Artificial intelligence (AI) in urology—Current use and future directions: An iTRUE study

Milap Shah1,2, Nithesh Naik2,3, Bhaskar K. Somani2,4, BM Zeeshan Hameed1,2,5

ABSTRACT

Objective: Artificial intelligence (AI) is used in various urological conditions such as urolithiasis, pediatric urology, urogynecology, benign prostate hyperplasia (BPH), renal transplant, and uro-oncology. The various models of AI and its application in urology subspecialties are reviewed and discussed.

Material and methods: Search strategy was adapted to identify and review the literature pertaining to the application of AI in urology using the keywords “urology,” “artificial intelligence,” “machine learning,” “deep learning,” “artificial neural networks,” “computer vision,” and “natural language processing” were included and categorized. Review articles, editorial comments, and non-urologic studies were excluded.

Results: The article reviewed 47 articles that reported characteristics and implementation of AI in urological cancer. In all cases with benign conditions, artificial intelligence was used to predict outcomes of the surgical procedure. In urolithiasis, it was used to predict stone composition, whereas in pediatric urology and BPH, it was applied to predict the severity of condition. In cases with malignant conditions, it was applied to predict the treatment response, survival, prognosis, and recurrence on the basis of the genomic and biomarker studies. These results were also found to be statistically better than routine approaches. Application of radiomics in classification and nuclear grading of renal masses, cystoscopic diagnosis of bladder cancers, predicting Gleason score, and magnetic resonance imaging with computer-assisted diagnosis for prostate cancers are few applications of AI that have been studied extensively.

Conclusions: In the near future, we will see a shift in the clinical paradigm as AI applications will find their place in the guidelines and revolutionize the decision-making process.

Keywords: Artificial intelligence; deep learning; machine learning; prostate cancer; urolithiasis; urology.

Introduction

Artificial intelligence (AI) refers to the computational capability of the machine to mimic and perform human cognitive tasks. It is causing a paradigm shift in terms of providing health care and decision-making for the clinicians. The advances in the medical technologies used in health care, such as electronic medical records (EMRs), are providing humongous amounts of data.1 This large amount of data allows computer-based predictions and decisions to be made in aid in better patient care (Figure 1). By 2025, the growth rate of AI applications in health care is expected to be 29.3% and the global revenue is estimated to increase by 40%.2 With the available patient data, the future health care system is likely to move toward AI outpatient clinics and preventive medicine. AI provides more accuracy and reliable clinical decisions; hence, it is possibly going to be an integral part of the health care system.

The four subfields of AI in health care are as follows:

1. Machine learning (ML): ML is statistical technique-based programming that allows a computer system to learn and recognize patterns to model without explicit instructions. ML uses procedural computer programs wherein machines are trained to learn, detect data patterns, compute, and infer from the
datasets provided. It is observed that the machines are able to generate results similar to those generated by human intelligence.

2. Natural language processing (NLP): It illustrates the ability of a computer to comprehend the written and spoken language. Some applications that are possible through NLP include language translation, text processing, and speech recognition. In order to extract useful information and reliable details from patient services and provide “virtual assistance” for physicians, a comprehensive data research such as electronic medical record (EMR), doctor’s notes, pharmaceutical products, and medical imaging can also be analyzed.

3. Deep learning (DL) and artificial neural networks (ANNs): In the network architectural layers, the ANN comprises of individual units that function like artificial neurons programmed to accomplish computer tasks and recognize complex patterns. DL requires training massive datasets of multilayered neural networks. Deep neural convolution network (DCNN) is a commonly used ANN, which is effective when used in digitized image pattern identification or recognition.

4. Computer vision: Computer vision technology is used for visual search, trend forecasting, augmented reality, and virtual reality. The radiological and pathological images and simple and complex endoscopic videos can be used by machines to understand the details and patterns in the images in order to identify the tumors or malignancy present in the diagnostic images. The latest experience at human level in diagnostic imaging has already shown that AI has extensive “knowledge” to identify tumors. Computer vision can also be used for analysis and grading of pathological tissue slides.

AI is increasingly applied not only to the diagnosis of urological conditions but also its management and predictive analysis. This article focuses on addressing the application of AI and AI algorithms in urological subspecialties. This article reviews the use of AI in various benign and malignant conditions such as urolithiasis, pediatric urology, urogynecology, benign enlargement of the prostate, renal transplant, and uro-oncology pertaining to the kidneys, bladder, prostate, and testes.

