World Journal of Urology
https://doi.org/10.1007/s00345-019-03000-5

TOPIC PAPER

Current and future applications of machine and deep learning in urology: a review of the literature on urolithiasis, renal cell carcinoma, and bladder and prostate cancer

Rodrigo Suarez-Ibarrola[1] · Simon Hein[1] · Gerd Reis[2] · Christian Gratzke[1] · Arkadiusz Miernik[1]

Received: 24 July 2019 / Accepted: 25 October 2019

Abstract

Purpose The purpose of the study was to provide a comprehensive review of recent machine learning (ML) and deep learning (DL) applications in urological practice. Numerous studies have reported their use in the medical care of various urological disorders; however, no critical analysis has been made to date.

Methods A detailed search of original articles was performed using the PubMed MEDLINE database to identify recent English literature relevant to ML and DL applications in the fields of urolithiasis, renal cell carcinoma (RCC), bladder cancer (BCa), and prostate cancer (PCa).

Results In total, 43 articles were included addressing these four subfields. The most common ML and DL application in urolithiasis is in the prediction of endourologic surgical outcomes. The main area of research involving ML and DL in RCC concerns the differentiation between benign and malignant small renal masses, Fuhrman nuclear grade prediction, and gene expression-based molecular signatures. BCa studies employ radiomics and texture feature analysis for the distinction between low- and high-grade tumors, address accurate image-based cytology, and use algorithms to predict treatment response, tumor recurrence, and patient survival. PCa studies aim at developing algorithms for Gleason score prediction, MRI computer-aided diagnosis, and surgical outcomes and biochemical recurrence prediction. Studies consistently found the superiority of these methods over traditional statistical methods.

Conclusions The continuous incorporation of clinical data, further ML and DL algorithm retraining, and generalizability of models will augment the prediction accuracy and enhance individualized medicine.

Keywords Artificial intelligence · Machine learning · Deep learning · Artificial neural network · Convolutional neural network · Prostate cancer · Bladder cancer · Renal cell carcinoma · Urolithiasis

Introduction

The term artificial intelligence (AI) commonly refers to the computational technologies that mimic or simulate intellectual processes typical of human cognitive function, such as reasoning, learning, and problem solving [1]. AI is a branch of computer science and part of a multidisciplinary approach adopting principles from the fields of mathematics, logic, computation, and biology in an attempt to build intelligent entities often represented as software programs [2, 3]. Given its broad, dynamic, and expanding computational power, AI has been revolutionizing and reshaping our health-care systems, allowing physicians to improve their ability to perform medical tasks [2]. As the medical community´s understanding and acceptance of AI grows, so does our imagination in ways to improve diagnostic accuracy, expedite clinical processes, and decrease human resource costs by assisting medical professionals in what once were time-consuming problems [4].

Machine learning (ML) is a subfield of AI involving the development and deployment of dynamic algorithms to analyze data and facilitate the identification of intricate patterns [5]. ML tends to improve or 'learn' as more

✉ Rodrigo Suarez-Ibarrola
rodrigo.suarez@uniklinik-freiburg.de

[1] Department of Urology, Faculty of Medicine, University of Freiburg-Medical Centre, Hugstetter Str. 55, 79106 Freiburg, Germany

[2] Department Augmented Vision, German Research Center for Artificial Intelligence, Kaiserslautern, Germany

Published online: 05 November 2019

data are incorporated by using decision trees to explicitly learn decision rules [6]. Recent advancements in AI have also been driven by deep learning (DL), which involves the training of artificial neural networks (ANN) with multiple layers on large datasets [7]. An ANN is a collection of individual information processing units or "artificial neurons" that are arranged and interconnected in network architectural layers to perform computational tasks and recognize complex patterns [8]. They are trained with input/output tuples where each input has a specific output assigned to arrange concepts or functions beyond the means of traditional statistical analysis methods. A frequently used ANN that is particularly efficient when applied to pattern recognition in digitized images is the deep convolutional neural network (DCNN). The depth and width of the network determine the complexity and 'learning potential' of the network. In health care, ML and DL have been increasingly and successfully applied to preventive medicine, image recognition diagnostics, personalized medicine, and clinical decision-making.

The aim of this review article is to address recent ML and DL applications in urolithiasis, renal cell carcinoma, bladder cancer, and prostate cancer to predict patient outcomes. Their utilization other than in these subfields is not in the scope of this work.

Methodology

A comprehensive review of current literature was performed using the PubMed-Medline database up to May 2019 using the term "urology", combined with one of the following terms: "machine learning", "deep learning" and "artificial neural network" in combination with "urolithiasis", "renal cell carcinoma", "bladder cancer", and "prostate cancer". To capture recent trends in ML and DL applications, the search was limited to articles published within the last 5 years, originally published in English. Review articles and editorials were excluded. Publications relevant to the subject and their cited references were retrieved and appraised independently by two authors (R.S. and A.M.). In accordance with the PRISMA criteria, Fig. 1 was included to delineate our article selection process. After full text evaluation, data were independently extracted by the authors for further assessment of qualitative and quantitative evidence synthesis. The following information was extracted from each study: name of author, journal and year of publication, AI method, number of participants per study, and outcome prediction accuracy.

Urinary stone disease

Despite novel instrumental advancements in urinary stone surgery, decision-making and patient counseling remain a challenge for clinicians. Several investigators have studied the prognostic role of preoperative parameters on surgical treatment outcomes in terms of stone-free rate (SFR) and the need for secondary procedures [9–13]. An accurate preoperative outcome prediction would assist urologists in optimizing patient selection, choosing ideal treatment options, and personalize patient counselling.

Kadlec et al. developed an ANN that predicted outcomes after various forms of endourologic intervention [9]. Input variables and outcome data from 382 endourologically treated renal units were used to assess SFR (defined by no visible stone on KUB or <4 mm on CT) and need for secondary procedures. The model predicted SFR with 75.3% sensitivity and 60.4% specificity, and the need for a secondary procedure with 30% sensitivity and 98.3% specificity, yielding a positive and negative predictive value of 60% and 94.2%, respectively. Aminsharifi et al. trained an ANN with pre- and postoperative data from 200 patients and used it to predict various outcomes for 254 patients after percutaneous nephrolithotomy (PCNL) [10]. The accuracy and sensitivity to predict SFR, blood transfusion, and post-PCNL ancillary procedures ranged from 81.0 to 98.2%. Stone burden and morphometry received the highest weight by the ANN as preoperative characteristics affecting postoperative outcomes. Recently, Choo et al. developed and validated a decision-support model using ML algorithms to predict treatment success after a single-session shock wave lithotripsy (SWL) in ureteral stone patients [11]. Using data from 791 patients, a model constructed with 15 variables exhibited 92.3% accuracy to predict SWL outcome. In the decision tree analysis, stone volume, length, and Hounsfield units were the top three most important preoperative variables. Similarly, Seckiner et al. collected data from 203 patients and developed an ANN to predict SFR and support SWL treatment planning [14]. ANN analysis demonstrated a prediction accuracy of 99.3% for SFR in the training group, 85.5% in the validation group, and 88.7% in the test group.

