Review

Artificial Intelligence in Renal Cell Carcinoma Histopathology: Current Applications and Future Perspectives

Alfredo Distante [1,12], Laura Marandino [2], Riccardo Bertolo [3], Alexandre Ingels [4], Nicola Pavan [5], Angela Pecoraro [6], Michele Marchioni [7], Umberto Carbonara [8], Selcuk Erdem [9], Daniele Amparore [6], Riccardo Campi [10], Eduard Roussel [11], Anna Caliò [12], Zhenjie Wu [13], Carlotta Palumbo [14], Leonardo D. Borregales [15], Peter Mulders [16] and Constantijn H.J. Muselaers [16]*, on behalf of the EAU Young Academic Urologists (YAU) Renal Cancer working group.

1. Department of Urology, Fondazione Policlinico Universitario A. Gemelli IRCCS, Università Cattolica del Sacro Cuore, 00168 Roma, Italy.
2. Department of Medical Oncology, IRCCS Ospedale San Raffaele, Milan, Italy.
3. Department of Urology, San Carlo Di Nancy Hospital, Rome, Italy.
4. Department of Urology, University Hospital Henri Mondor, APHP, Créteil, France.
5. Department of Surgical, Oncological and Oral Sciences, Section of Urology, University of Palermo, Palermo, Italy
6. Department of Urology, San Luigi Gonzaga Hospital, University of Turin, Orbassano, Turin, Italy.
7. Department of Medical, Oral and Biotechnological Sciences, G. d'Annunzio University of Chieti, Chieti, Italy.
8. Department of Emergency and Organ Transplantation-Urology, Andrology and Kidney Transplantation Unit, University of Bari, Bari, Italy.
9. Division of Urologic Oncology, Department of Urology, Istanbul University Istanbul Faculty of Medicine, Istanbul, Turkey
10. Urological Robotic Surgery and Renal Transplantation Unit, Careggi Hospital, University of Florence, Italy.
11. Department of Urology, University Hospitals Leuven, Leuven, Belgium.
12. Department of Diagnostic and Public Health, Section of Pathology, University of Verona, Verona 37134, Italy.
13. Department of Urology, Changhai Hospital, Naval Medical University, Shanghai, China.
14. Division of Urology, Maggiore della Carità Hospital of Novara, Department of Translational Medicine, University of Eastern Piedmont, Novara, Italy
15. Department of Urology, Well Cornell Medicine, New York-Presbyterian Hospital, NY, USA
16. Department of Urology, Radboud University Medical Center, Nijmegen, The Netherlands.

* Correspondence: A. Distante, Department of Urology, Radboud University Medical Center, Geert Grooteplein 10, 6525 GA, Nijmegen, The Netherlands. Phone: +31243613735. E-mail: a.distante1@gmail.com

Keywords: Artificial intelligence; pathology; renal cell carcinoma; kidney cancer

1. INTRODUCTION

Renal cell carcinoma (RCC) is among the top 10 most common cancers in both men and women. The incidence of RCC is gradually rising over the years, which results in increased demands on healthcare systems in time, effort, and cost [1]. Adequate diagnosis and treatment planning of RCC relies on adequate clinical data, imaging, histology, and molecular profiling [2,3].

Histological analysis, supported by genetic and cytogenetic analysis, is crucial for RCC diagnosis, subtyping, and defining features with high prognostic and therapeutic impact [4,5]. These features include tumor grade, RCC subtype, lymphovascular invasion, tumor necrosis, sarcomatoid dedifferentiation, and others [6–8]. RCC histological diagnosis and classification in particular, can be a daunting task, as it encompasses a broad spectrum of histopathological entities, which have recently been subject to changes [9].

Over the years, the daily clinical practice of treating patients with RCC has changed from paper charts, analogue radiographs, and light microscopes to their more modern

 doi:10.20944/preprints202306.0221.v1

counterparts, such as electronic health records and digitalized radiology and virtual pathology. This has resulted in an enormous amount of digital data, which can be utilized by data-characterization algorithms or artificial intelligence (AI) [10,11].

Machine learning (ML) is a subfield of AI. It employs algorithms that enable computers to learn from digital images of tissue samples. In histopathology, it can be utilized in a number of ways. These consist in digital analysis of images of tissue samples, identification of different structures or cell types, and classification or segmentation of different regions in the tissue sample [12]. The capabilities of ML have increased with the development of deep learning (DL), a section of ML focused on creating a virtual neural network with multiple layers, inspired by how biological neurons communicate [13]. DL models are well-suited for feature extraction and learning from data because they can automatically identify complex patterns and relationships within large and diverse datasets, such as those used in cancer diagnostics.

The choice of the best algorithm for AI applications in histopathology is still difficult. There are three primary types of learning: supervised learning, which uses labeled data for training; unsupervised learning, which finds patterns without labels; and weakly supervised learning, which strikes a medium ground by using partially labeled data.

AI in radiology, also known as radiomics, has shown excellent diagnostic accuracy for detecting RCC and can even provide information regarding RCC subtyping, nuclear grade prediction, gene mutations, and gene expression–based molecular signatures [14]. In line with AI in radiology, efforts to use AI in RCC histopathology have been undertaken in recent years. This relatively new field, called pathomics or computational pathology, can be used to improve efficiency, accessibility, cost-effectiveness, and time consumption and enhance accuracy and reproducibility with lower subjectivity [10,15–17]. In addition, Whole Slide Imaging (WSI) technology allows machine learning in pathology by providing an enormous amount of high-quality information for training and testing AI models to identify specific features and patterns that can even be complex for the human eye to discern [11,18,19]. Ultimately, AI aims to assist pathologists in making more accurate and consistent diagnoses in a shorter time and is a valuable implement to undercover the information cited above [20,21].

In this literature review, we aim to provide an overview of the current evidence regarding the use of computational pathology in histopathology in RCC.

2. EVIDENCE ACQUISITION

We conducted a narrative review of the literature concerning all the possible applications of AI in the histopathological analysis of RCC specimens.

The Medline database was screened, and literature research was restricted to articles published in English between January 1st, 2017, and January 1st, 2023, since most of the relevant literature in this field has been published in this timeframe.

We used a structured search strategy (Supplementary material), obtaining 98 results that were reviewed, and references to the retrieved articles were hand-searched to identify additional reports that met the scope of this review.

Original studies and case series were selected for inclusion, while reviews, editorials, and letters to the editor were excluded. Finally, references to the retrieved articles were hand-searched to identify additional reports that met the scope of this review.

The titles and abstracts of all papers included were independently assessed against the inclusion and exclusion criteria using Rayyan (Rayyan Systems, Cambridge, MA, USA).

3. ARTIFICIAL INTELLIGENCE AIDED DIAGNOSIS OF RCC SUBTYPES

Although several advances have been made in RCC diagnostics in the last decade, especially in imaging techniques, histopathological diagnosis based on a pathologist's eye and experience remains the current clinical practice in distinguishing RCC from normal renal tissue on the microscopic level [14,24–26].

 doi:10.20944/preprints202306.0221.v1

However, RCCs can have complicated characteristics that make the diagnosis difficult, laborious, and time-consuming, even for experienced pathologists. This is known to lead to a moderate inter-reader agreement for the RCC subtype [27–29]. In addition, several studies demonstrated how computational pathology could be a solution to more uniform specimen readings and reduce intra and inter-observer variability [30–32].

3.1. RCC diagnosis and subtyping in biopsy specimens

RCC varies in its biological behavior, ranging from indolent to aggressive tumors. Currently, no reliable predictive models to distinguish among different clinical types are available to be used in the preoperative setting, creating concerns about under and overtreatment, especially in small renal masses (SRMs), which now represent up to 50% of renal lesions [33–37]. Therefore this can lead to overdiagnosis and overtreament, as up to date there are no highly reliable biomarkers or imaging methods that can correctly differentiate beningn from malignant lesions [38–40] As a result, there has been a growing trend in using renal mass biopsy (RMB) over the past decade to address this challenge [41,42].

However, RMBs have some limitations as they are nondiagnostic in approximately 10-15% of the cases and remain intrinsically invasive [43]. The main reason for the high percentage of nondiagnostic results is an inadequate sampling of tumors [44]. Another crucial issue in RMB is a fair degree of interobserver variability [45], a concern also found in breast, prostate, and melanoma biopsies [46–48].

To tackle these problems, Fenstermaker et al. developed a DL-based algorithm for RCC diagnosis, grading, and subtype assessment [49]. Their method reached a high accuracy level when using only a 100 square micrometers (μm2) patch, making it a potentially valuable tool in RMB analysis. In addition, although their method has been trained on whole-mount surgical specimens, a computational method trained and tested on small tissue samples may reduce the need for repeat biopsies by decreasing insufficient tissue sampling and reducing interobserver variability.

However, this study focused on identifying the three main subtypes of RCC without considering benign tumors or oncocytomas. A significant proportion of small renal masses (SRMs) are benign, with oncocytoma being the most frequent benign, contrast-enhancing renal mass found. A well-known problem for pathologists is differentiating oncocytomas from chromophobe RCC [50–52]. Zhu et al. reported favorable results in RCC subtyping in surgical resection and RMB specimens, promising results in oncocytoma diagnosis in RMB [53]. The group trained and tested a model on an internal dataset of renal resections. In addition, they tested this model on 79 RCC biopsy slides, 24 of which were diagnosed as renal oncocytoma, and also on an external dataset, achieving good performance, as shown in Table 1.