Clinical and Research Consequences

Application of AI in Benign Urological Conditions

Urolithiasis
In the past few decades, there has been a rapid transition in the analysis, treatment and monitoring of cases with urolithiasis, the recent entry being applications of AI to identify the stone from computed tomography (CT) and ultrasound (US) images\cite{45}, detecting stone composition\cite{6,7}, predicting spontaneous stone passage\cite{8,9} and even the outcomes of endourological procedures (Table 1).\cite{10-13}
<table>
<thead>
<tr>
<th>Study</th>
<th>Objective</th>
<th>Study Design</th>
<th>Algorithm/Model</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Urolithiasis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parakh et al.[4]</td>
<td>Urinary stone detection on unenhanced CT images</td>
<td>• 535 patients (279 stones present; 256 stones absent) • 100 scans (test data)</td>
<td>Convolutional neural network (CNN)</td>
<td>&gt;90%</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Chiang et al.[5]</td>
<td>To identify association between stone disease and genetic polymorphisms, patient habits</td>
<td>• 151 (calcium oxalate stone) patients • 105 healthy controls</td>
<td>• Discriminant analysis • Artificial Neural network (ANN)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Kazemi et al.[6]</td>
<td>For early detection of the type of kidney stone and the most influential parameters</td>
<td>• 936 patients dataset • 42 features</td>
<td>• Bayesian model-Decision Trees • ANN • Rule-based classifiers</td>
<td>97.1% (ensemble model)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Kriegshauser et al.[7]</td>
<td>To investigate use of single-source dual-energy computed tomography (ssDECT) for the characterization of renal stones.</td>
<td>• 32 stone dataset</td>
<td>• Multiparametric algorithms</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Eken et al.[8]</td>
<td>To compare AI models with LR by applying on medical dataset</td>
<td>• 227 patients • 176 urinary stones • 51 no stones</td>
<td>• ANN • Genetic algorithm (GA) • Logistic regression analysis (LR)</td>
<td>NA</td>
<td>ANN: 94.9%</td>
<td>GA: 76%</td>
</tr>
<tr>
<td>Dal Moro et al.[9]</td>
<td>To predict the spontaneous passage of ureteral stones in patients with renal colic</td>
<td>• 1163 patients (402 found valuable) • nine clinical factors</td>
<td>Linear programming support vector machine (LPSVM)</td>
<td>NA</td>
<td>84.5%</td>
<td>86.9%</td>
</tr>
<tr>
<td>Shabaniyan et al.[10]</td>
<td>To predict postoperative outcome of PCNL</td>
<td>• 254 patients • 26 variables</td>
<td>Machine learning (ML) techniques such as sequential forward selection and Fisher’s discriminant analysis</td>
<td>94.8%</td>
<td>Requirement of Stent placement: 85.2%</td>
<td>Requirement of blood transfusion: 95%</td>
</tr>
<tr>
<td>Study</td>
<td>Objective</td>
<td>Study Design</td>
<td>Algorithm/Model</td>
<td>Accuracy</td>
<td>Sensitivity</td>
<td>Specificity</td>
</tr>
<tr>
<td>-------</td>
<td>-----------</td>
<td>--------------</td>
<td>----------------</td>
<td>----------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Aminsharifi et al.[11]</td>
<td>To predict multiple outcomes after percutaneous nephrolithotomy (PCNL) and compare with GSS and CROES nomogram</td>
<td>• 146 adult patients</td>
<td>• ML-based SVM model</td>
<td>80%-95.1%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Stone free status: 92% Need for repeat PCNL: 97% Need for ESWL: 82% Need for URS: 91%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kadlec et al.[12]</td>
<td>For outcome prediction after various forms of endourological intervention</td>
<td>• 382 renal units</td>
<td>• nonlinear LR model</td>
<td>Classification accuracy of 69.6%</td>
<td>Stone free status: 75.3% Need for sec. procedure: 30%</td>
<td>Stone free status: 60.4% Need for sec. procedure: 98.3%</td>
</tr>
<tr>
<td>Seckiner et al.[13]</td>
<td>To predict stone free status after ESWL</td>
<td>• 139 patients (training set) • 32 patients (validation set) • 32 patients (test set)</td>
<td>ANN</td>
<td>88.7% in the test group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Torshizi et al.[14]</td>
<td>To diagnose the severity of BPH and suggest appropriate treatment</td>
<td>• 44 patients</td>
<td>Fuzzy system expert</td>
<td>~90%</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Sonke et al.[15]</td>
<td>To diagnose BPH and compare with regression analysis</td>
<td>• 1903 patients</td>
<td>ANN</td>
<td>NA</td>
<td>71%</td>
<td>69%</td>
</tr>
<tr>
<td>Bagli et al.[16]</td>
<td>To predict sonographic outcome after pyeloplasty in children with ureteropelvic junction obstruction</td>
<td>• 84 children training set • 16 children test set</td>
<td>ANN</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Logvinenko et al.[17]</td>
<td>To predict patients at high risk of VUR abnormalities, based on RBUS findings</td>
<td>• 2259 patients</td>
<td>ANN Multivariate LR analysis</td>
<td>NA</td>
<td>For any grade VUR • ANN: 64% • MLR: 84%</td>
<td>For any grade VUR • ANN: 60% • MLR: 25%</td>
</tr>
<tr>
<td>Blum et al.[18]</td>
<td>To predict the need for surgery in UPJO cases based on dynamics of renogram</td>
<td>• 55 patients • 45 features</td>
<td>• Linear support vector machine (SVM)</td>
<td>93%</td>
<td>91%</td>
<td>96%</td>
</tr>
<tr>
<td>Sabadell et al.[20]</td>
<td>To predict occurrence of SUI after prolapse surgery and as a diagnostic tool</td>
<td>• 169 patients</td>
<td>ML algorithm</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>
Parakh et al.\textsuperscript{[4]} studied the diagnostic performance of the convolutional neural network (CNN) on CT images for detection of urinary stones in 535 adult patients assumed to have renal calculi using two scanners. The first scanner identified the urinary tract, and the next one detected the stone. Using nine different variation models, it achieved an accuracy of more than 90%. The study concluded that the efficiency of CNNs can be improved by the use of transfer learning with datasets augmented with labeled images.