Other studies have investigated computer-assisted detection using image features for supporting radiologists in identifying stones. Längkvist et al. developed a DCNN to differentiate ureteral stones from phleboliths in thin slice CT volumes due to their similarity in shape and intensity [15]. The DCNN was evaluated on a database consisting of 465 clinically acquired abdominal CT scans of patients suffering from suspected renal colic. The model achieved 100% sensitivity and an average of 2.68 false positives per patient on a test set of 88 scans. Kazemi et al. derived an ANN for the early detection of kidney stone type and most influential parameters to provide a decision-support system [16]. Information pertaining to 936 patients who underwent treatment for kidney stones was collected and included 42 image features. The model resulted in 97.1% accuracy for predicting kidney stone type and identified gender, calcium level, uric acid condition, hypertension, diabetes, nausea/

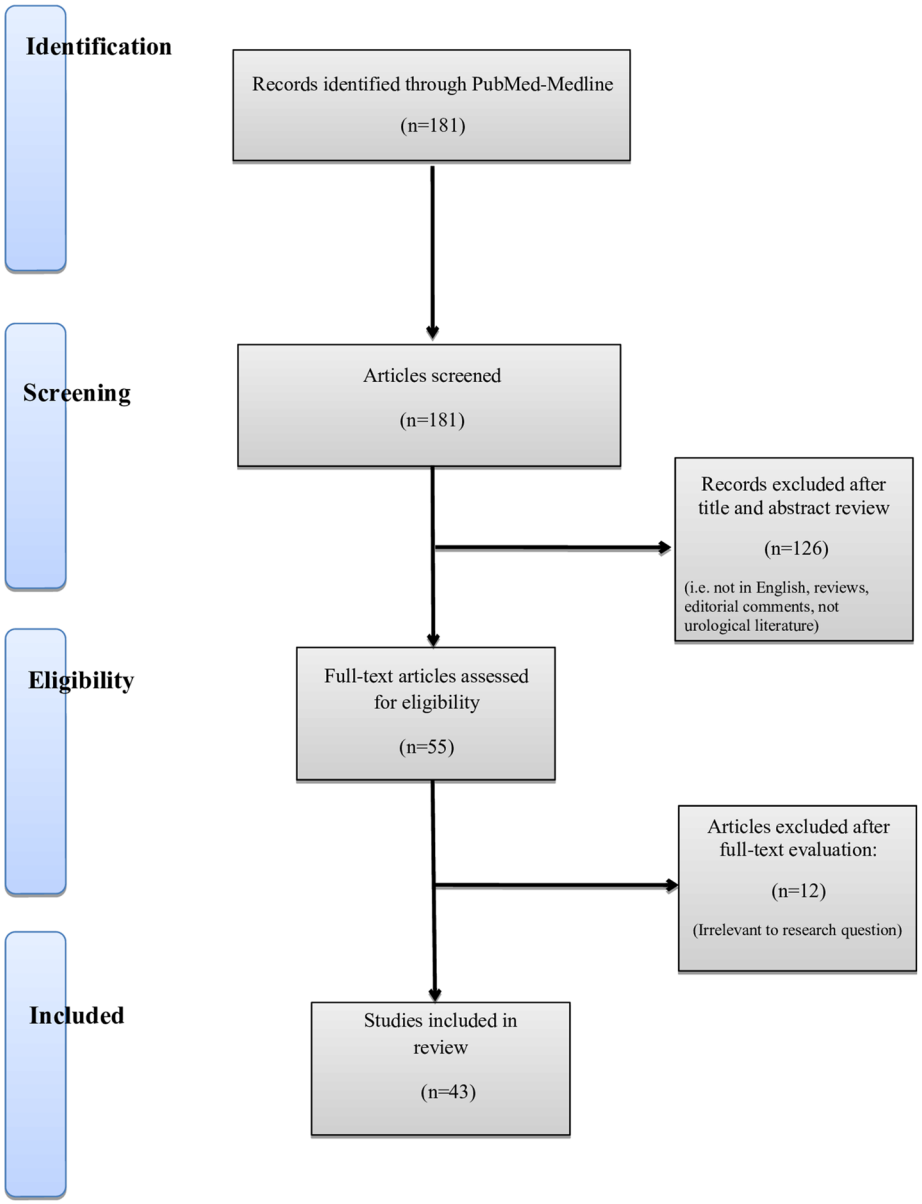

Fig. 1 Summary of study selection process

vomiting, flank pain, and urinary tract infection as the most vital parameters for predicting the chance of nephrolithiasis.

A future common goal is for ANNs to be exchanged between institutions to overcome the limitation of having networks trained with data from just one center. To this extent, ML and DL methods hold promise for multi-institutional dataset expansion in national registries and the development of future predictive nomograms. Table 1 provides a summary of studies using ML and DL methods applied in various modalities of urolithiasis therapy.

Renal cell carcinoma

The incidence of renal cell carcinoma (RCC) has steadily increased over the past decades as a result of incidental small renal mass (SRM) detection via cross-sectional imaging [17]. Surgical series have shown that 20–30% of SRMs ≤ 4 cm are benign, while 20% exhibit potentially aggressive behavior [18]. However, there are currently no clinical or radiographic features that accurately predict histologic analysis. Magnetic resonance imaging (MRI) and computed tomography (CT) have been employed in an attempt to noninvasively differentiate tumors by their degree, pattern, and heterogeneity of enhancement. While promising, these approaches remain suboptimal as clinical tools for differentiating SRMs. Recently, powerful ML algorithms are being used to explore complex interactions in clinical and imaging data to provide diagnosis, prognosis, treatment planning and assist in shared decision-making.

Table 1 Machine and deep learning applications in urolithiasis

Author	AI method	Patients	Outcome prediction accuracy
Kadlec et al. Urolithiasis 2014 [8]	Outcome prediction in SWL, URS, PCNL	382 treated renal units	SFR sensitivity 75.3% and specificity 60.4% (PPV 75.3% and NPV 60.4%) SP sensitivity 30% and specificity 98.3% (PPV 60% and NPV 94.2%)
Aminsharifi et al. J Endourol 2017 [9]	ANN prediction of outcome variables in PCNL	Data from 200 patients to predict 254 subsequent patients	Overall SFR was 76.4%. 54 out of 254 (21.3%) patients required ancillary procedures (SWL 5.9%, URS 10.6% and repeat PCNL 4.7%) Accuracy and sensitivity of the system in predicting different postoperative variables ranged from 81 to 98.2%
Choo et al. J Urol 2018 [10]	ML prediction of SWL success in ureteral stones	791 patients with ureteral stones	Applying decision tree analysis predicted treatment outcome with 92.29% accuracy after single session and ROC AUC of 0.951
Mannil et al. Abdom Radiol 2018 [11]	ML texture analysis for successful SWL	34 urinary stones in in vitro setting	Accuracy of AUC 0.84, sensitivity of 94% and specificity of 59% for successful stone disintegration
Mannil et al. J Urol 2018 [12]	ML prediction of SWL success in renal stones	51 patients with untreated renal stones	SWL success based on SSD, BMI, and stone size showed AUCs of 0.63, 0.68 and 0.58, respectively. Combining ML 3D texture analysis with clinical variables improved accuracy to an AUC of 0.85 for SSD, 0.8 for BMI, and 0.81 for stone size
Seckiner et al. Int Braz J Urol 2017 [13]	ANN predicting SFR after SWL	203 patients with renal stones	ANN analysis demonstrated that the prediction accuracy of the stone-free rate was 99.25% in the training group, 85.48% in the validation group, and 88.70% in the test group
Längkvist et al. Comput Biol Med 2018 [14]	CAD detection of ureteral stones using CT volumes on CNN	465 clinically acquired high-resolution images of the urinary tract	The best model using 2.5D input data and anatomical information achieved a sensitivity of 100% and an average of 2.68 false positives per patient on a test set of 88 scans
Kazemi et al. Artif Intell Med 2018 [15]	ANN for predicting kidney stone type	936 patients and 42 clinical features	Model predicted chance of developing nephrolithiasis with 97.1% accuracy. Parameters such as gender, hypertension, diabetes, nausea, vomiting, flank pain, urinary tract infection, and calcium and uric acid levels were the most vital for predicting the chance to develop nephrolithiasis