3.2. RCC diagnosis and subtyping in surgical resection specimens

Despite the recent increased use of RMB and the enormous advancement in diagnostic accuracy [54,55], approximately 73% of surveyed urologists would not perform a RMB for various reasons [56]. Currently, the standard of treatment for non-metastatic RCC is surgical resection, either with a radical or partial nephrectomy, and in some selected cases of metastatic RCC [57]. However, examining and analyzing the complex histological patterns of RCC surgical resection specimens under a microscope can be challenging and time-consuming for pathologists for many reasons. For instance, nephrectomy specimens exhibit substantial heterogeneity, exemplifying the wide variation observed within RCC surgical resection samples [58]. Moreover, variability among different observers and even within the same observer has been reported [28].

Good results were obtained by Tabibu et al. in distinguishing ccRCC and chRCC from normal tissue by using two pre-trained convolutional neural networks (CNN) and replacing the last layers with two output layers and fine-tuned it on RCC data [59]. Moreover, for subtype classification, the group introduced a so-called Directed Acyclic Graph Support Vector Machine (DAG-SVM) on the top of the deep network obtaining good

 doi:10.20944/preprints202306.0221.v1

accuracy in this task. Unlike Tabibu et al. model, Chen et al. developed a DL algorithm to detect RCC that was externally validated on an independent dataset [60]. To accomplish this task, they used LASSO (Least Absolute Shrinkage and Selection Operator), a method used in ML to select from a more extensive set of features, the most important in predicting outcomes. Through LASSO analysis, they identified various image features based on the "The Cancer Genome Atlas" (TCGA) cohort to distinguish ccRCC from normal renal parenchyma and ccRCC from pRCC and chRCC, obtaining high accuracy in test and external validation cohorts.

Also, Marostica et al. created a pipeline using transfer learning to identify cancerous regions from slide images and classify the three major subtypes obtaining good performance in the test set and two external independent datasets (Table 3) [61].

RCC classification is a challenging task not only due to the complexity of the procedure itself but also because the classification system is subject to periodic updates [62,63]. For example, only in recent years has clear cell papillary renal cell carcinoma (ccpRCC) been recognized as a specific entity [64]. This subtype of RCC histologically resembles both ccRCC and pRCC and has clear cell changes. However, ccpRCC has distinct immunohistochemical and genetic profiles compared to ccRCC and pRCC [65]. It also carries a favorable prognosis compared to the latter; therefore, the last World Health Organization changed the denomination to clear cell papillary renal cell tumor [66]. Abdeltawab et al. developed a computational model that could classify between ccRCC and ccpRCC, obtaining an accuracy of 91% on the institution files in identifying ccpRCC and 90% in diagnosing ccRCC on an external dataset [67].

The abovementioned studies are mainly supervised and highly defined for RCC approaches, making them time-consuming. However, the capability to apply knowledge gained from previous experiences to novel situations is a vital skill for human beings. As an example, pathologists can use lessons learned outside their specific subspecialty because several cancer types exhibit common hallmarks of malignancy, as demonstrated by Faust et al., who attested whether a previously trained AI developed for recognizing brain tumor features could be applied to cluster and analyze RCC specimens in an unsupervised fashion [68]. The results showed that grouping cancer regions from non-neoplastic tissue elements matched expert annotations in multiple randomly selected cases. This hypothetically represents a way to demonstrate that unsupervised ML-based methods, built for other cancers' diagnosis, can also be used for RCC, reducing developing time and work amount.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 June 2023 doi:10.20944/preprints202306.0221.v1

Table 1. Overview of studies on AI models for diagnosis and subtyping.

Group	Aim	Number of patients	Accuracy on the test set	External validation (N of patients)	Accuracy on the external validation cohort	Algorithm
Fenstermaker et al. [49]	1) RCC diagnosis, 2) subtyping, 3) grading	15 ccRCC 15 pRCC 12 chRCC	1) 99.1%; 2) 97.5%; 3) 98.4%	N.A.	N.A.	CNN
Zhu et al. [53]	RCC subtyping	486 SR (30 NT, 27 RO, 38 chRCC, 310 ccRCC, 81 pRCC), 79 RMB (24 RO, 34 ccRCC, 21 pRCC)	1) 97% on SRS, 2) 97% on RMB	0 RO 109 ChRCC 505 ccRCC 294 pRCC:	95% accuracy *(only SRs)*	DNN
Chen et al. [60]	1) RCC diagnosis, 2) subtyping, 3) survival prediction	1) & 2) 362 NT, 362ccRCC, 128pRCC, 84chRCC 3) 283ccRCC	1) 94.5% vs. NT 2) 97% vs. pRCC and chRCC 3) 88.8%, 90.0%, 89.6% in 1-3-5 y DFS	1) & 2) 150 NP 150 ccRCC 52 pRCC 84 chRCC 3) 120ccRCC	1)87.6% vs. NP 2)81.4% vs. pRCC and chRCC 3) 72.0%, 80.9%, 85.9% in 1-3-5 y DFS	CNN
Tabibu et al. [59]	1) RCC diagnosis, 2) subtyping,	509 NT 1027 ccRCC 303 pRCC 254 chRCC	1)93.9% ccRCC vs. NP 87.34% chRCC vs. NP 2)92.16% subtyping	N.A.	N.A.	CNN (Resnet 18 and 34 architecture based); DAG-SVM on top of CNN for subtyping
Abdeltawab et al. [67]	RCC subtyping	27 ccRCC 14 ccpRCC	91% in ccpRCC	10 ccRCC.	90% in ccRCC	CNN

ccRCC=clear cell renal cell carcinoma, ccpRCC=clear cell papillary renal cell carcinoma, chRCC=chromophobe renal cell carcinoma, CNN=convolutional neural network, DFS=disease-free-survival, DNN=deep neural network, N.A.=not applicable, NT=normal tissue, pRCC=papillary renal cell carcinoma, RMB=renal mass biopsy, SR=surgical resection,.

4. PATHOMICS IN DISEASE PROGNOSIS

The prognosis for RCC depends on several factors, including anatomical and clinical factors, but also histological and molecular factors play important prognostic roles in both non-metastatic and mRCC [69].

4.1. Cancer grading

Tumor grading is considered one of the most critical factors in prognosis prediction, as the 5-year survival rate for patients with low-grade RCC is around 90%, while in high-grade RCC is about 12% [69–71].

Although largely replaced by the WHO/ISUP grading classification, the Fuhrman grade still plays an independent factor in determining a higher risk of recurrence and a lower chance of survival [72–76]. The Fuhrman grading system predominantly focuses on the morphology of the nucleus (size and shape) and the existence of prominent nucleoli, but inter- and intra-observer variability represents an issue [77–79]. Yeh et al. trained a support vector machine (SVM) classifier that performed well in identifying, size-estimating, calculating spatial distribution, and distinguishing low vs. high grades on ccRCC

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 June 2023 doi:10.20944/preprints202306.0221.v1

specimens [80]. However, it couldn't differentiate between specific grades (e.g., III and IV), and no analyses of patients' survival were presented.

Unlike the Fuhrman grading, WHO/ISUP system relies solely on nucleolar prominence for grade 1-3 tumors, allowing for less inter-observer variation [81]. Therefore, Holdbrook et al. developed a model that detected prominent nucleoli and quantified nuclear pleomorphic patterns by concatenating features (i.e., combining different features (or variables) into a single input representation for the model) extracted from prominent nucleoli and classifying them as either high or low-grade [82]. The model also showed excellent grade classification accuracy and prognosis prediction by comparing these results to a multigene score.

The beforementioned computational systems are unique in many features like image processing, feature extraction, classification method, and predicting 2 -tiered grades (which demonstrated to perform well in cancer-specific-survival (CSS) prediction). [83]. Tian et al. used 395 ccRCC cases from the TCGA dataset reviewed by a pathologist and stratified in the 2-tiered system: low or high-grade [84]. Of these, 277 had concordance between the TCGA and the pathologist's assigned grade and were used to train the model by extracting different histomic features for each patch. They used LASSO regression to select the most associated with grading ones, obtaining a model that predicted 2-tiered ccRCC grading with good agreement with manual grades. It also showed a significant association between predicted grade and overall survival, even when adjusting for age and gender. Furthermore, the model's predicted grade was better for overall survival prediction than TCGA and pathologist grade in discordant cases. This study was different from Yeh et al. [80] who only evaluated one feature (i.e., maximum nuclei size) to predict 2-tiered grade, and Holdbrook et al. [82] who used up to 4 concatenate feature vectors to calculate F-scores before classification into low or high grade. The features in the model of Holdbrook et al. [82] are unspecified.

In addition, Tian et al. and Holdbrook et al. showed that the predicted grade had prognostic value, whereas Yeh et al. did not report any association between their grade and prognosis.

Tian et al. study used a conventional image-analysis technique for nuclei segmentation. However, DL-based techniques of nuclei segmentation might be a solution like Yeh et al. and Song et al. method for this task [80,85]. The results of the studies above are summarized in Table 2.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 June 2023 doi:10.20944/preprints202306.0221.v1

Table 2. Overview of studies on AI models for RCC grading.

Group	Aim	Number of patients	Accuracy on the test set	External validation (N of patients)	Accuracy on the external validation cohort	Algorithm
Yeh et al. [80]	RCC grading	39 ccRCC	AUC: 0.97	N.A.	N.A.	SVM
Holdbrook et al. [82]	1) RCC grading, 2) survival prediction	59 ccRCC	1) F-score: 0.78 – 0.83 grade prediction 2) High degree of correlation (R = 0.59) with a multigene score	N.A.	N.A.	DNN – features concatenation
Tian et al. [84]	1) RCC grading, 2) survival prediction	395 ccRCC	1) 84.6% sensitivity and 81.3% specificity grade prediction 2) predicted grade associated with overall survival (HR: 2.05; 95% CI 1.21-3.47)	N.A.	N.A.	DNN - LASSO model

AUC= area under curve, ccRCC= clear cell renal cell carcinoma, DNN= deep neural network, N.A.= not applicable, SVM= support vector machine.