Shabaniyan et al.\textsuperscript{[10]} developed a decision support system using ML techniques to predict the outcomes of surgical treatment for renal calculus. The algorithm was trained with a dataset of 254 patients and 26 parameters, which comprised variables from patients’ history, renal calculus composition, and laboratory investigations. This model achieved an accuracy of 94.8\%, 85.2\%, and 95\% in predicting outcomes of a procedure, predicting whether patient will require a stent after the procedure, and predicting the need for blood transfusion, respectively. Aminsharifi et al.\textsuperscript{[11]} studied data of 146 adult patients in whom percutaneous nephrolithotomy (PCNL) was done to validate efficiency of a machine-based learning algorithm for predicting the outcomes after PCNL and to compare the results with Clinical Research Office of Endourological Society nomogram and Guy’s Stone Score (GSS). This program predicted the PCNL results with an accuracy of up to 95\%.

Benign Enlargement of the Prostate

Many questionnaires are available for the clinical prediction of benign prostatic hyperplasia (BPH), yet the results are unreliable and inaccurate. Various AI techniques and ANN models such as multilayered back propagation method to predict the severity of obstruction on the basis of noninvasive tests have been used (Table 1).\textsuperscript{[14,15]} Torshizi et al.\textsuperscript{[14]} applied fuzzy intelligent systems in predicting the severity of BPH and also recommended the treatment required for it. The study consisted of two models. The first model predicted the severity, whereas the second model helped to make a treatment decision. The results were then compared for accuracy and validation with an expert panel. The accuracy achieved was nearly 90\%.

Pediatric Urology

AI has been used in the field of pediatric urology for predicting the outcome of surgical procedures\textsuperscript{[16]}, severity of the condition on the basis of imaging as well as detecting abnormalities in imaging (Table 1).\textsuperscript{[17,18]} Bagri et al.\textsuperscript{[16]} applied computerized