AI artificial intelligence, *ML* machine learning, *ANN* artificial neural network, *CNN* convolutional neural network, *SWL* shock wave lithotripsy, *fURS* flexible ureteroscopy, *PCNL* percutaneous nephrolithotomy, *SFR* stone-free rates, *SP* secondary procedures, *SSD* stone-to-skin distance, *BMI* body mass index, *AUC* area under the curve, *CAD* computer-aided detection

Given the limitations of conventional medical imaging, there has been increasing interest in radiomics, which involves automatically extracting quantitative features from medical images. Radiomics may provide a novel approach to develop predictive tools by correlating imaging features to tumor characteristics including histology, tumor grade, genetic patterns and molecular phenotypes, as well as clinical outcomes in patients with renal masses. Pixel distribution and pattern-based texture analysis have emerged as practical quantitative methods to build image processing algorithms for the detection of tissue differences that cannot be determined by subjective visual assessments [19].

Several studies have shown that texture analysis has potential in differentiating SRM [20–22]. Yan et al. showed that texture analysis may be a reliable quantitative strategy to differentiate between angiomyolipoma (AML), clear cell RCC (ccRCC), and papillary RCC (pRCC) with an accuracy in the range of 90.7–100% based on the analysis of three-phase CT scans [23]. Feng et al. achieved a higher accuracy and area under the curve (AUC) of 93.9% using a similar ML strategy [24]. Cui et al. proposed an automatic computer-identification system to differentiate AML from whole-tumor CT images using an over-sampling technique to increase the sample volume of AML [25]. Yu et al. evaluated the utility of texture analysis for the distinction of renal tumors, including various RCC subtypes and oncocytoma. The ability of ML to distinguish ccRCC and pRCC from oncocytoma was excellent with AUC of 0.93 and 0.99, respectively [22]. Coy et al. investigated the diagnostic value and feasibility of a DL-based renal lesion classifier to differentiate ccRCC from oncocytoma in 179 patients with pathologically confirmed renal masses on routine four-phase multiple detector CT scans [26]. When using the entire tumor volume, the excretory phase showed the best classification performance with 74.4% accuracy, 85.8% sensitivity, and PPV of 80.1%.

Furthermore, the nuclear grade of a tumor is widely recognized as one of the most important independent prognostic factors [27]. Determination of the Fuhrman grade by percutaneous renal biopsy suffers from significant sampling bias, making the preoperative recognition of biological aggressiveness challenging. Studies have shown that ML models constructed from CT imaging texture features can accurately distinguish between ccRCC high and low grades, with accuracy ranging from 0.73 to 0.93 [19, 28–32]. Ding et al. showed high prediction accuracy in identifying ccRCC grade and their results were superior to those obtained from CT image features or the RENAL nephrometry score for high- and low-grade ccRCC predictions [29].

In recent years, biomarkers and multiple gene expression-based signatures have been developed to predict survival and disease prognosis in ccRCC. Li et al. developed a prognostic model based on 15 survival-related genes from The Cancer Genome Atlas and showed that patients in the model´s high-risk group had significantly worse survival than those in the low-risk group. Risk group was independent of age and sex, but was significantly associated with hemoglobin level, primary tumor size, and grade [33]. Radiogenomics is a field investigating the potential associations between a disease's imaging features and the underlying genetic patterns or molecular phenotype. Kocak et al. evaluated the potential of quantitative CT scan texture analysis to predict the presence of PBRM1 mutations, which is the second most commonly identified mutation in ccRCC, using ANNs and ML algorithms [34]. Overall, the ANN correctly classified 88.2% of ccRCC with regard to PBRM1 mutation status, while the random forest ML algorithm correctly classified 95% of ccRCC.

These are promising results for developing noninvasive imaging biomarkers of histopathologic subtypes, prognosis, and treatment response. Moreover, they demonstrate that noninvasive ML and DL models constructed from radiomics features have comparable performance to percutaneous renal biopsy in predicting the International Society of Urological Pathology (ISUP) grading. Accurate preoperative nuclear grading may substantially aid in risk assessment, patient stratification, and treatment planning of RCC patients. Table 2 summarizes the most significant findings of MDL applications in the field of RCC.

Bladder cancer

The diagnosis and tumor staging of bladder cancer (BCa) ultimately depend on cystoscopic examination of the bladder and histological evaluation of sampled tissue by transurethral resection (TURB). The main limitation of cystoscopy is its difficulty in discriminating between areas of malignancy and healthy urothelium given the multifocal nature of the disease and inconspicuous but significant lesions such as CIS. However, CT/MRI image-based 3D texture feature analysis of the bladder wall has demonstrated its potential as a non-invasive, image-based strategy to accurately identify heterogeneous tumor distribution and preoperatively discriminate BCa from normal wall tissue [35]. MRI textural features extracted from cancerous volumes of interest and incorporated into ML models have further demonstrated their ability to preoperatively distinguish low- and high-grade BCa with 83% accuracy [36]. DCNNs have also been used to classify and predict cystoscopic findings with a high degree of accuracy [37]. Such a DL model can be integrated into an AI-aided imaging diagnostic tool to support urologists during cystoscopic examinations. 'AI cystoscopy' may serve as an adjunct during surgical training and medical education to help differentiate benign from malignant lesions using visual evaluation and thereby reduce the number of unnecessary biopsies. A different approach for image-based diagnosis

Table 2 Machine and deep learning applications in renal cell carcinoma

Author	Patients	AI method	Parameters	Predicted outcome accuracy
Bektas et al. Eur Radiol 2018 [18]	53 RCC patients with 54 pathologically proven ccRCCs	Various ML classifiers[a]	279 texture features extracted from contrast-enhanced CT scans	Fuhrman nuclear grade prediction SVM model scored highest with overall accuracy, sensitivity for detecting high-grade cc-RCCs, specificity for detecting high-grade cc-RCCs, and overall AUC of 85.1%, 91.3%, 80.6%, and 0.860, respectively
Kocak et al. Eur J Radiol 2018 [19]	68 RCC patients	ANN and ML-based SVM[a] classifiers	275 texture features extracted and analyzed from 3-phase CT images	Distinguishing between three major subtypes of RCC ANN discrimination of non-ccRCC from ccRCC subtypes with an external validation accuracy, sensitivity, and specificity of 84.6%, 69.2%, and 100%, respectively SVM discrimination of pRCC from other RCC subtypes with an external validation accuracy, sensitivity, and specificity of 69.2%, 71.4%, and 100%, respectively
Kunapuli et al. J Digit Imaging 2018 [20]	150 post-resected patients	Various ML radiomics-based models	204 texture features extracted from preoperative 4-phase CT images	Identifying malignancy of renal masses Visual classification by experts achieves an AUC of 0.65 using pathology as gold standard. RFGB model outperformed several conventional ML algorithms ($p = 0.05$) with an accuracy of 0.82, and a precision and recall of 0.87
Yu et al. Abdom Radiol 2017 [21]	119 RCC patients	CT texture analysis with ML-based SVM classifier	Tumor images from 4-phase CT manually segmented and 43 texture features analyzed	TA for renal tumor differentiation Excellent discriminators of tumors were identified with AUC of 0.91 and 0.93 ($p < 0.0001$), respectively for differentiating ccRCC from oncocytoma. AUC of 0.99 ($p < 0.0001$) for differentiating pRCC from oncocytoma and an AUC of 0.92 for differentiating oncocytoma from other tumors. The ability of ML to distinguish ccRCC from other tumors and pRCC from other tumors was excellent with AUC of 0.91 and 0.92, respectively