4.2. Molecular-morphological Connections and AI-based therapy response prediction

Recent developments in predicting RCC survival have suggested molecular differences within subtypes that affect prognosis, as well as potentially predictive molecular biomarkers and marker signatures, even though there is no definitive evidence to date supporting the routine clinical use of biomarkers for treatment selection in metastatic RCC (mRCC) [86–91].

As the finding of predictive biomarkers still represents an unmet clinical need, AI can be used to explore connections between molecular biomarkers and morphological features on histopathology images, overcoming traditional biomarker analysis limitations such as the high cost (both financially and in terms of time), limited sample size, and lack of standardization [92–95].

Among the many possible genetic aberrations in RCC, one crucial type of mutation is copy number alterations (CNAs) associated with RCC's development, treatment response, and prognosis [96,97]. Marostica et al. used transfer learning to develop CNAs and somatic mutations image-based prediction models. They demonstrated that CNAs in several genes, including KRAS, EGFR, and VHL, could affect quantitative histopathology patterns [61]. Furthermore, the group also leveraged a framework to predict ccRCC tumor mutational burden, a potential yet controversial biomarker for immune checkpoint blockade response [98], obtaining good performances on this task. It is important to note that this approach was weakly supervised and did not need a slide-level label with detailed region or pixel-level segmentation, making it readily applicable for clinical use.

Although immunotherapy has changed the field of mRCC over the last years, TKI monotherapy still plays an essential role in patients unable to receive or tolerate

 doi:10.20944/preprints202306.0221.v1

checkpoint inhibitors and as a later-line therapy [69]. Go et al. developed an ML-based method to identify which mRCC patients will respond to VEGFR-TKI treatment by analyzing clinical, pathology, and molecular data from 101 patients [99]. Specimens of the primarily resected tissue were collected and retrospectively divided into clinical and non-clinical benefit groups. The authors developed a predictive classifier obtaining a prediction accuracy of 0.87.

As stated, gene expression signatures are commonly used as predictive biomarkers. Endothelial cells and vascular architecture are known to play a role in the biological behavior of the tumor [100]. Ing et al. used ML to analyze tumor vasculature for prognostic insight [101]. They used ccRCC cases from the TCGA database to train their algorithm and discovered nine vascular features correlated with clinical outcomes. They found that 4 of these features had more significant variation in individuals with poor outcomes than favorable outcomes, linking variation in vascular structure to worse results. Ing et al. identified 14 genes that correlated strongly to these features and built two ML-based models with satisfactory prediction outcomes comparable to traditional gene signatures. Further efforts are needed to develop models using morphologic and genomic biomarkers for improved patient prognosis and treatment options.

Another active area of RCC research is in the field of epigenetics [102–106]. Zheng et al. investigated possible interactions between histopathologic features and epigenetic changes in RCC [107]. Using morphometric features extracted from histopathology images, they employed ML models to accurately forecast differential methylation values for specific genes or gene clusters. Further, prospective studies are needed to predict mechanisms underlying cancer progression using predicted genes [108].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 June 2023 doi:10.20944/preprints202306.0221.v1

Table 3. Studies aimed to uncover molecular-morphological connections and/or AI-based therapy response prediction.

Group	Aim	Number of patients	Accuracy on the test set	External validation (N of patients)	Accuracy on the external validation cohort	Algorithm
Marostica et al. [61]	1) RCC diagnosis 2) RCC subtyping, 3) CNAs identification 4) RCC survival prediction 5) tumor mutation burden prediction	1) & 2) 537 ccRCC, 288 pRCC, 103 chRCC 3) 528 ccRCC, 288 pRCC, 66 chRCC 4) 269 stage I ccRCC 5) 302 ccRCC	1)AUC: 0.990 **ccRCC**, 1.00 **pRCC**, 0.9998 **chRCC** 2) AUC: 0.953 3) **ccRCC** KRAS CNA: AUC=0.724; **pRCC** somatic mutations: AUC: 0.419 – 0.684; 4) short vs. Long-term survivors log-rank test P = 0.02, n = 269 5) Spearman correlation coefficient: 0.419	1)&2) 841 ccRCC, 41 pRCC, 31 chRCC	1) 0.964 – 0.985 ccRCC; 2) 0.782 – 0.993	DCNN
Go et al. [99]	RCC VEGFR-TKI response classifier; survival prediction	101 m-ccRCC	Apparent accuracy of the model: 87.5%; C-index = 0.7001 for PFS; C-index of 0.6552 for OS	N.A.	N.A.	SVM
Ing et al. [101]	1) RCC vascular phenotypes; 2) survival prediction; 3) identification of prognostic gene signature 4) prediction models	1); 2) & 3)64 ccRCC 4) 301 ccRCC	1) AUC = 0.79; 2) log-rank p = 0.019, HR = 2.4 3) Wilcoxon rank-sum test $p < 0.0511$ 4) C-Index: Stage = 0.7, Stage + 14VF = 0.74, Stage + 14GT = 0.74	N.A.	N.A.	1) SVM; Random Forest classifier 2) correlation analysis and information gain 3) two generalized linear models with elastic net regularization
Zheng et al. [107]	RCC methylation profile	326 RCC *(also tested on glioma)*	average AUC and F1 score higher than 0.6	N.A.	N.A.	Classic ML, FCNN

AUC= area under curve, ccRCC= clear cell renal cell carcinoma, chRCC=chromophobe renal cell carcinoma, CNA = copy number alteration, DCNN = deep convolutional neural network, DFS = disease-free survival, FCNN = fully-connected neural network, ML = machine learning, N.A.= not applicable, OS = overall survivale, PFS = Progression-free survival, pRCC=papillary renal cell carcinoma, SVM= support vector machine, VEGFR-TKI = Vascular Endothelial Growth Factor Receptor-Tyrosine Kinase Inhibitor.

4.3. PROGNOSIS PREDICTION MODELS BASED ON COMPUTATIONAL PATHOLOGY

In the past, several models have been developed and externally validated for the prediction of the prognosis of RCC patients. These models, currently used for both localized and metastatic RCC, are mainly based on clinicopathological data, both for localized and mRCC [109–113]. Currently, the prognostic models of localized ccRCC mainly include the Leibovich score [112] and the UISS score [113]. The latter is primarily based on

 doi:10.20944/preprints202306.0221.v1

clinicopathological data, making a pathologist's experience one of the limitations in their performances [114,115]. All the mentioned models incorporate clinical parameters within their framework; however models based exclusive on pathological data have been validated [116], Regarding mRCC, risk groups assigned by the Memorial Sloan Kettering Cancer Center (MSKCC) and the International Metastatic Renal Cell Carcinoma Database Consortium (IMDC) may differ in up to 23% of cases [69]. Although these models have shown reasonably good performance in the past, there is still room for improvement [117]. AI multimodal approaches applied to medical issues can raise accuracy by up to 27.7% compared to a single modality [118]. Specifically, integrating an ML-based algorithm that predicts RCC survival from histopathology to other known prognosis modalities improved in more studies prediction accuracy [119,120].

Cheng et al. was the first to combine features from the gene data and histopathologic data for ccRCC prognosis [121], generating a risk index strongly correlated with survival and outperforming prediction based on considering morphologic features or eigengenes separately. The predicted risk could also stratify early-stage patients (stage I and II), whereas no significant difference in survival outcomes when using stage alone. In the Cheng et al. study, microenvironment and radiologic imaging information were not integrated into the prognostic model. At the same time, the latter proved to be the single modality with the best predictive performance in a computational method presented by Ning et al. This method combined features extracted from CT, histopathological images, and clinical and genomic data [122]. However, Ning et al. method also had limitations, such as a small sample size and the lack of external validation. Another algorithm from Chen et al. was trained on ccRCC images from the TCGA cohort and validated on Shangai General Hospital images to identify substantial survival-related digital pathological factors and combine them with clinicopathological factors (age, stage, and grade) [60]. The integration nomogram developed showed good ability in predicting 1- 3- and 5-year DFS (Table 1). The study defined the cut-off value for high and low-risk scores as the median score for each cohort. Therefore, external validation using a larger cohort or a prospective study would be necessary to confirm the novel computational recognition model's validity and determine the optimal cut-off value for high and low-risk scores.

Another study by Schulz et al. reported on a multimodal deep learning model trained on multiscale histopathological images, CT/MRI scans, and genomic data from whole exome sequencing [123]. The model showed excellent performance in terms of 5-year survival status prediction, outperforming other parameters (T-stage, N-stage, M-stage, grading). They also investigated the possibility of predicting the 5-year survival status, obtaining a significant difference in the survival curves after dividing the cohorts into low and high-risk patients, even after evaluating only M0 or M+ patients. Also, this study had limitations: it needed to compare other clinical tools that consider factors such as performance status and calcium levels incorporated in the current, widely used prognostic models. Additionally, the external validation sample size is relatively small, and further research is required to confirm the generalizability of their approach.

The abovementioned and future new models should be externally validated, used in prospective cohorts, and compared to current prognostic models regarding discrimination, calibration, and net benefit [69].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 June 2023 doi:10.20944/preprints202306.0221.v1

Table 4. Prognostic models.

Group	Aim	Number of patients	Accuracy on the test set	External validation (N of patients)	Accuracy on the external validation cohort	Algorithm
Ning et al. [122]	RCC prognosis prediction	209 ccRCC	Mean C-index = 0.832 (0.761–0.903)	N.A.	N.A.	CNN; BFPS algorithm for feature selection
Cheng et al. [121]	RCC prognosis prediction	410 ccRCC	log-rank test *P* values<0.05	N.A.	N.A.	lmQCM – gene coexpression and analysis; ML – LASSO-Cox model for prognosis prediction
Schulz et al. [123]	RCC prognosis prediction	248 ccRCC	Mean C-index of 0.7791 and a mean accuracy of 83.43%. *(prognosis prediction)*	18 ccRCC	Mean C-index reached 0.799 ± 0.060 with a maximum of 0.8662. Accuracy averaged at 79.17% ± 9.8% with a maximum of 94.44%.	CNN consisting of one individual 18-layer residual network (ResNet) per image modality and a dense layer for genomic data

BFPS = block filtering post-pruning search, ccRCC= clear cell renal cell carcinoma, CNN= convolutional neural network, lmQCM=local maximum quasi-clique merging, ML=machine learning, N.A.= not applicable, SVM= support vector machine.