\begin{table}[h]
\centering
\begin{tabular}{|l|l|l|l|l|l|l|}
\hline
\textbf{Study} & \textbf{Objective} & \textbf{Study Design} & \textbf{Algorithm/Model} & \textbf{Accuracy} & \textbf{Sensitivity} & \textbf{Specificity} \\
\hline
Jelovsek et al.\textsuperscript{[21]} & To predict recurrence, complications, and health status improvement after prolapse surgery & • 1301 patients & LR models & NA & NA & NA \\
\hline
E. Renal Transplant
\hline
Atallah et al.\textsuperscript{[22]} & To predict 5-year graft survival & • 2728 patients (70\% training and 30\% testing) & • Naïve Bayes Based Feature Selector (NBBFS) Algorithm & 80.77\% & 81.2\% & NA \\
\hline
Greco et al.\textsuperscript{[23]} & To predict graft failure and association with BMI and other risk factors & • 194 patients & • ML algorithms & NA & 88.2\% & 73.8\% \\
\hline
Goldfarb et al.\textsuperscript{[24]} & To predict cadaveric graft survival over three years based on pretransplant variables & 37,407 patients dataset & Logistic regression based model & 65\% & NA & NA \\
\hline
\end{tabular}
\caption{Studies looking at applications of AI in benign urological conditions (Continue)}
\end{table}

ANN to predict the outcomes after pyeloplasty on the basis of US findings in children with uretero-pelvic junction obstruction. The prediction was based on whether postoperatively the results were “significantly improved,” “improved,” “same,” or “worse.” The results showed 100% sensitivity and specificity for all four-outcome measures. Both multivariate analysis and ML algorithms were used by Logvinenko et al. [17] to evaluate whether the renal and bladder ultrasound (RBUS) could predict the abnormalities on voiding cystourethrogram (VCUG) for conditions such as vesico-ureteric reflux and congenital urethral abnormalities. The results showed that RBUS was a poor predictor of the abnormalities of VCUG and both could only complement each other but cannot replace them.

Urogynecology
On the basis of urinary incontinence data obtained from the wearable devices, AI techniques were applied to predict the time and number of incontinence episodes and the outcome of conservative or medical management for stress urinary incontinence (SUI). [18] Models to predict the occurrence of complications such as SUI after prolapse surgery, recurrence, and overall outcomes of surgery were also studied on the basis of the data available from various randomized controlled trials (Table 1). [20,21] In the near future, AI applications can be used to provide personalized care based on the patient demographics and clinical characteristics of every individual.

Renal Transplant
The outcome of kidney transplant prediction is very important. Various studies have been conducted to predict the outcomes of kidney transplantation using ANN and ML algorithms, as described (Table 1). [22-24] Atallah et al. [22] proposed a prediction method by combining two methods-Bayes and k-nearest neighbor—which achieved more accuracy by choosing minimum number of features. It was based on data mining techniques to predict five-year graft survival after transplantation. This new proposed prediction method comprises three stages: data preparation stage, feature selection stage, and prediction stage. This prediction method can be used in other transplant datasets to measure the graft survival.

Application of AI in Uro-Oncology
Testicular Malignancy
Not much has been studied about the applications of AI in testicular malignancy. Baessler et al. [25] applied ML-based CT radiomics to determine whether the lymph nodes dissected in patients with metastatic or advanced nonseminomatous testicular germ cell tumor were malignant or benign. The model correctly classified with an accuracy of 0.81 (area under the curve [AUC]), 88% sensitivity, and 72% specificity. 

Renal Cell Carcinoma
ML and DL algorithms based on CT-texture analysis were applied for differentiating renal masses such as angiomylipoma, clear cell renal cell carcinoma (ccRCC), papillary renal cell carcinoma, and oncocytoma [26-28] to predict the nuclear grade and to identify certain genetic mutations to predict the prognosis, recurrence, and survival outcomes (Table 2). Kocak et al. [26] used CT-texture analysis, applied ML techniques to predict and identify the nuclear grade (Furhman) of ccRCC, and compared the results with those obtained with percutaneous biopsy. The results were comparable, and the maximum predictive value was achieved with the use of the support vector machine (SVM). The algorithm could differentiate nuclear grades in 85.1% of ccRCC cases. Ding et al. [29] also conducted a similar study showing increased precision in classifying the grade of ccRCC.

Biomarkers and signatures based on more than one gene expression have been developed in recent years for predicting the ccRCC overall survival (OS) and prognosis of the disease. Li et al. [30] developed a model based on 15 genes, which could help predict the prognosis and survival. They found that the group with a higher risk had substantially poorer prognosis and survival than the group with patients having lower risk. The risk groups were not associated with patient characteristics such as sex or age but were related to hemoglobin levels. They were also associated with tumor features such as size and grade.

PBRM1 mutations are the second most common mutations found in ccRCC. Kocak et al. [31] applied ANN- and ML-based algorithms to identify PBRM1 mutations based on CT scan texture analysis. Overall, 88% of ccRCC with PBRM1 mutation status was correctly identified by ANN. On the basis of these results, future studies can be conducted to develop noninvasive biomarkers for identifying histopathological subtypes to predict the prognosis and response to treatment.