Table 2 (continued)

Author	Patients	AI method	Parameters	Predicted outcome accuracy
Yan et al. Acad Radiol 2015 [22]	18 AML, 18 ccRCC and 14 pRCC pathologically proven patients	ANN classifier	Images from 3-phase CT scan analyzed with TA software	TA to differentiate between angiomyolipoma, ccRCC, and pRCC Excellent classification results (0–9.3% error) were obtained for all three groups, independently of CT phase used
Feng et al. Eur Radiol 2018 [23]	58 patients with SRM: 17 AML and 41 RCC	ML-based SMV classifier for quantitative TA	42 manually segmented texture features extracted from 3-phase CT scan tumor regions of interest	Differentiating angiomyolipoma from RCC Of 42 extracted features, 16 showed significant intergroup differences ($p < 0.05$). The SMV-RFE + SMOTE classifier achieved the best performance in discriminating between small AML and RCC, with the highest accuracy, sensitivity, specificit,y and AUC of 93.9%, 87.8%, 100% and 0.955, respectively
Cui et al. Acta Radiol 2019 [24]	171 pathologically proven renal masses	ML-based SMV established differentiation classifiers	Texture features extracted from whole tumor images in 3-phase CT scans	Differentiation of angiomyolipoma from all RCC subtypes Differentiating AML from all-RCC (AUC = 0.96) and ccRCC (AUC 0.97) was higher than AML from non-ccRCC (AUC = 0.89). Morphological interpretation by radiologists achieved lower performance differentiating AML from all-RCC (AUC = 0.067), ccRCC (AUC = 0.68), and non-ccRCC (AUC = 0.64)
Coy et al. Abdom Radiol 2019 [25]	179 patients with 179 renal lesions	ANN trained with 4000 iterations	Feature extraction representative of the entire renal mass	Differentiating oncocytoma from ccRCC Excretory phase of entire tumor volume achieved highest 74.4% accuracy, 85.8% sensitivity, and 80.1% PPV. When combined with tumor mid-slices a PPV of 82.5%

Table 2 (continued)

Author	Patients	AI method	Parameters	Predicted outcome accuracy
Holdbrook et al. JCO Clin Cancer Inform 2018 [27]	Histopathologic slides of 59 ccRCC patients	ML-based SMV and logistic regression models	Image-based texture feature extraction for nuclear analysis	Quantification of nuclear pleomorphic patterns of clear cell RCC Objective and fully automated process to analyze pathologic slides. Automated image classification pipeline score correlated ($R=0.59$) with existing multigene assay-based scoring system which has demonstrated to be a strong indicator of prognosis
Ding et al. Eur J Radiol 2018 [28]	114 ccRCC patients treated by partial or total nephrectomy	ML models for non-texture and texture feature variables	Six non-texture features: pseudocapsule, roundness, max. diameter, intratumoral artery, enhancement value, and relative enhancement. 184 texture features extracted from 3-phase CT scans	High-grade ccRCC prediction Texture score-based models with or without non-texture features showed a significant discrimination of the high- from low-grade ccRCC ($p<0.05$)
Kocak et al. AJR 2019 [29]	81 ccRCC patients (56 high and 25 low grade)	ANN and binary logistic regression models	744 texture features extracted from unenhanced CT scan images	ccRCC histopathologic nuclear grade prediction ANN correctly classified 81.5% of ccRCCs (AUC = 0.714). Logistic regression correctly classified 75.3% of all tumors (AUC = 0.656). ANN outperformed logistic regression model ($p<0.05$)
Lin et al. Abdom Radiol 2019 [30]	231 patients with 232 pathologically proven lesions	Various ML classifiers	134 texture features extracted from 3-phase CT images. Single and 3-phase CT images compared with each other	ccRCC Fuhrman grade prediction ML model based on 3-phase CT images achieved best diagnostic performance for differentiating low- from high-grade ccRCC (AUC = 0.87), followed by single nephrogenic phase (AUC = 0.84), corticomedullary phase (AUC = 0.80), and pre-contrast phase (AUC = 0.82)
Sun et al. Medicine 2019 [31]	227 patients with ISUP confirmed ccRCC	ML-based SMV model	1029 imaging features extracted from regions of interest extracted from 3-phase CT scan images	Prediction of ISUP grading of ccRCC SVM model effectively distinguishes between high- and low-grade ccRCC. AUC training and validation values were 0.88 and 0.91, respectively

Table 2 (continued)

Author	Patients	AI method	Parameters	Predicted outcome accuracy
Li et al. Medicine 2018 [32]	15 survival- selected genes from TCGA dataset	ML-based random forest[a] variable hunting	Risk score model to predict survival of ccRCC patients in the TCGA dataset	Fifteen gene expression survival prediction Patients in the high-risk group had significantly worse overall survival than those in the low-risk group ($p = 5.6e{-}16$); recurrence-free survival showed a similar pattern. Risk was independent of age and sex, but was significantly associated with hemoglobin level, primary tumor size, and grade
Kocak et al. AJR Am J Roentgenol 2019 [33]	45 ccRCC patients: 16 with and 29 without PBRM1 mutation	ML-based high-dimensional quantitative CT texture analysis	828 texture features extracted per lesion	PBRM1 mutation prediction in ccRCC 759 texture features (91.7%) had excellent reproducibility. ANN algorithm correctly classified 88.2% of ccRCCs in terms of PBRM1 mutation status. The sensitivity, specificity, and precision for predicting the ccRCCs with the *PBRM1* mutation were 87.8%, 88.5%, and 86.7%, respectively. Random forest algorithm correctly classified 95% of ccRCCs with AUC value of 0.987. The sensitivity, specificity, and precision for predicting the ccRCCs with the *PBRM1* mutation were 94.6%, 95.4%, and 94.6%, respectively

ccRCC clear cell renal cell carcinoma, *pRCC* papillary RCC, *AML* angiomyolipoma, *SRM* small renal mass, *TA* texture analysis, *CT* computed tomography, *ML* machine learning, *SMV* support vector machine, *ANN* artificial neural network, *ISUP* International Society of Urological Pathology, *TCGA* The Cancer Genome Atlas

[a]Classifiers in ML refer to the insertion of a new observation into the appropriate category among others that were based on trained data sets of known observations. Support vector machines are supervised learning methods for classification that learn the optimal difference between features of each class. Random forest is supervised learning methods for classification that is based on decision trees

has focused on the nanoscale-resolution scanning of cell surfaces collected from urine [38]. Atomic force microscopy coupled to ML analysis has been shown as a noninvasive method to detect BCa with 94% accuracy when five cells per patient's urine sample are examined. Moreover, it demonstrated a statistically significant improvement in diagnostic accuracy compared to cystoscopy alone. ML-based methods have been further applied to accurately quantify tumor buds from immunofluorescence-labeled slides of muscle-invasive BCa (MIBC) patients [39]. Tumor budding was found to correlate with TNM staging and patients of all stages were stratified into three new staging criteria based on disease-specific death. Tumor bud quantification through automated slide analysis may provide an alternate staging model with prognostic value for MIBC patients.