5. FUTURE PERSPECTIVES

According to currently available data, AI and ML in RCC pathology ('pathomics') holds promise for the future, as they might help to overcome several problems in classic histopathology, such as intra and interobserver variability and time consumption. Currently, several AI methods can be reliable in RCC diagnosis and, on some occasions, appear capable of predicting clinical outcomes in seconds. This could be of great help for pathologists in times were the incidence of RCC is still rising. However, this exciting field is still relatively new and not without teething troubles, both in general as specifically within the realm of RCC [124,125]

In this review, we reported on the excellent results achieved by AI in several tasks like staging and grading. Supervised learning methods efficiently perform these tasks but cannot be visually authenticated. In simple terms, the machine generates an answer (i.e., low or high grade or subtype) according to its learned algorithms that humans cannot survey. These algorithms are often referred to as black box algorithms [126]. This makes them prone to doubt by the pathology community, as the pathologist must have faith in the findings before approving and discussing a report in multidisciplinary meetings [127]. One possible solution might be implementing tools that bring transparency to non-linear machine learning techniques. For instance, gradient-weighted class activation mapping (grad-CAM) is a tool that can overlay images and heatmaps to improve the visualization of the cell type or region in which the informative features were expressed [128]. Another possible solution can be "searching and matching" instead of "classifying" in an unsupervised fashion like the group of Faust et al. did for RCC diagnosis [68]. With unsupervised learning, computers can search and cluster images with matching features in a dataset

 doi:10.20944/preprints202306.0221.v1

without labeling the data, which can be labor-intensive and potentially biased [129]. This method more or less resembles the current workflow, as pathologists often use atlases to compare images found in the specimen to match certain previously described conditions. Alternatively, consultation with other experts for a second opinion may be asked. However, this approach doesn't exclude the intervention of human experts since a pathologist still needs to inspect and interpret the images visually.

Another possible drawback of computational pathology is the current lack of generalization due to potentially biased input used in the training process of models. For example, using cross-validation, ML models are validated on a set different from the training set, which can lead to biased evaluation if the input data is biased. Therefore, a recommended step before model training is to always examine for any potential sample bias and assess whether there may be any issues related to sample size [130,131], heterogeneity [132], noise [133], and confounding factors [134].

Moreover, supposing the data is derived from one pathology laboratory, the algorithm may only be able to account for some variations and artifacts arising from different institutions. For example, the color distribution of WSIs varies across different pathology laboratories due to the staining process.

Once the data is adequately processed, the model is trained on the training set, and its performance is evaluated on the validation set. The so-called 'overfitting,' can occur when a model is too finely tuned to a particular dataset that it fails to generalize well to new, unseen data. Overfitting is like memorizing answers to a test rather than understanding the material. Once the training process is complete, the final performance of the model is evaluated on the test set, which contains data that the model has not seen before. This final evaluation estimates the model's performance on new and unseen data [22]. But if the model is overfitting it can still perform well if the data derives from the same laboratory.

This leads to inter-center variability that impacts the accuracy of machine learning algorithms used to analyze WSIs automatically. This includes state-of-the-art CNN-based algorithms, which often exhibit reduced performance when applied to images from a different center than the one they were trained on [12,135–137]. Therefore, a global standard for tissue processing, staining, slide preparation in surgical pathology, and even digital acquisition would greatly help [138]. Existing solutions to reduce generalization error in this setting can be categorized into stain color augmentation and stain color normalization, with ML-based methods to perform stain color normalization using a neural network being proposed [139]. One of the most effective methods to mitigate overfitting is external validation, testing the method on a group of new patients distinct from the initial set, thus assessing the model's generalization [22]

The critical evidence for generalizability would be introducing external validation. Any features selected based on idiosyncrasies of the original training data, such as technical or sampling biases, would likely not hold up. As a result, adequate performance on a reasonably extensive external validation set is seen as good evidence of a model's generalizability (fig. 1 and 2) [140].

To conclude, AI is a promising tool still under investigation in the diagnosis, grading, prognosis assessment, and treatment of kidney neoplasms. Results of new AI algorithms are encouraging since they are on par with or outperform current state-of-the-art methods. However, most available technologies are currently unavailable for widespread clinical use, and further evidence is needed. Therefore, more advancements in this exciting field are eagerly awaited.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 June 2023 doi:10.20944/preprints202306.0221.v1

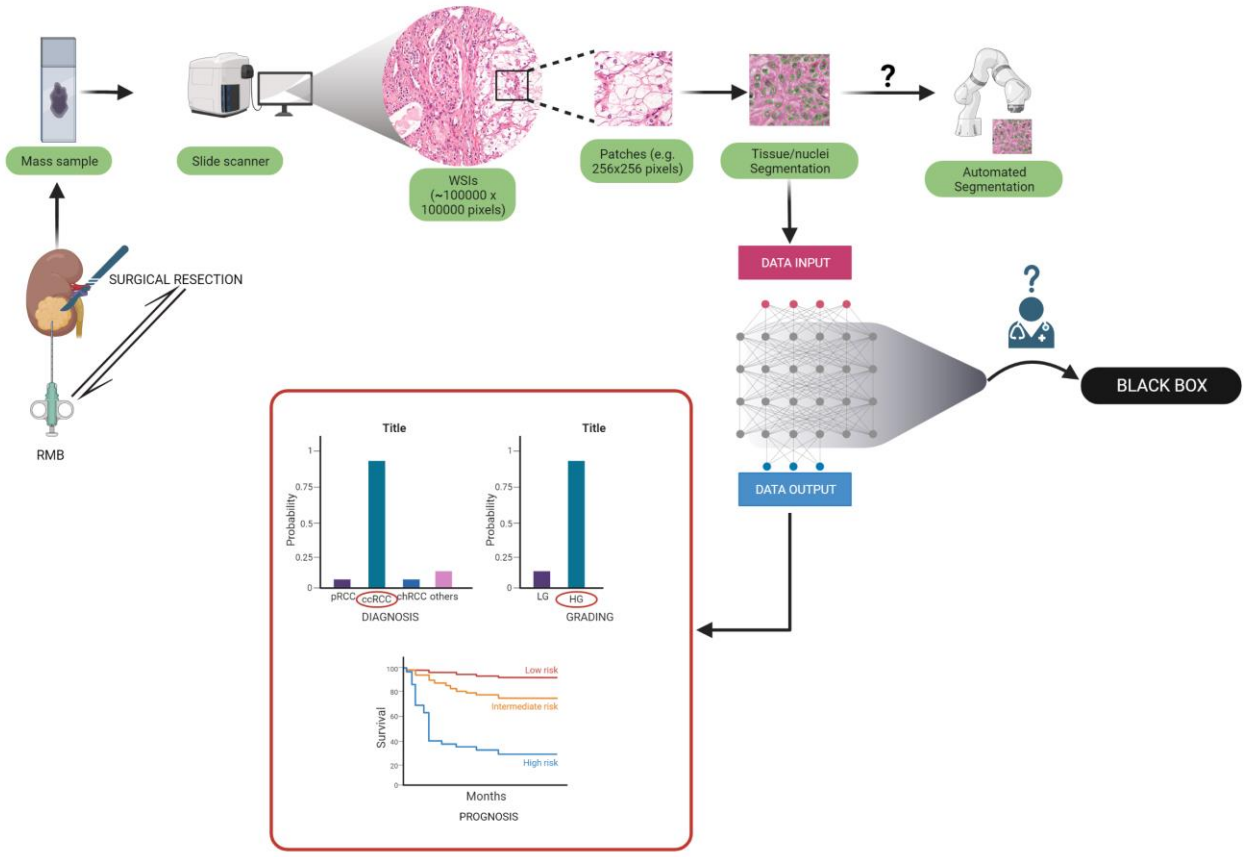

Figure 1. Pathway for the development of pathomics algorithms. After the sample is obtained by surgical resection or biopsy, through a digital scanner the WSI is created and derived patches utilized for training the algorithm to define diagnostic, prognostic or predictive models. Supervised learning based algorithm could carry the issue of "black box" (see paragraph 6).

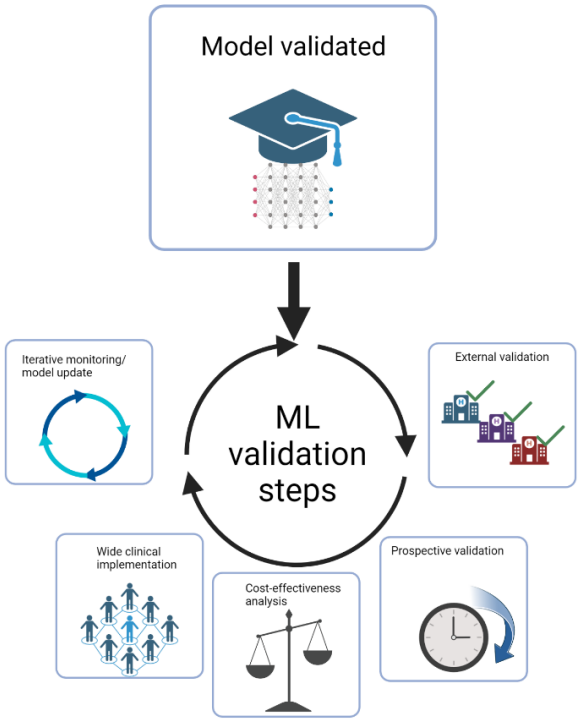

Figure 2. Challenges in clinical translation after the development of a new ML algorithm.