Bladder Cancer
ML algorithms, DCNN models, genetic algorithms, and SVMs have been applied in bladder cancer for improving cystoscopic diagnosis and prediction of prognosis and survival (Table 2) [32-35] Ikeda et al. [32] made a competent CNN by training it with 2102 cystoscopic pictures with an aim to increase the efficiency in diagnosis of bladder cancer using AI. It achieved sensitivity and specificity of 89.7% and 94.0%, respectively. Lorencin et al. [33] used the data of 1997 and 986 images with and without bladder cancer, respectively, to train multilayer perceptron along with DCNN for the diagnosis of bladder malignancy. It showed promising results, with AUC value reaching up to 0.99. Wang et al. [36] achieved more than 75% accuracy by using least squares SVM in predicting the five-year overall and cancer-specific mortality of patients post radical cystectomy.
<table>
<thead>
<tr>
<th>Study</th>
<th>Objective</th>
<th>Study Design</th>
<th>Algorithm/Model</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Renal Cell Carcinoma (RCC)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Kocak et al.[26] | To distinguish the major subtypes of RCC                                    | • 68 RCC patients for internal validation  
• 26 RCC patients for external validation  
• 275 CT images for texture features | Artificial Neural Network (ANN) Support vector machine (SVM)  
• ANN: 84.60%  
• SVM: 69%  
• ANN: 69%  
• SVM: 71%  
• ANN: 100%  
• SVM: 100% |                   |             |             |
| Feng et al.[27]  | To differentiate angiomyolipoma (AML) and RCC based on texture analysis of CT images | • 58 patients  
• 42 features  
• Machine Learning (ML) based quantitative texture analysis  
• SVM with recursive feature elimination  
• Synthetic minority oversampling technique (SMOTE) | 93.9%  
87.8%  
100% |                   |             |             |
| Coy et al.[28]  | To distinguish ccRCC and oncocytoma from MDCT images                        | • 4000 iterations (90% training and 10% validation)  
• 179 patients  
• 161 labeled segmentations (87 without mutation; 74 with mutation) | Deep Learning (DL) based Google TensorFlow software  
74.4%  
85.8%  
NA |                   |             |             |
| Ding et al.[29] | To preoperatively distinguish high nuclear grade from low nuclear grade in ccRCC | • 92 cases (for validation)  
• Logistic Regression (LR) model  
• Least absolute shrinkage and selection operator (LASSO) for texture score | NA  
NA  
NA |                   |             |             |
| Li et al.[30]   | To predict survival in patients with ccRCC based on gene expression         | N=533 (training dataset) Risk score model based on 15 genes  
N=101 (test dataset)  
• ML-based random forest variable hunting Cox regression analysis | NA  
NA  
NA |                   |             |             |
| Kocak et al.[31] | To identify the mutation status of PBRM1 gene in ccRCC patients             | • 45 patients (29 without mutation; 16 with mutation)  
• 161 labeled segmentations (87 without mutation; 74 with mutation)  
• ML-based quantitative CT-texture analysis such as  
• Random Forest (RF) algorithm  
• ANN | ANN: 88.22%  
RF: 95%  
NA  
NA |                   |             |             |
<table>
<thead>
<tr>
<th>Study</th>
<th>Objective</th>
<th>Study Design</th>
<th>Algorithm/Model</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. Bladder Carcinoma</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ikeda et al. [32]</td>
<td>To improve cystoscopic diagnosis of bladder cancer using AI</td>
<td>2102 images (1671 normal tissue; 431 tumor lesions)</td>
<td>Convolutional Neural Network (CNN)</td>
<td>NA</td>
<td>89.7%</td>
<td>94%</td>
</tr>
<tr>
<td>Lorencin et al. [33]</td>
<td>To use multilayer perceptron method for diagnosis of bladder cancer from cystoscopic images</td>
<td>1997 bladder cancer images, 986 non-cancer tissues images</td>