ML algorithms have been employed to create recurrence and survival predictive models from imaging and operative data. Patient recurrence and survival at 1, 3 and 5 years after cystectomy was predicted with greater than 70% sensitivity and specificity [40]. Such predictive models may help dictate patients´ follow-up schedules, adjuvant treatments, and also provide opportunities for improving care by optimally utilizing operative data collection. ML algorithms used to identify genes at initial presentation that are most predictive of recurrence can be applied as molecular signatures to predict the risk of recurrence within 5 years after TURB [41]. Whole genome profiling from frozen non-muscle-invasive BCa specimens was integrated into a genetic programming algorithm to generate classifier mathematical models for outcome prediction. The model identified 21 key genes that are associated with recurrence from which an optimal three-gene rule [TMEM205 × (NFKBIA × KRT17)] was developed to predict recurrence with 70.6% sensitivity and 66.7% specificity on the test set.

An unmet need in BCa treatment is the early assessment of chemotherapeutic efficacy and prediction of treatment failure at an early phase to reduce unnecessary morbidity, improve patients´ quality of life, and reduce costs. Therefore, the development of accurate predictive models to determine the effectiveness of neoadjuvant chemotherapy is of critical importance in BCa management. Computerized decision support systems (CDSS) have been developed to provide noninvasive, objective, and reproducible decision support for identifying non-responders, so that treatment may be suspended early to preserve their physical condition or to distinguish full responders for organ preservation. Wu et al. compared the performance of different DCNN models and showed that they effectively predicted a bladder lesion´s response to chemotherapy and compared favorably to radiologists´ performance [42]. Cha et al. developed a CT-based CDSS to improve the identification of patients who responded completely to neoadjuvant chemotherapy and found that physicians´ diagnostic accuracy significantly increased with the aid of CDSS [43]. Thus, computer-aided treatment prediction using DL algorithms may prove to be invaluable to medical professionals as a decision support tool for improving the selection of patients considering bladder-sparing therapy for MIBC and avoiding adverse effects in non-responders.

Despite several ML and DL research efforts in predicting BCa patients' outcomes, there is scarce adoption of such models in clinical practice. The main challenges ahead before such models can be deployed successfully in a clinical setting are the inclusion of standardized parameters, adjusting for the equipment variance, and the collection of multi-institutional data to ensure the generalizability of the models. Once these issues are addressed, ML and DL models can be trained using BCa datasets to accurately predict an individual patient´s outcome using pre-, peri-, and postoperative data. Table 3 summarizes the most significant findings of MDL applications in the field of BCa.

Prostate cancer

There is an unmet need for definitive diagnosis besides transrectal imaging and biopsy for men with prostate cancer (PCa). Although a biopsy is necessary for a conclusive diagnosis, patients with low cancer risk could avoid this procedure due to the potential complications that may arise. To achieve this goal, prediction models have been developed to determine patients' cancer risk on the basis of clinical characteristics. Multilayer ANNs have predicted patients' prostate biopsy results more accurately while assessing large numbers of variables than traditional statistical methods ranging from non-linear relationships to logistic regression [44, 45]. Despite MRI having improved PCa detection and thereby reducing the number of unnecessary biopsies, excessive variation in its performance and interpretation is a major barrier for global standardization. Computer-aided diagnostic (CAD) systems with DL architecture have been applied to diminish variation in the interpretation of prostatic MRI. Among the advantages of this approach are consistent diagnoses, cost-effectiveness, and improved efficiency. Ishioka et al. developed DCNN algorithms that estimate the area in which a targeted biopsy may detect the presence of cancer and in its execution decrease the number of patients mistakenly diagnosed as having cancer [46]. However, other studies have shown no added benefit of radiomic ML when compared with mean apparent diffusion coefficient in differentiating benign versus malignant prostate lesions [47]. Provided that the diagnostic precision of CAD systems exceeds that attained by humans, and the pathological diagnosis can be predicted with high accuracy, it may be reasonable to confirm clinically significant PCa solely based on MRI images rather than with biopsy.

Table 3 Machine and deep learning applications in bladder cancer

Author	AI method	Patients	Parameters	Predicted outcome accuracy
Xu et al. Int J CARS 2017 [34]	3D texture features in MRI	62 BCa patients	VOIs extracted from T2 MRI datasets	3D texture analysis to discriminate bladder tumors From each VOI, 58 texture features were derived. 37 features showed significant interclass differences ($p \leq 0.01$). 29 optimal features by RFE-SVM resulted in sensitivity, specificity, accuracy, and AUC of 0.90, 0.85, 0.88 and 0.90, respectively
Zhang et al. J Magn Reson Imaging 2017 [35]	ML-based SVM classification	61 BCa patients	DWI and ADC maps to distinguish low- from high-grade Bca	Textural features distinguish low-grade from high-grade Bca From each VOI, 102 features were derived. 47 features showed significant inter-group difference ($p \leq 0.05$). 22 optimal features by SVM-RFE resulted in best performance in bladder cancer grading with sensitivity, specificity, accuracy, and AUC of 0.78, 0.87, 0.83 and 0.86, respectively
Eminaga et al. JCO 2018 [36]	CNN models	479 patients with 44 cystoscopic findings	18,681 cystoscopy images	Diagnostic classification of cystoscopic images Over 99% accuracy for classifying images was achieved by four models. 7.68% of images showing bladder stones with indwelling catheter and 1.43% of diverticulum images were falsely classified
Sokolov et al. PNAS 2018 [37]	Atomic force microcopy and ML analysis	43 controls and 25 cases	Surface parameters of cells collected from urine	Detection of bladder cancer in voided urine cytology using atomic force microscopy 94% diagnostic accuracy when examining five cells per patient´s urine sample. A statistically significant improvement ($p < 0.05$) in diagnostic accuracy of AFM (AUC = 0.92) compared to cystoscopy (AUC = 0.77)

Table 3 (continued)