Author Contributions: All authors read and approved the final version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors certify that there is no conflict of interest with any financial organization regarding the material discussed in the manuscript.

Funding: Not applicable.

References

[1] Capitanio U, Bensalah K, Bex A, Boorjian SA, Bray F, Coleman J, et al. Epidemiology of Renal Cell Carcinoma. Eur Urol 2019;75:74. https://doi.org/10.1016/J.EURURO.2018.08.036.
[2] Garfield K, LaGrange CA. Renal Cell Cancer. StatPearls 2022.
[3] Bukavina L, Bensalah K, Bray F, Carlo M, Challacombe B, Karam JA, et al. Epidemiology of Renal Cell Carcinoma: 2022 Update. Eur Urol 2022;82:529–42. https://doi.org/10.1016/J.EURURO.2022.08.019.
[4] Moch H, Cubilla AL, Humphrey PA, Reuter VE, Ulbright TM. The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs-Part A: Renal, Penile, and Testicular Tumours. Eur Urol 2016;70:93–105. https://doi.org/10.1016/J.EURURO.2016.02.029.
[5] Cimadamore A, Caliò A, Marandino L, Marletta S, Franzese C, Schips L, et al. Hot topics in renal cancer pathology: implications for clinical management. Expert Rev Anticancer Ther 2022;22:1275–87. https://doi.org/10.1080/14737140.2022.2145952.
[6] Fuhrman SA, Lasky LC, Limas C. Prognostic significance of morphologic parameters in renal cell carcinoma. Am J Surg Pathol 1982;6:655–63. https://doi.org/10.1097/00000478-198210000-00007.
[7] Zhang L, Zha Z, Qu W, Zhao H, Yuan J, Feng Y, et al. Tumor necrosis as a prognostic variable for the clinical outcome in patients with renal cell carcinoma: a systematic review and meta-analysis. BMC Cancer 2018;18. https://doi.org/10.1186/S12885-018-4773-Z.
[8] Sun M, Shariat SF, Cheng C, Ficarra V, Murai M, Oudard S, et al. Prognostic factors and predictive models in renal cell carcinoma: a contemporary review. Eur Urol 2011;60:644–61. https://doi.org/10.1016/J.EURURO.2011.06.041.
[9] Hora M, Albiges L, Bedke J, Campi R, Capitanio U, Giles RH, et al. European Association of Urology Guidelines Panel on Renal Cell Carcinoma Update on the New World Health Organization Classification of Kidney Tumours 2022: The Urologist's Point of View. Eur Urol 2023;83:97–100. https://doi.org/10.1016/J.EURURO.2022.11.001.
[10] Baidoshvili A, Bucur A, van Leeuwen J, van der Laak J, Kluin P, van Diest PJ. Evaluating the benefits of digital pathology implementation: time savings in laboratory logistics. Histopathology 2018;73:784–94. https://doi.org/10.1111/HIS.13691.
[11] Shmatko A, Ghaffari Laleh N, Gerstung M, Kather JN. Artificial intelligence in histopathology: enhancing cancer research and clinical oncology. Nat Cancer 2022;3:1026–38. https://doi.org/10.1038/S43018-022-00436-4.
[12] Komura D, Ishikawa S. Machine Learning Methods for Histopathological Image Analysis. Comput Struct Biotechnol J 2018;16:34–42. https://doi.org/10.1016/J.CSBJ.2018.01.001.
[13] Lecun Y, Bengio Y, Hinton G. Deep learning. Nature 2015 521:7553 2015;521:436–44. https://doi.org/10.1038/nature14539.
[14] Roussel E, Capitanio U, Kutikov A, Oosterwijk E, Pedrosa I, Rowe SP, et al. Novel Imaging Methods for Renal Mass Characterization: A Collaborative Review. Eur Urol 2022;81:476–88. https://doi.org/10.1016/J.EURURO.2022.01.040.
[15] Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology. Nat Rev Clin Oncol 2019;16:703–15. https://doi.org/10.1038/S41571-019-0252-Y.
[16] Niazi MKK, Parwani A v., Gurcan MN. Digital pathology and artificial intelligence. Lancet Oncol 2019;20:e253–61. https://doi.org/10.1016/S1470-2045(19)30154-8.
[17] Colling R, Pitman H, Oien K, Rajpoot N, Macklin P, Bachtiar V, et al. Artificial intelligence in digital pathology: a roadmap to routine use in clinical practice. J Pathol 2019;249:143–50. https://doi.org/10.1002/PATH.5310.
[18] Volpe A, Patard JJ. Prognostic factors in renal cell carcinoma. World J Urol 2010;28:319–27. https://doi.org/10.1007/S00345-010-0540-8.
[19] Tucker MD, Rini BI. Predicting Response to Immunotherapy in Metastatic Renal Cell Carcinoma. Cancers (Basel) 2020;12:1–20. https://doi.org/10.3390/CANCERS12092662.
[20] Wang D, Khosla A, Gargeya R, Irshad H, Beck AH, Israel B. Deep Learning for Identifying Metastatic Breast Cancer 2016. https://doi.org/10.48550/arxiv.1606.05718.
[21] Hayashi Y. Black Box Nature of Deep Learning for Digital Pathology: Beyond Quantitative to Qualitative Algorithmic Performances. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 2020;12090 LNCS:95–101. https://doi.org/10.1007/978-3-030-50402-1_6/COVER.
[22] Tougui I, Jilbab A, Mhamdi J el. Impact of the Choice of Cross-Validation Techniques on the Results of Machine Learning-Based Diagnostic Applications. Healthc Inform Res 2021;27:189–99. https://doi.org/10.4258/HIR.2021.27.3.189.
[23] Cabitza F, Campagner A, Soares F, García de Guadiana-Romualdo L, Challa F, Sulejmani A, et al. The importance of being external. methodological insights for the external validation of machine learning models in medicine. Comput Methods Programs Biomed 2021;208:106288. https://doi.org/10.1016/J.CMPB.2021.106288.

[24] Creighton CJ, Morgan M, Gunaratne PH, Wheeler DA, Gibbs RA, Robertson G, et al. Comprehensive molecular characterization of clear cell renal cell carcinoma. Nature 2013 499:7456 2013;499:43–9. https://doi.org/10.1038/nature12222.
[25] Krajewski KM, Pedrosa I. Imaging Advances in the Management of Kidney Cancer. Journal of Clinical Oncology 2018;36:3582. https://doi.org/10.1200/JCO.2018.79.1236.
[26] Roussel E, Campi R, Amparore D, Bertolo R, Carbonara U, Erdem S, et al. Expanding the Role of Ultrasound for the Characterization of Renal Masses. J Clin Med 2022;11. https://doi.org/10.3390/JCM11041112.
[27] Shuch B, Hofmann JN, Merino MJ, Nix JW, Vourganti S, Linehan WM, et al. Pathologic validation of renal cell carcinoma histology in the Surveillance, Epidemiology, and End Results program. Urol Oncol 2014;32:23.e9-23.e13. https://doi.org/10.1016/J.UROLONC.2012.08.011.
[28] Al-Aynati M, Chen V, Salama S, Shuhaibar H, Treleaven D, Vincic L. Interobserver and Intraobserver Variability Using the Fuhrman Grading System for Renal Cell Carcinoma. Arch Pathol Lab Med 2003;127:593–6. https://doi.org/10.5858/2003-127-0593-IAIVUT.
[29] Williamson SR, Rao P, Hes O, Epstein JI, Smith SC, Picken MM, et al. Challenges in pathologic staging of renal cell carcinoma: A study of interobserver variability among urologic pathologists. American Journal of Surgical Pathology 2018;42:1253–61. https://doi.org/10.1097/PAS.0000000000001087.
[30] Gavrielides MA, Gallas BD, Lenz P, Badano A, Hewitt SM. Observer variability in the interpretation of HER2/neu immunohistochemical expression with unaided and computer-aided digital microscopy. Arch Pathol Lab Med 2011;135:233–42. https://doi.org/10.5858/135.2.233.
[31] Ficarra V, Martignoni G, Galfano A, Novara G, Gobbo S, Brunelli M, et al. Prognostic role of the histologic subtypes of renal cell carcinoma after slide revision. Eur Urol 2006;50:786–94. https://doi.org/10.1016/J.EURURO.2006.04.009.
[32] Multicenter determination of optimal interobserver agreement using the Fuhrman grading system for renal cell carcinoma - Lang - 2005 - Cancer - Wiley Online Library n.d. https://acsjournals.onlinelibrary.wiley.com/doi/full/10.1002/cncr.20812 (accessed February 1, 2023).
[33] Smaldone MC, Egleston B, Hollingsworth JM, Hollenbeck BK, Miller DC, Morgan TM, et al. Understanding Treatment Disconnect and Mortality Trends in Renal Cell Carcinoma Using Tumor Registry Data. Med Care 2017;55:398–404. https://doi.org/10.1097/MLR.0000000000000657.
[34] Kutikov A, Smaldone MC, Egleston BL, Manley BJ, Canter DJ, Simhan J, et al. Anatomic Features of Enhancing Renal Masses Predict Malignant and High-Grade Pathology: A Preoperative Nomogram Using the RENAL Nephrometry Score. Eur Urol 2011;60:241. https://doi.org/10.1016/J.EURURO.2011.03.029.
[35] Pierorazio PM, Patel HD, Johnson MH, Sozio SM, Sharma R, Iyoha E, et al. Distinguishing malignant and benign renal masses with composite models and nomograms: A systematic review and meta-analysis of clinically localized renal masses suspicious for malignancy. Cancer 2016;122:3267–76. https://doi.org/10.1002/CNCR.30268.
[36] Joshi S, Kutikov A. Understanding Mutational Drivers of Risk: An Important Step Toward Personalized Care for Patients with Renal Cell Carcinoma. Eur Urol Focus 2017;3:428–9. https://doi.org/10.1016/J.EUF.2016.07.008.
[37] Nguyen MM, Gill IS, Ellison LM. The evolving presentation of renal carcinoma in the United States: trends from the Surveillance, Epidemiology, and End Results program. J Urol 2006;176:2397–400. https://doi.org/10.1016/J.JURO.2006.07.144.
[38] Sohlberg EM, Metzner TJ, Leppert JT. The Harms of Overdiagnosis and Overtreatment in Patients with Small Renal Masses: A Mini-review. Eur Urol Focus 2019;5:943–5. https://doi.org/10.1016/J.EUF.2019.03.006.
[39] Campi R, Stewart GD, Staehler M, Dabestani S, Kuczyk MA, Shuch BM, 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. https://doi.org/10.1016/J.EUO.2020.12.011.
[40] Warren H, Palumbo C, Caliò A, Tran MGB, Campi R, Courcier J, et al. Oncocytoma on renal mass biopsy: why is surgery even performed? World J Urol 2023:1–2. https://doi.org/10.1007/S00345-023-04402-2/METRICS.
[41] Kutikov A, Smaldone MC, Uzzo RG, Haifler M, Bratslavsky G, Leibovich BC. Renal Mass Biopsy: Always, Sometimes, or Never? Eur Urol 2016;70:403–6. https://doi.org/10.1016/J.EURURO.2016.04.001.
[42] Lane BR, Samplaski MK, Herts BR, Zhou M, Novick AC, Campbell SC. Renal mass biopsy--a renaissance? J Urol 2008;179:20–7. https://doi.org/10.1016/J.JURO.2007.08.124.
[43] Marconi L, Dabestani S, Lam TB, Hofmann F, Stewart F, Norrie J, et al. Systematic Review and Meta-analysis of Diagnostic Accuracy of Percutaneous Renal Tumour Biopsy. Eur Urol 2016;69:660–73. https://doi.org/10.1016/J.EURURO.2015.07.072.
[44] Evans AJ, Delahunt B, Srigley JR. Issues and challenges associated with classifying neoplasms in percutaneous needle biopsies of incidentally found small renal masses. Semin Diagn Pathol 2015;32:184–95. https://doi.org/10.1053/J.SEMDP.2015.02.009.
[45] Kümmerlin I, ten Kate F, Smedts F, Horn T, Algaba F, Trias I, et al. Core biopsies of renal tumors: a study on diagnostic accuracy, interobserver, and intraobserver variability. Eur Urol 2008;53:1219–27. https://doi.org/10.1016/J.EURURO.2007.11.054.
[46] Elmore JG, Longton GM, Carney PA, Geller BM, Onega T, Tosteson ANA, et al. Diagnostic concordance among pathologists interpreting breast biopsy specimens. JAMA 2015;313:1122–32. https://doi.org/10.1001/JAMA.2015.1405.
[47] Elmore JG, Barnhill RL, Elder DE, Longton GM, Pepe MS, Reisch LM, et al. Pathologists' diagnosis of invasive melanoma and melanocytic proliferations: observer accuracy and reproducibility study. BMJ 2017;357. https://doi.org/10.1136/BMJ.J2813.
[48] Shah MD, Parwani A v., Zynger DL. Impact of the Pathologist on Prostate Biopsy Diagnosis and Immunohistochemical Stain Usage Within a Single Institution. Am J Clin Pathol 2017;148:494–501. https://doi.org/10.1093/AJCP/AQX103.