<td>Multilayer Perceptron (MLP)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Hashemi et al. [34]</td>
<td>To classify cystoscopic bladder images using AI</td>
<td>540 cystoscopic bladder images</td>
<td>Multilayer neural networks, Genetic algorithm (GA)</td>
<td>7% decrease in error on classification as compared with other methods</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Eminaga et al. [35]</td>
<td>To perform diagnostic classification based on cystoscopic images using DL-CNN</td>
<td>479 patients, 18,681 images (generated with 10 degree grades), 60% training set-10% validation set, 30% test set</td>
<td>Deep Learning CNN (DL-CNN), Xception model, ResNet50 model, InceptionV3, VGG-19, VGG-16</td>
<td>F1 scores Xception: 99.52% ResNet: 99.48%</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Gavriel et al. [37]</td>
<td>To predict five-year prognosis of bladder cancer</td>
<td>78 patients diagnosed with MIBC</td>
<td>ML-based ensemble model</td>
<td>94.8%</td>
<td>89.5%</td>
<td>97.4%</td>
</tr>
<tr>
<td>Study</td>
<td>Objective</td>
<td>Study Design</td>
<td>Algorithm/Model</td>
<td>Accuracy</td>
<td>Sensitivity</td>
<td>Specificity</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------</td>
<td>----------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Hasnain et al. [38]</td>
<td>To predict postcystectomy recurrence and survival</td>
<td>• Dataset of 3503 patients</td>
<td>Ensemble ML-based models • Support vector machine (SVM) • K-nearest neighbor Algorithm (KNN) • Random Forest • Gradient-boosted trees (GBT)</td>
<td>NA</td>
<td>&gt;70%</td>
<td>&gt;70%</td>
</tr>
<tr>
<td>Bartsch et al. [39]</td>
<td>To predict recurrence of NMIBC based on genome profile</td>
<td>• 112 frozen NMIBC specimens • 21 gene classifier set</td>
<td>ML cased genetic programming algorithm</td>
<td>NA</td>
<td>Test Set Five gene combined rule: 69% Three gene combined rule: 71% Test Set Five gene combined rule: 62% Three gene combined rule: 67%</td>
<td></td>
</tr>
<tr>
<td>Wu et al. [40]</td>
<td>To compare different DL-CNN models to predict response to treatment in bladder cancer (T0 prediction)</td>
<td>• 123 CT scans (pre and post-treatment)</td>
<td>Multiple DL-CNN models with structure modification and layer freezing</td>
<td>Base DL-CNN 70%</td>
<td>Base DL-CNN 60%</td>
<td>Base DL-CNN 80%</td>
</tr>
<tr>
<td>C. Prostate Carcinoma</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ström et al. [41]</td>
<td>To diagnose and grade prostate cancer in biopsies</td>
<td>Training set-976 patients (6682 slides) Test set-246 patients (1631 slides)</td>
<td>ANN</td>
<td>NA</td>
<td>99%</td>
<td>94.9%</td>
</tr>
<tr>
<td>Bulten et al. [42]</td>
<td>To assign Gleason grade to prostate biopsies using AI</td>
<td>1243 patients (5759 biopsies)</td>
<td>DL system</td>
<td>Benign versus malignant: 96%-97% Grade group 2 or more: 79%-83% Grade group 3 or more: 76%-82%</td>
<td>Benign versus malignant: 83%-100% Grade group 2 or more: 52%-70% Grade group 3 or more: 72%-782%</td>
<td></td>
</tr>
<tr>
<td>Viswanath et al. [44]</td>
<td>To compare various classifier in detecting CaP on t2W MRI images using radiomic texture features</td>
<td>• 85 T2W MRI datasets</td>
<td>• Quadratic Discriminant Analysis (QDA) • Support Vector Machines (SVMs) • Naïve Bayes Decision Trees (NBDT)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>
Gavriel et al. proposed an ensemble system comprising ML-based algorithms to predict five-year prognosis with different combinations of image, clinical, and spatial features and quantify potential prognostic markers related to lymphocytes, macrophages, tumor buds, and PD-L1. The method successfully classified 71.4% of the patients who succumbed to muscle invasive bladder cancer (MIBC) within five years, significantly higher than the 28.6% of the current clinical gold standard, the tumour, node, metastasis (TNM) staging system.

