Sun TL, Lin PC, *et al*. A deep learning-based precision volume calculation approach for kidney and tumor segmentation on computed tomography images. *Comput Methods Programs Biomed* 2022; 221: 106861.

142. Houshyar R, Glavis-Bloom J, Bui TL, *et al*. Outcomes of artificial intelligence volumetric assessment of kidneys and renal tumors for preoperative assessment of nephron-sparing interventions. *J Endourol* 2021; 35: 1411–1418.

143. Waite S, Grigorian A, Alexander RG, *et al*. Analysis of perceptual expertise in radiology – current knowledge and a new perspective. *Front Hum Neurosci* 2019; 13: 213.

144. Kowalewski KF, Egen L, Fischetti CE, *et al*. Artificial intelligence for renal cancer: from imaging to histology and beyond. *Asian J Urol* 2022; 9: 243–252.

145. Toda N, Hashimoto M, Arita Y, *et al*. Deep learning algorithm for fully automated detection of small (⩽4 cm) renal cell carcinoma in contrast-enhanced computed tomography using a multicenter database. *Invest Radiol* 2022; 57: 327–333.

146. Kart T, Fischer M, Küstner T, *et al*. Deep learning-based automated abdominal organ segmentation in the UK Biobank and German National Cohort Magnetic Resonance Imaging Studies. *Invest Radiol* 2021; 56: 401–408.

147. Zhao Y, Peng C, Wang S, *et al*. The feasibility investigation of AI -assisted compressed sensing in kidney MR imaging: an ultra-fast T2WI imaging technology. *BMC Med Imaging* 2022; 22: 119.

148. Alessandrino F, Shinagare AB, Bossé D, *et al*. Radiogenomics in renal cell carcinoma. *Abdom Radiol N Y* 2019; 44: 1990–1998.

149. Kocak B, Durmaz ES, Erdim C, *et al*. Radiomics of renal masses: systematic review of reproducibility and validation strategies. *AJR Am J Roentgenol* 2020; 214: 129–136.

150. Ravi D, Wong C, Deligianni F, *et al*. Deep learning for health informatics. *IEEE J Biomed Health Inform* 2017; 21: 4–21.