[49] Fenstermaker M, Tomlins SA, Singh K, Wiens J, Morgan TM. Development and Validation of a Deep-learning Model to Assist With Renal Cell Carcinoma Histopathologic Interpretation. Urology 2020;144:152–7. https://doi.org/10.1016/J.UROLOGY.2020.05.094.
[50] van Oostenbrugge TJ, Fütterer JJ, Mulders PFA. Diagnostic Imaging for Solid Renal Tumors: A Pictorial Review. Kidney Cancer 2018;2:79–93. https://doi.org/10.3233/KCA-180028.
[51] Williams GM, Lynch DT. Renal Oncocytoma. StatPearls 2022.
[52] Leone AR, Kidd LC, Diorio GJ, Zargar-Shoshtari K, Sharma P, Sexton WJ, et al. Bilateral benign renal oncocytomas and the role of renal biopsy: Single institution review. BMC Urol 2017;17:1–6. https://doi.org/10.1186/S12894-016-0190-2/TABLES/3.
[53] Zhu M, Ren B, Richards R, Suriawinata M, Tomita N, Hassanpour S. Development and evaluation of a deep neural network for histologic classification of renal cell carcinoma on biopsy and surgical resection slides. Sci Rep 2021;11. https://doi.org/10.1038/S41598-021-86540-4.
[54] Volpe A, Mattar K, Finelli A, Kachura JR, Evans AJ, Geddie WR, et al. Contemporary results of percutaneous biopsy of 100 small renal masses: a single center experience. J Urol 2008;180:2333–7. https://doi.org/10.1016/J.JURO.2008.08.014.
[55] Wang R, Wolf JS, Wood DP, Higgins EJ, Hafez KS. Accuracy of Percutaneous Core Biopsy in Management of Small Renal Masses. Urology 2009;73:586–90. https://doi.org/10.1016/J.UROLOGY.2008.08.519.
[56] Barwari K, de La Rosette JJ, Laguna MP. The penetration of renal mass biopsy in daily practice: a survey among urologists. J Endourol 2012;26:737–47. https://doi.org/10.1089/END.2011.0407.
[57] Escudier B. Emerging immunotherapies for renal cell carcinoma. Annals of Oncology 2012;23:viii35–40. https://doi.org/10.1093/ANNONC/MDS261.
[58] Bertolo R, Pecoraro A, Carbonara U, Amparore D, Diana P, Muselaers S, et al. Resection Techniques During Robotic Partial Nephrectomy: A Systematic Review. Eur Urol Open Sci 2023;52:7–21. https://doi.org/10.1016/J.EUROS.2023.03.008.
[59] Tabibu S, Vinod PK, Jawahar C v. Pan-Renal Cell Carcinoma classification and survival prediction from histopathology images using deep learning. Sci Rep 2019;9. https://doi.org/10.1038/S41598-019-46718-3.
[60] Chen S, Zhang N, Jiang L, Gao F, Shao J, Wang T, et al. Clinical use of a machine learning histopathological image signature in diagnosis and survival prediction of clear cell renal cell carcinoma. Int J Cancer 2021;148:780–90. https://doi.org/10.1002/IJC.33288.
[61] Marostica E, Barber R, Denize T, Kohane IS, Signoretti S, Golden JA, et al. Development of a Histopathology Informatics Pipeline for Classification and Prediction of Clinical Outcomes in Subtypes of Renal Cell Carcinoma. Clin Cancer Res 2021;27:2868–78. https://doi.org/10.1158/1078-0432.CCR-20-4119.
[62] Pathology Outlines - WHO classification n.d. https://www.pathologyoutlines.com/topic/kidneytumorWHOclass.html (accessed January 24, 2023).
[63] Cimadamore A, Cheng L, Scarpelli M, Massari F, Mollica V, Santoni M, et al. Towards a new WHO classification of renal cell tumor: what the clinician needs to know—a narrative review. Transl Androl Urol 2021;10:1506. https://doi.org/10.21037/TAU-20-1150.
[64] Moch H, Cubilla AL, Humphrey PA, Reuter VE, Ulbright TM. The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs-Part A: Renal, Penile, and Testicular Tumours. Eur Urol 2016;70:93–105. https://doi.org/10.1016/J.EURURO.2016.02.029.
[65] Weng S, DiNatale RG, Silagy A, Mano R, Attalla K, Kashani M, et al. The Clinicopathologic and Molecular Landscape of Clear Cell Papillary Renal Cell Carcinoma: Implications in Diagnosis and Management. Eur Urol 2021;79:468–77. https://doi.org/10.1016/j.eururo.2020.09.027.
[66] Williamson SR, Eble JN, Cheng L, Grignon DJ. Clear cell papillary renal cell carcinoma: Differential diagnosis and extended immunohistochemical profile. Modern Pathology 2013;26:697–708. https://doi.org/10.1038/modpathol.2012.204.
[67] Abdeltawab HA, Khalifa FA, Ghazal MA, Cheng L, El-Baz AS, Gondim DD. A deep learning framework for automated classification of histopathological kidney whole-slide images. J Pathol Inform 2022;13:100093. https://doi.org/10.1016/J.JPI.2022.100093.
[68] Faust K, Roohi A, Leon AJ, Leroux E, Dent A, Evans AJ, et al. Unsupervised Resolution of Histomorphologic Heterogeneity in Renal Cell Carcinoma Using a Brain Tumor-Educated Neural Network. JCO Clin Cancer Inform 2020;4:811–21. https://doi.org/10.1200/CCI.20.00035.
[69] Renal Cell Carcinoma EAU Guidelines on 2022.
[70] Gelb AB. C O M M U N I C A T I O N Union Internationale Contre le Cancer (UICC) and the American Joint Renal Cell Carcinoma Committee on Cancer (AJCC) Current Prognostic Factors BACKGROUND. Renal cell carcinomas include several distinct entities with a range n.d. https://doi.org/10.1002/(SICI)1097-0142(19970901)80:5.
[71] Beksac AT, Paulucci DJ, Blum KA, Yadav SS, Sfakianos JP, Badani KK. Heterogeneity in renal cell carcinoma. Urologic Oncology: Seminars and Original Investigations 2017;35:507–15. https://doi.org/10.1016/J.UROLONC.2017.05.006.
[72] Dall'Oglio MF, Ribeiro-Filho LA, Antunes AA, Crippa A, Nesrallah L, Gonçalves PD, et al. Microvascular Tumor Invasion, Tumor Size and Fuhrman Grade: A Pathological Triad for Prognostic Evaluation of Renal Cell Carcinoma. J Urol 2007;178:425–8. https://doi.org/10.1016/J.JURO.2007.03.128.
[73] Tsui KH, Shvarts O, Smith RB, Figlin RA, Dekernion JB, Belldegrun A. PROGNOSTIC INDICATORS FOR RENAL CELL CARCINOMA: A MULTIVARIATE ANALYSIS OF 643 PATIENTS USING THE REVISED 1997 TNM STAGING CRITERIA. J Urol 2000;163:1090–5. https://doi.org/10.1016/S0022-5347(05)67699-9.