Author	AI method	Patients	Parameters	Predicted outcome accuracy
Brieu et al. Sci Rep 2019 [38]	ML-based methodology	100 MIBC patients	Tumor bud quantification across immunofluorescence labeled slides	ML tumor bud quantification for TNM staging in MIBC Tumor budding correlated to TNM ($p=0.0009$) and pT staging ($p=0.008$). Decision support tree model reported that tumor budding was the most significant feature (HR = 2.59, $p=0.009$)
Hasnain et al. PLoS One 2019 [39]	ML-based SMV model	3499 post-cystectomy patients	Pre- and operative data Adjuvant therapy cycles	Post-cystectomy 1-, 3-, and 5-year recurrence and survival prediction Patient recurrence and survival 1, 3, and 5 years after cystectomy can be predicted with greater than 70% sensitivity and specificity. Pathologic stage subgroup (organ confined, extravesica,l and node+) and TNM stage were ranked the most important predictors
Bartsch et al. J Urol 2016 [40]	Whole genome profiling	112 NMIBC specimens	Key genes involved in BCa recurrence	Gene expression profiling to predict recurrent NMIBC 21 genes predicted recurrence. A 5-gene algorithm yielded 77% sensitivity and 85% specificity to predict recurrence in a training set, and 69% and 62%, respectively, in a test set. Singular 3-gene predicted recurrence with 80% sensitivity and 90% specificity in a training set, and 71% and 67%, respectively, in a test set
Wu et al. Tomography 2019 [41]	DL-CNN	CT scans of 123 patients with 129 cancer foci	Extraction of ROIs from segmented lesions of pre- and posttreatment scans	Assessment of BCa treatment response using a DL approach The base DL-CNN with pre-trained weights and transfer learning achieved test AUC of 0.79. The radiologists' AUCs were 0.76 and 0.77. DL-CNN performed better with pre-trained than randomly initialized weights

Table 3 (continued)

Author	AI method	Patients	Parameters	Predicted outcome accuracy
Cha et al. Acad Radiol 2018 [42]	DL-CNN	CT scans of 123 subjects with 157 MIBC foci	Radiomics feature-based analysis	CDSS for predicting BCa treatment response The mean AUCs for assessment of T0 disease were 0.80 for CDSS alone, 0.74 for physicians not using CDSS, and 0.77 for physicians using CDSS. The increase in the physicians´ performance was statistically significant ($p < 0.05$)

BCa bladder cancer, *RC* radical cystectomy, *MIBC* muscle-invasive bladder cancer, *NMIBC* non-muscle-invasive bladder cancer, *ML* machine learning, *DL* deep learning, *CNN* deep convolutional neural network, *SVM* support vector machine, *VOI* volume of interest, *ROI* region of interest, *DWI* diffuse-weighted images, *ADC* apparent diffusion coefficient, *AUC* area under the curve, *CDSS* computer decision support system

Automated computational methods applied to digital pathology images have shown the ability to overcome Gleason score ambiguity, convey reproducible results, and generate large amounts of data. Arvanati et al. trained a DCNN as Gleason score annotator and used the model's predictions to assign patients into low-, intermediate-, and high-risk groups, achieving pathology expert-level stratification results [48]. Accurate post-surgical risk stratification is essential to identify patients at high risk of PCa-specific mortality who would benefit from early intervention. Donovan et al. introduced an innovative platform which accurately discriminates between low-, intermediate-, and high-risk PCa, and predicts the likelihood of significant clinical failure within 8 years [49]. By combining ML-guided image analysis with biological attributes, the authors provided a risk assignment that is unbiased, broadly applicable, and independent of interpretive histology.

While information from clinical registry data assists physicians to make data-driven decisions, there is limited opportunity for patients to access these registries to help them make informed decisions. Auffenberg et al. utilized data from a prospective cancer registry comprising 7543 men diagnosed with PCa to train an ML model to help newly diagnosed men to view predicted treatment decisions based on patients with similar characteristics [50]. Their personalized model was highly accurate with age, followed by number of positive cores and Gleason score resulting as the most important variables that influenced patient treatment decisions.

Treatment response prediction using MRI images has been shown as an efficient clinical decision-making tool. Abdollahi et al. developed various radiomics models based on pre- and post-intensity-modulated radiotherapy (IMRT) MRI data for individualized treatment response prediction in PCa patients [51]. Their results showed that the features extracted from pre-treatment MRI images predicted early IMRT response with reliable performance. Moreover, Hung et al. presented a novel ML method of processing automated performance metrics to evaluate surgical performance and predict clinical outcomes after robot-assisted radical prostatectomy (RARP) [52]. Their model predicted length of hospital stay, operative time, Foley catheter duration, and urinary continence with over 85% accuracy [53]. In a recent study, Wong et al. used three ML algorithms for the prediction of early biochemical recurrence and showed with an AUC > 0.95 to outperform traditional statistical regression models [54]. Such methodology can be employed as potentially more accurate for identifying patients at risk and equip patients and physicians alike with prognostic information to provide individualized health care (Table 4).

Table 4 Machine and deep learning applications in prostate cancer

Author	AI method	Patients	Parameters	Predicted outcome accuracy
Tekeuchi et al. CUAJ 2018 [43]	Multilayer ANN	334 patients mpMRI transrectal biopsy	Twenty-two selected variables	Prostate cancer prediction ANN prevented 48% of patients without PCa from undergoing prostate biopsy. It missed 16% of PCa and 6% of GS ≥ 7. NPV was 76% for any PCa and 94% with GS ≥ 7
Zhang et al. Oncotarget 2016 [44]	Ml-based SVM classifier	205 PCa patients	Clinicopathological and MR imaging datasets	Image-based clinical outcome prediction: SVM vs. logistic regression and D`Amico risk groups SVM had significantly higher AUC (0.959 vs. 0.886, $p=0.007$), sensitivity (93.3% vs 83.3%, $p=0.025$), specificity (91.7% vs. 77.2%, $p=0.009$) and accuracy (92.2% vs. 79%, $p=0.006$) than logistic regression analysis. D'Amico scheme effectively improved by adding MRI-derived variables
Ishioka et al. BJUI 2018 [45]	CNN algorithm	335 patients with MRI and extended systematic prostate biopsy	MR images labelled as 'cancer' or 'no cancer'	Computer-aided diagnosis algorithm for detecting PCa on MRI Graphics processing unit required 5.5 h to learn to analyze 2 million images. Algorithms required 30 ms/image for evaluation and showed AUC of 0.645 and 0.636. 16/17 and 7/17 patients mistakenly diagnosed as having PCa in two validation datasets
Bonekamp et al. Eur Radiol 2018 [46]	ML-based radiomics models	316 men with MRI-transrectal US fusion biopsy	Lesion segmentation and radiomics analysis	Radiomic characterization of prostate lesions with MRI: ML vs ADC comparison AUC for the mean ADC ($AUC_{global}=0.84$; $AUC_{zone\text{-}specific}\leq 0.87$) vs the RML ($AUC_{global}=0.88$, $p=0.176$; $AUC_{zone\text{-}specific}\leq 0.89$, $p\geq 0.493$) showed no significantly different performance

Table 4 (continued)