[74] Ficarra V, Righetti R, Pillonia S, D'amico A, Maffei N, Novella G, et al. Prognostic Factors in Patients with Renal Cell Carcinoma: Retrospective Analysis of 675 Cases. Eur Urol 2002;41:190–8. https://doi.org/10.1016/S0302-2838(01)00027-6.
[75] Scopus preview - Scopus - Document details - Prognostic significance of morphologic parameters in renal cell carcinoma n.d. https://www.scopus.com/record/display.uri?eid=2-s2.0-2642552183&origin=inward&txGid=18f4bff1afabc920febe75bb222fbbab (accessed January 18, 2023).
[76] Prognostic value of nuclear grade of renal cell carcinoma n.d. https://acsjournals.onlinelibrary.wiley.com/doi/epdf/10.1002/1097-0142(19951215)76:12%3C2543::AID-CNCR2820761221%3E3.0.CO;2-S?src=getftr (accessed January 18, 2023).
[77] Intraobserver and Interobserver Variability of Fuhrman and Modified Fuhrman Grading Systems for Conventional Renal Cell Carcinoma - Bektas - 2009 - The Kaohsiung Journal of Medical Sciences - Wiley Online Library n.d. https://onlinelibrary.wiley.com/doi/abs/10.1016/S1607-551X(09)70562-5 (accessed January 18, 2023).
[78] Al-Aynati M, Chen V, Salama S, Shuhaibar H, Treleaven D, Vincic L. Interobserver and Intraobserver Variability Using the Fuhrman Grading System for Renal Cell Carcinoma. Arch Pathol Lab Med 2003;127:593–6. https://doi.org/10.5858/2003-127-0593-IAIVUT.
[79] Multicenter determination of optimal interobserver agreement using the Fuhrman grading system for renal cell carcinoma n.d. https://acsjournals.onlinelibrary.wiley.com/doi/epdf/10.1002/cncr.20812?src=getftr (accessed January 18, 2023).
[80] Yeh F-C, Parwani A V., Pantanowitz L, Ho C. Automated grading of renal cell carcinoma using whole slide imaging. J Pathol Inform 2014;5:23. https://doi.org/10.4103/2153-3539.137726.
[81] Paner GP, Stadler WM, Hansel DE, Montironi R, Lin DW, Amin MB. Updates in the Eighth Edition of the Tumor-Node-Metastasis Staging Classification for Urologic Cancers. Eur Urol 2018;73:560–9. https://doi.org/10.1016/J.EURURO.2017.12.018.
[82] Holdbrook DA, Singh M, Choudhury Y, Kalaw EM, Koh V, Tan HS, et al. Automated Renal Cancer Grading Using Nuclear Pleomorphic Patterns. JCO Clin Cancer Inform 2018;2:1–12. https://doi.org/10.1200/CCI.17.00100.
[83] Qayyum T, McArdle P, Orange C, Seywright M, Horgan P, Oades G, et al. Reclassification of the Fuhrman grading system in renal cell carcinoma-does it make a difference? Springerplus 2013;2:1–4. https://doi.org/10.1186/2193-1801-2-378/TABLES/3.
[84] Tian K, Rubadue CA, Lin DI, Veta M, Pyle ME, Irshad H, et al. Automated clear cell renal carcinoma grade classification with prognostic significance. PLoS One 2019;14. https://doi.org/10.1371/JOURNAL.PONE.0222641.
[85] Song J, Xiao L, Lian Z. Contour-Seed Pairs Learning-Based Framework for Simultaneously Detecting and Segmenting Various Overlapping Cells/Nuclei in Microscopy Images. IEEE Trans Image Process 2018;27:5759–74. https://doi.org/10.1109/TIP.2018.2857001.
[86] Role of VHL gene mutation in human renal cell carcinoma | SpringerLink n.d. https://link.springer.com/article/10.1007/s13277-011-0257-3 (accessed January 18, 2023).
[87] Nogueira M, Kim HL. Molecular markers for predicting prognosis of renal cell carcinoma. Urologic Oncology: Seminars and Original Investigations 2008;26:113–24. https://doi.org/10.1016/J.UROLONC.2007.03.028.
[88] Roussel E, Beuselinck B, Albersen M. Tailoring treatment in metastatic renal cell carcinoma. Nat Rev Urol 2022;19:455–6. https://doi.org/10.1038/S41585-022-00615-Z.
[89] Funakoshi T, Lee CH, Hsieh JJ. A systematic review of predictive and prognostic biomarkers for VEGF-targeted therapy in renal cell carcinoma. Cancer Treat Rev 2014;40:533–47. https://doi.org/10.1016/J.CTRV.2013.11.008.
[90] Rodriguez-Vida A, Strijbos M, Hutson T. Predictive and prognostic biomarkers of targeted agents and modern immunotherapy in renal cell carcinoma. ESMO Open 2016;1:e000013. https://doi.org/10.1136/ESMOOPEN-2015-000013.
[91] Motzer RJ, Robbins PB, Powles T, Albiges L, Haanen JB, Larkin J, et al. Avelumab plus axitinib versus sunitinib in advanced renal cell carcinoma: biomarker analysis of the phase 3 JAVELIN Renal 101 trial. Nat Med 2020;26:1733. https://doi.org/10.1038/S41591-020-1044-8.
[92] Schimmel H, Zegers I, Emons H. Standardization of protein biomarker measurements: Is it feasible? 2010. https://doi.org/10.3109/00365513.2010.493362.
[93] Mayeux R. Biomarkers: Potential Uses and Limitations. NeuroRx 2004;1:182. https://doi.org/10.1602/NEURORX.1.2.182.
[94] Singh NP, Bapi RS, Vinod PK. Machine learning models to predict the progression from early to late stages of papillary renal cell carcinoma. Comput Biol Med 2018;100:92–9. https://doi.org/10.1016/J.COMPBIOMED.2018.06.030.
[95] Bhalla S, Chaudhary K, Kumar R, Sehgal M, Kaur H, Sharma S, et al. Gene expression-based biomarkers for discriminating early and late stage of clear cell renal cancer. Scientific Reports 2017 7:1 2017;7:1–13. https://doi.org/10.1038/srep44997.
[96] Fernandes FG, Silveira HCS, Júnior JNA, da Silveira RA, Zucca LE, Cárcano FM, et al. Somatic Copy Number Alterations and Associated Genes in Clear-Cell Renal-Cell Carcinoma in Brazilian Patients. Int J Mol Sci 2021;22:1–14. https://doi.org/10.3390/IJMS22052265.
[97] D'Avella C, Abbosh P, Pal SK, Geynisman DM. Mutations in renal cell carcinoma. Urologic Oncology: Seminars and Original Investigations 2020;38:763–73. https://doi.org/10.1016/J.UROLONC.2018.10.027.
[98] Havel JJ, Chowell D, Chan TA. The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy. Nature Reviews Cancer 2019 19:3 2019;19:133–50. https://doi.org/10.1038/s41568-019-0116-x.
[99] Go H, Kang MJ, Kim PJ, Lee JL, Park JY, Park JM, et al. Development of Response Classifier for Vascular Endothelial Growth Factor Receptor (VEGFR)-Tyrosine Kinase Inhibitor (TKI) in Metastatic Renal Cell Carcinoma. Pathol Oncol Res 2019;25:51–8. https://doi.org/10.1007/S12253-017-0323-2.

[100] Padmanabhan RK, Somasundar VH, Griffith SD, Zhu J, Samoyedny D, Tan KS, et al. An Active Learning Approach for Rapid Characterization of Endothelial Cells in Human Tumors. PLoS One 2014;9:e90495. https://doi.org/10.1371/JOURNAL.PONE.0090495.
[101] Ing N, Huang F, Conley A, You S, Ma Z, Klimov S, et al. A novel machine learning approach reveals latent vascular phenotypes predictive of renal cancer outcome. Sci Rep 2017;7. https://doi.org/10.1038/S41598-017-13196-4.
[102] Herman JG, Latif F, Weng Y, Lerman MI, Zbar B, Liu S, et al. Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma. Proc Natl Acad Sci U S A 1994;91:9700. https://doi.org/10.1073/PNAS.91.21.9700.
[103] Yamana K, Ohashi R, Tomita Y. Contemporary Drug Therapy for Renal Cell Carcinoma— Evidence Accumulation and Histological Implications in Treatment Strategy. Biomedicines 2022;10:2840. https://doi.org/10.3390/biomedicines10112840.
[104] Zhu L, Wang J, Kong W, Huang J, Dong B, Huang Y, et al. LSD1 inhibition suppresses the growth of clear cell renal cell carcinoma via upregulating P21 signaling. Acta Pharm Sin B 2019;9:324–34. https://doi.org/10.1016/J.APSB.2018.10.006.
[105] Chen W, Zhang H, Chen Z, Jiang H, Liao L, Fan S, et al. Development and evaluation of a novel series of Nitroxoline-derived BET inhibitors with antitumor activity in renal cell carcinoma. Oncogenesis 2018 7:11 2018;7:1–11. https://doi.org/10.1038/s41389-018-0093-z.
[106] Joosten SC, Smits KM, Aarts MJ, Melotte V, Koch A, Tjan-Heijnen VC, et al. Epigenetics in renal cell cancer: mechanisms and clinical applications. Nature Reviews Urology 2018 15:7 2018;15:430–51. https://doi.org/10.1038/s41585-018-0023-z.
[107] Zheng H, Momeni A, Cedoz PL, Vogel H, Gevaert O. Whole slide images reflect DNA methylation patterns of human tumors. NPJ Genom Med 2020;5. https://doi.org/10.1038/S41525-020-0120-9.
[108] Singh NP, Vinod PK. Integrative analysis of DNA methylation and gene expression in papillary renal cell carcinoma. Mol Genet Genomics 2020;295:807–24. https://doi.org/10.1007/S00438-020-01664-Y.
[109] Guida A, le Teuff G, Alves C, Colomba E, di Nunno V, Derosa L, et al. Identification of international metastatic renal cell carcinoma database consortium (IMDC) intermediate-risk subgroups in patients with metastatic clear-cell renal cell carcinoma. Oncotarget 2020;11:4582–92. https://doi.org/10.18632/ONCOTARGET.27762.
[110] Zigeuner R, Hutterer G, Chromecki T, Imamovic A, Kampel-Kettner K, Rehak P, et al. External validation of the Mayo Clinic stage, size, grade, and necrosis (SSIGN) score for clear-cell renal cell carcinoma in a single European centre applying routine pathology. Eur Urol 2010;57:102–11. https://doi.org/10.1016/J.EURURO.2008.11.033.
[111] Prediction of progression after radical nephrectomy for patients with clear cell renal cell carcinoma: a stratification tool for prospective clinical trials - PubMed n.d. https://pubmed.ncbi.nlm.nih.gov/12655523/ (accessed March 5, 2023).
[112] Leibovich BC, Blute ML, Cheville JC, Lohse CM, Frank I, Kwon ED, et al. Prediction of progression after radical nephrectomy for patients with clear cell renal cell carcinoma: a stratification tool for prospective clinical trials. Cancer 2003;97:1663–71. https://doi.org/10.1002/CNCR.11234.
[113] Zisman A, Pantuck AJ, Dorey F, Said JW, Shvarts O, Quintana D, et al. Improved prognostication of renal cell carcinoma using an integrated staging system. J Clin Oncol 2001;19:1649–57. https://doi.org/10.1200/JCO.2001.19.6.1649.
[114] Lubbock ALR, Stewart GD, O'Mahony FC, Laird A, Mullen P, O'Donnell M, et al. Overcoming intratumoural heterogeneity for reproducible molecular risk stratification: A case study in advanced kidney cancer. BMC Med 2017;15:1–12. https://doi.org/10.1186/S12916-017-0874-9/FIGURES/5.
[115] Heng DYC, Xie W, Regan MM, Harshman LC, Bjarnason GA, Vaishampayan UN, et al. External validation and comparison with other models of the International Metastatic Renal-Cell Carcinoma Database Consortium prognostic model: a population-based study. Lancet Oncol 2013;14:141–8. https://doi.org/10.1016/S1470-2045(12)70559-4.
[116] Erdem S, Capitanio U, Campi R, Mir MC, Roussel E, Pavan N, et al. External validation of the VENUSS prognostic model to predict recurrence after surgery in non-metastatic papillary renal cell carcinoma: A multi-institutional analysis. Urol Oncol 2022;40:198.e9-198.e17. https://doi.org/10.1016/J.UROLONC.2022.01.006.
[117] di Nunno V, Mollica V, Schiavina R, Nobili E, Fiorentino M, Brunocilla E, et al. Improving IMDC Prognostic Prediction Through Evaluation of Initial Site of Metastasis in Patients With Metastatic Renal Cell Carcinoma. Clin Genitourin Cancer 2020;18:e83–90. https://doi.org/10.1016/J.CLGC.2019.08.007.
[118] Huang SC, Pareek A, Seyyedi S, Banerjee I, Lungren MP. Fusion of medical imaging and electronic health records using deep learning: a systematic review and implementation guidelines. Npj Digital Medicine 2020 3:1 2020;3:1–9. https://doi.org/10.1038/s41746-020-00341-z.
[119] Wessels F, Schmitt M, Krieghoff-Henning E, Kather JN, Nientiedt M, Kriegmair MC, et al. Deep learning can predict survival directly from histology in clear cell renal cell carcinoma. PLoS One 2022;17. https://doi.org/10.1371/JOURNAL.PONE.0272656.
[120] Chen S, Jiang L, Gao F, Zhang E, Wang T, Zhang N, et al. Machine learning-based pathomics signature could act as a novel prognostic marker for patients with clear cell renal cell carcinoma. Br J Cancer 2022;126:771–7. https://doi.org/10.1038/S41416-021-01640-2.
[121] Cheng J, Zhang J, Han Y, Wang X, Ye X, Meng Y, et al. Integrative Analysis of Histopathological Images and Genomic Data Predicts Clear Cell Renal Cell Carcinoma Prognosis. Cancer Res 2017;77:e91–100. https://doi.org/10.1158/0008-5472.CAN-17-0313.
[122] Ning Z, Ning Z, Pan W, Pan W, Chen Y, Xiao Q, et al. Integrative analysis of cross-modal features for the prognosis prediction of clear cell renal cell carcinoma. Bioinformatics 2020;36:2888–95. https://doi.org/10.1093/BIOINFORMATICS/BTAA056.
[123] Schulz S, Woerl AC, Jungmann F, Glasner C, Stenzel P, Strobl S, et al. Multimodal Deep Learning for Prognosis Prediction in Renal Cancer. Front Oncol 2021;11. https://doi.org/10.3389/FONC.2021.788740.

[124] Khene ZE, Kutikov A, Campi R. Machine learning in renal cell carcinoma research: the promise and pitfalls of 'renal-izing' the potential of artificial intelligence. BJU Int 2023. https://doi.org/10.1111/BJU.16016.
[125] Wu Z, Carbonara U, Campi R. Re: Criteria for the Translation of Radiomics into Clinically Useful Tests. Eur Urol 2023;0. https://doi.org/10.1016/j.eururo.2023.02.020.
[126] Shortliffe EH, Sepúlveda MJ. Clinical Decision Support in the Era of Artificial Intelligence. JAMA 2018;320:2199–200. https://doi.org/10.1001/JAMA.2018.17163.
[127] Durán JM, Jongsma KR. Who is afraid of black box algorithms? On the epistemological and ethical basis of trust in medical AI. J Med Ethics 2021;47:329–35. https://doi.org/10.1136/MEDETHICS-2020-106820.
[128] Teo YY (Alan), Danilevsky A, Shomron N. Overcoming Interpretability in Deep Learning Cancer Classification. Methods in Molecular Biology 2021;2243:297–309. https://doi.org/10.1007/978-1-0716-1103-6_15/COVER.
[129] Das S, Moore T, Wong WK, Stumpf S, Oberst I, McIntosh K, et al. End-user feature labeling: Supervised and semi-supervised approaches based on locally-weighted logistic regression. Artif Intell 2013;204:56–74. https://doi.org/10.1016/J.ARTINT.2013.08.003.
[130] Krzywinski M, Altman N. Points of significance: Power and sample size. Nat Methods 2013;10:1139–40. https://doi.org/10.1038/NMETH.2738.
[131] Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ, et al. Power failure: why small sample size undermines the reliability of neuroscience. Nature Reviews Neuroscience 2013 14:5 2013;14:365–76. https://doi.org/10.1038/nrn3475.
[132] Wang L. Heterogeneous Data and Big Data Analytics. Automatic Control and Information Sciences 2017;3:8–15. https://doi.org/10.12691/acis-3-1-3.
[133] Borodinov N, Neumayer S, Kalinin S v., Ovchinnikova OS, Vasudevan RK, Jesse S. Deep neural networks for understanding noisy data applied to physical property extraction in scanning probe microscopy. NPJ Comput Mater 2019;5. https://doi.org/10.1038/S41524-019-0148-5.
[134] Goh WW bin, Wong L. Dealing with Confounders in Omics Analysis. Trends Biotechnol 2018;36:488–98. https://doi.org/10.1016/J.TIBTECH.2018.01.013.
[135] Deep Learning n.d. https://mitpress.mit.edu/9780262035613/deep-learning/ (accessed April 23, 2023).
[136] Veta M, Heng YJ, Stathonikos N, Bejnordi BE, Beca F, Wollmann T, et al. Predicting breast tumor proliferation from whole-slide images: The TUPAC16 challenge. Med Image Anal 2019;54:111–21. https://doi.org/10.1016/J.MEDIA.2019.02.012.
[137] Sirinukunwattana K, Pluim JPW, Chen H, Qi X, Heng PA, Guo YB, et al. Gland Segmentation in Colon Histology Images: The GlaS Challenge Contest. Med Image Anal 2016;35:489–502. https://doi.org/10.1016/j.media.2016.08.008.
[138] Yagi Y. Color standardization and optimization in Whole Slide Imaging. Diagn Pathol 2011;6:S15. https://doi.org/10.1186/1746-1596-6-S1-S15.
[139] Tellez D, Litjens G, Bándi P, Bulten W, Bokhorst JM, Ciompi F, et al. Quantifying the effects of data augmentation and stain color normalization in convolutional neural networks for computational pathology. Med Image Anal 2019;58:101544. https://doi.org/10.1016/J.MEDIA.2019.101544.
[140] Ho SY, Phua K, Wong L, bin Goh WW. Extensions of the External Validation for Checking Learned Model Interpretability and Generalizability. Patterns 2020;1:100129. https://doi.org/10.1016/J.PATTER.2020.100129.