Author	AI method	Patients	Parameters	Predicted outcome accuracy
Arvaniti et al. Sci Rep 2018 [47]	CNN algorithm	System trained on 641 patients and tested on 245 patients	Annotated dataset of PCa tissue microarrays	Automated Gleason score grading Agreements between the model and each pathologist were 0.75 and 0.71. Model´s GS ssignment achieved pathology expert-level stratification of patients into prognostically distinct groups on the basis of disease-specific survival
Donovan et al. Prostate Cancer Prostatic Dis 2018 [48]	ML/microscopic pattern analysis	892 post radical prostatectomy patients	Digital image analysis	Automated Gleason score grade and molecular profile In training, model predicted significant clinical failure with C-index of 0.82 and HR 6.7. Results confirmed in validation with C-index 0.77 and HR 5.4 for discriminating low from intermediate high- risk prostate cancer
Auffenberg et al. Eur Urol 2018 [49]	Random forest ML	7,543 PCa patients	3413 (45%) underwent RP, 2289 (30%) AS, 1280 (17%) RT, 422 (5.6%) ADT, and 139 (1.8%) WW	ML model to inform patients of PCa treatments chosen by similar men Personalized prediction for patients in the validation cohort was highly accurate (AUC = 0.81). Age, number of positive cores, and Gleason score were the most influential variables influencing patient treatment decisions
Abdollahi et al. Radiol Med 2019 [50]	ML-based radiomics models	33 PCa patients included	Radiomic features extracted from MRI T2W and ADC images	Prediction of intensity-modulated radiation therapy response Twenty highly correlated radiomics features with IMRT response. For GS prediction, T2W radiomics models were found to be more predictive (mean AUC 0.739), while for stage prediction, ADC models had a higher performance (AUC 0.675)
Hung et al. J Endourol 2018 [51]	3 ML algorithms	75 RARP cases	Automated performance metrics and hospital length of stay	RARP performance and outcome prediction Algorithm predicted length of stay with 87.2% accuracy. Patient outcomes predicted by the algorithm had significant association with 'ground truth' in surgery time, length of stay and Foley duration

Table 4 (continued)

Author	AI method	Patients	Parameters	Predicted outcome accuracy
Hung et al. BJUI 2019 [52]	DL model (DeepSurv)	100 RARP cases	Automated performance metrics, patient clinicopathological and continence data	Prediction of urinary continence recovery after RARP Continence was attained in 79% after median of 126 days. DL model achieved C-index of 0.6 in predicting continence. Automated performance metrics were ranked higher by the model than clinicopathological features
Wong et al. BJUI 2019 [53]	3 ML algorithms	338 RARP cases	19 different training variables	Early biochemical recurrence prediction after RARP All three ML models outperformed the conventional statistical regression model (AUC 0.865). Accuracy prediction scores for K-nearest neighbor, random forest tree, and logistic regression were 0.976, 0.953 and 0.976, respectively

PCa prostate cancer, *RP* radical prostatectomy, *AS* active surveillance, *RT* radiotherapy, *ADT* androgen deprivation therapy, *WW* watchful waiting, *ANN* artificial neural network, *CNN* convolutional neural network, *GS* Gleason score, *IMRT* intensity-modulated radiation therapy, *ADC* apparent diffusing coefficient, *RARP* robot-assisted radical prostatectomy, *ML* machine learning, *DL* deep learning, *SVM* support vector machine

ML and DL limitations

AI technologies have been attracting substantial attention in urology; however, their real-life implementation still faces obstacles. Several limitations exist in most studies applying ML and DL methods to urological diseases. First, the variability in study design, algorithms employed, training features used, and observed end points make it difficult to perform quantitative analysis. Second, most algorithms in these studies were validated with their dataset; therefore, they lack external validation and the generalizability of their results across other datasets is not applicable. Third, further algorithm development and research are particularly required in the field of urolithiasis to outperform conventional statistical methods as observed in urooncological investigations to reduce procedural costs and maximize patient outcomes. Lastly, some studies did not compare AI with conventional statistical analysis, since these methods only allow a limited number of training features, whereas AI can process big data and can thus be trained with a greater number of training features. For this reason, a comparison between any two techniques is challenging [55].

Future directions

Future research should focus on the construction of larger medical databases and further development of AI techniques. Once developed, the use of improved algorithms should not require large computer centers, but be performed on mobile devices or by access to cloud services. Specialized AI-based software for image-guided, real-time, intraoperative decisions will require appropriate regulatory approvals to function with robotic platforms and expand to operating rooms worldwide [56]. Issues remain regarding the trustworthiness of a computer´s diagnosis and that programming biases do not interfere with diagnoses. Human intuition, experience, and common sense will remain to play a crucial role in future AI developments to ensure that these systems are operating as intended and to deal with undesired consequences in a timely fashion.

Conclusion

The predictive precision of ML and DL will continue to provide and enhance personalized medicine with the further inclusion of data and model retraining. Larger patient datasets and electronic medical records can be semi-automated to provide instant predictive analytics that can be used to obtain insights into a variety of disease processes. Predictive accuracy, however, is highly dependent on efficient data integration obtained from different sources to enable it to be generalized. Although the shared decision-making will not be replaced by these models, it may complement the information patients obtain from traditional methods. While this is the beginning and further validation is required, there are limitless future applications for artificial intelligence in the field of urology.

Author contributions Project development: RS and AM. Literature review and data extraction: RS. Manuscript drafting: RS, GR, and AM. Manuscript editing: SH, CG, and AM.

Funding This research received no financial or other support.

Compliance with ethical standards

Conflict of interest The authors declare no conflicts of interest.

Human and animal rights statement This research did not involve human subjects or animals.

Ethical approval As this is a review of the literature, no ethical approval was necessary.

References

1. Nuffield Council on Bioethics (2018) Bioethics briefing notes: artificial intelligence (AI) in healthcare and research. https://nuffieldbioethics.org/wp-content/uploads/Artificial-Intelligence-AI-in-healthcare-and-research.pdf. Accessed 21 Dec 2018
2. Frankish K, Ramsey WM (eds) (2014) Introduction. The Cambridge handbook of artificial intelligence. Cambridge University Press, Cambridge, pp 1–14
3. Stuart R, Norvig P (eds) (2010) Artificial intelligence—a modern approach, 3rd edn. Prentice Hall, Upper Saddle River
4. Tran BX et al (2019) Global evolution of research in artificial intelligence in health and medicine: a bibliometric study. J Clin Med 8(3):360
5. Goldenberg SL, Nir G, Salcudean SE (2019) A new era: artificial intelligence and machine learning in prostate cancer. Nat Rev Urol 16(7):391–403
6. Yu KH, Beam AL, Kohane IS (2018) Artificial intelligence in healthcare. Nat Biomed Eng 2(10):719–731
7. Curran Associates Inc. (2014) Advances in neural information processing systems 26: 27th annual conference on neural information processing systems 2014, December 8–13. Curran Associates Inc., vol 1
8. Abbod MF et al (2007) Application of artificial intelligence to the management of urological cancer. J Urol 178(4 Pt 1):1150–1156
9. Kadlec AO et al (2014) Nonlinear logistic regression model for outcomes after endourologic procedures: a novel predictor. Urolithiasis 42(4):323–327
10. Aminsharifi A et al (2017) Artificial neural network system to predict the postoperative outcome of percutaneous nephrolithotomy. J Endourol 31(5):461–467
11. Choo MS et al (2018) A prediction model using machine learning algorithm for assessing stone-free status after single session shock wave lithotripsy to treat ureteral stones. J Urol 200(6):1371–1377
12. Mannil M et al (2018) Prediction of successful shock wave lithotripsy with CT: a phantom study using texture analysis. Abdom Radiol (NY) 43(6):1432–1438

13. Mannil M et al (2018) Three-dimensional texture analysis with machine learning provides incremental predictive information for successful shock wave lithotripsy in patients with kidney stones. J Urol 200(4):829–836
14. Seckiner I et al (2017) A neural network-based algorithm for predicting stone-free status after ESWL therapy. Int Braz J Urol 43(6):1110–1114
15. Langkvist M et al (2018) Computer aided detection of ureteral stones in thin slice computed tomography volumes using Convolutional Neural Networks. Comput Biol Med 97:153–160
16. Kazemi Y, Mirroshandel SA (2018) A novel method for predicting kidney stone type using ensemble learning. Artif Intell Med 84:117–126
17. Richard PO et al (2015) Renal tumor biopsy for small renal masses: a single-center 13-year experience. Eur Urol 68(6):1007–1013
18. Mir MC et al (2018) Role of active surveillance for localized small renal masses. Eur Urol Oncol 1(3):177–187
19. Bektas CT et al (2019) Clear cell renal cell carcinoma: machine learning-based quantitative computed tomography texture analysis for prediction of Fuhrman nuclear grade. Eur Radiol 29(3):1153–1163
20. Kocak B et al (2018) Textural differences between renal cell carcinoma subtypes: machine learning-based quantitative computed tomography texture analysis with independent external validation. Eur J Radiol 107:149–157
21. Kanapuli G et al (2018) A decision-support tool for renal mass classification. J Digit Imaging 31(6):929–939
22. Yu H et al (2017) Texture analysis as a radiomic marker for differentiating renal tumors. Abdom Radiol (NY) 42(10):2470–2478
23. Yan L et al (2015) Angiomyolipoma with minimal fat: differentiation from clear cell renal cell carcinoma and papillary renal cell carcinoma by texture analysis on CT images. Acad Radiol 22(9):1115–1121
24. Feng Z et al (2018) Machine learning-based quantitative texture analysis of CT images of small renal masses: differentiation of angiomyolipoma without visible fat from renal cell carcinoma. Eur Radiol 28(4):1625–1633
25. Cui EM et al (2019) Differentiation of renal angiomyolipoma without visible fat from renal cell carcinoma by machine learning based on whole-tumor computed tomography texture features. Acta Radiol 60(11):1543–1552
26. Coy H et al (2019) Deep learning and radiomics: the utility of Google TensorFlow Inception in classifying clear cell renal cell carcinoma and oncocytoma on multiphasic CT. Abdom Radiol 44(6):2009–2020
27. Minardi D et al (2005) Prognostic role of Fuhrman grade and vascular endothelial growth factor in pT1a clear cell carcinoma in partial nephrectomy specimens. J Urol 174(4 Pt 1):1208–1212
28. Holdbrook DA et al (2018) Automated renal cancer grading using nuclear pleomorphic patterns. JCO Clin Cancer Inform 2:1–12
29. Ding J et al (2018) CT-based radiomic model predicts high grade of clear cell renal cell carcinoma. Eur J Radiol 103:51–56
30. Kocak B et al (2019) Unenhanced CT texture analysis of clear cell renal cell carcinomas: a machine learning-based study for predicting histopathologic nuclear grade. AJR Am J Roentgenol 212:W1–W8
31. Lin F et al (2019) CT-based machine learning model to predict the Fuhrman nuclear grade of clear cell renal cell carcinoma. Abdom Radiol 44(7):2528–2534
32. Sun X et al (2019) Prediction of ISUP grading of clear cell renal cell carcinoma using support vector machine model based on CT images. Medicine (Baltimore) 98(14):e15022
33. Li P et al (2018) Fifteen-gene expression based model predicts the survival of clear cell renal cell carcinoma. Medicine (Baltimore) 97(33):e11839
34. Kocak B et al (2019) Radiogenomics in clear cell renal cell carcinoma: machine learning-based high-dimensional quantitative CT texture analysis in predicting PBRM1 mutation status. AJR Am J Roentgenol 212(3):W55–W63
35. Xu X et al (2017) Three-dimensional texture features from intensity and high-order derivative maps for the discrimination between bladder tumors and wall tissues via MRI. Int J CARS 12(4):645–656
36. Zhang X et al (2017) Radiomics assessment of bladder cancer grade using texture features from diffusion-weighted imaging. J Magn Reson Imaging 46(5):1281–1288
37. Eminaga O et al (2018) Diagnostic classification of cystoscopic images using deep convolutional neural networks. JCO Clin Cancer Inform 2:1–8
38. Sokolov I et al (2018) Noninvasive diagnostic imaging using machine-learning analysis of nanoresolution images of cell surfaces: detection of bladder cancer. Proc Natl Acad Sci USA 115(51):12920–12925
39. Brieu N et al (2019) Automated tumour budding quantification by machine learning augments TNM staging in muscle-invasive bladder cancer prognosis. Sci Rep 9(1):5174
40. Hasnain Z et al (2019) Machine learning models for predicting post-cystectomy recurrence and survival in bladder cancer patients. PLoS ONE 14(2):e0210976
41. Bartsch G Jr et al (2016) Use of artificial intelligence and machine learning algorithms with gene expression profiling to predict recurrent nonmuscle invasive urothelial carcinoma of the bladder. J Urol 195(2):493–498
42. Wu E et al (2019) Deep learning approach for assessment of bladder cancer treatment response. Tomography 5(1):201–208
43. Cha KH et al (2018) Diagnostic accuracy of CT for prediction of bladder cancer treatment response with and without computerized decision support. Acad Radiol 26:1137–1145
44. Takeuchi T et al (2019) Prediction of prostate cancer by deep learning with multilayer artificial neural network. Can Urol Assoc J 13(5):E145–E150
45. Zhang YD et al (2016) An imaging-based approach predicts clinical outcomes in prostate cancer through a novel support vector machine classification. Oncotarget 7(47):78140–78151
46. Ishioka J et al (2018) Computer-aided diagnosis of prostate cancer on magnetic resonance imaging using a convolutional neural network algorithm. BJU Int 122(3):411–417
47. Bonekamp D et al (2018) Radiomic machine learning for characterization of prostate lesions with MRI: comparison to ADC values. Radiology 289(1):128–137
48. Arvaniti E et al (2018) Automated Gleason grading of prostate cancer tissue microarrays via deep learning. Sci Rep 8(1):12054
49. Donovan MJ et al (2018) Development and validation of a novel automated Gleason grade and molecular profile that define a highly predictive prostate cancer progression algorithm-based test. Prostate Cancer Prostatic Dis 21(4):594–603
50. Auffenberg GB et al (2019) askMUSIC: leveraging a clinical registry to develop a new machine learning model to inform patients of prostate cancer treatments chosen by similar men. Eur Urol 75(6):901–907
51. Abdollahi H et al (2019) Machine learning-based radiomic models to predict intensity-modulated radiation therapy response, Gleason score and stage in prostate cancer. Radiol Med 124(6):555–567
52. Hung AJ et al (2018) Utilizing machine learning and automated performance metrics to evaluate robot-assisted radical prostatectomy performance and predict outcomes. J Endourol 32(5):438–444
53. Hung AJ et al (2019) A deep-learning model using automated performance metrics and clinical features to predict urinary continence recovery after robot-assisted radical prostatectomy. BJU Int 124(3):487–495

54. Wong NC et al (2019) Use of machine learning to predict early biochemical recurrence after robot-assisted prostatectomy. BJU Int 123(1):51–57
55. Chen J et al (2019) Current status of artificial intelligence applications in urology and their potential to influence clinical practice. BJU Int [Epub ahead of print]
56. Goldenberg SL, Nir G, Salcudean SE (2019) A new era: artificial intelligence and machine learning in prostate cancer. Nat Rev Urol 16(7):391–403