Several studies have applied ML-based algorithms and models to identify genes that could predict the recurrence of disease or the future progression. Slides of patients diagnosed with MIBC were labeled with immunofluorescence (IF) and used for measuring the tumor buds, to determine the effectiveness of neoadjuvant chemotherapy, and to identify patients who were not responding to the treatment. This was done to stop the treatment prematurely to avoid the adverse effects of chemotherapy.
Prostate Carcinoma

AI applications are on the verge of revolutionizing the current practice in carcinoma prostate (CaP) in terms of diagnosis, treatment decisions, and even predicting the disease-free survival. There is high observer-dependent variability in Gleason grading because of the subjective nature of the analysis of biopsy specimens. Considering this, Ström et al.\cite{strom2004} developed an AI model for identification, Gleason grading, and localization of prostate cancer. The model was trained with 6682 digitized slides of 976 men and tested on 1631 biopsy specimens from 246 men. It achieved an accuracy of 0.997 (AUC) to differentiate between a malignant and a benign tumor. The results in terms of Gleason grading were also comparable to those achieved by the expert pathologists. In various studies, DL methods to calculate the Gleason Grading have been applied (Table 2).\cite{deng2020}

Multiparametric imaging uses multiple modalities or techniques before making the ultimate diagnosis, and this adds to the burden of the radiologist. However, in the current era, computer-aided diagnosis is possible because of progress in AI, which eventually helps in making the diagnosis by image interpretation. This is particularly useful in situations where multiple modalities, parameters, or techniques are involved in diagnosing a condition.\cite{wang2020}

Application of a Quadrant Discriminant Classifier to the radiomic features derived from T2-weighted MRI images for detection of CaP\cite{bibault2020} and application of ML-based random Forrest classification algorithm to localize CaP on transrectal ultrasonogram\cite{shah2020} have been studied (Table 2).

In view of the toxic effects of docetaxel chemotherapy, 20% of the patients undergo therapeutic failure in metastatic castrate resistant CaP. Deng et al.\cite{deng2020} developed an AI-based computational model that could differentiate patients in two groups, docetaxel-tolerable and docetaxel-intolerable, for better and individualized treatment for the patients in this category. Identification of the presence of biomarkers on tissue microarrays can predict the risk of recurrence and metastasis. Biomarker identification under IF microscope by the human eye is subjective as well as time-consuming. Hence, an automated method using DL algorithms was developed for analysis of biomarkers using 648 samples and IF staining with anti-Ki-67, ERG antibodies. The results were promising, with only 5% difference between manual and algorithm-based biomarker detection and 100% accuracy in identification of tumors positive for ERG.\cite{bibault2020}

Bibault et al.\cite{bibault2020} used data from the prospective clinical trial Prostate Lung Colorectal and Ovarian cancer screening, selected patients who were diagnosed with CaP during follow-up, and trained two models to predict ten-year cancer-specific survival (CSS) and OS. Of the 8776 patients diagnosed with PCa on follow-up, training of the models was done with 7021 and tested on dataset of 1755. It achieved an accuracy of 0.87 and 0.98 for OS and CSS, respectively. These models can be used online to provide predictions and support informed decision-making in CaP treatment.

Limitations

AI applications are gaining significant interest in urology, but their real-world implementation still faces an uphill task. There are limitations to some studies that use AI algorithms and its subsets in urological diseases. The key challenges that can be addressed before being integrated into the clinical setting are the incorporation of standardized criteria, the correction for system variation, and the data collection from multiple institutions in various geographical locations, so that the results can be generalized and applied to the real-world scenario.\cite{shah2020}

Future Considerations

The President of the World Economic Forum, Klaus Schwabe, made the following announcements at the Davos Summit just a few years ago: “We stand on the brink of a technological revolution that will fundamentally alter the way we live, work, and relate to one another. In its scale, scope, and complexity, the transformation will be unlike anything human kind has experienced before”.\cite{schwab2020}

Future work will concentrate on creating larger medical databases and expanding AI techniques further. The use of enhanced algorithms will take place on smartphones or can be accessed through the cloud. Applications for clinical decision-making and its use in the real world require appropriate permissions from the regulatory bodies. Issues exist concerning the reliability of a machine diagnosis and that prejudices of programming do not create hindrances in the diagnosis.

Conclusion

In the near future, we will see a shift in the clinical paradigm as AI applications will find their place in the guidelines and revolutionize the decision-making process. Having said that, human qualities of intelligence, adaptation, and sense of duty will prove to be important factors in further development of AI.

Peer-review: This manuscript was prepared by the invitation of the Editorial Board and its scientific evaluation was carried out by the Editorial Board.

Conflict of Interest: The authors have no conflicts of interest to declare.

Financial Disclosure: The authors declared that this study has received no financial support.

References

1. Beam A, Kohane I. Big Data and Machine Learning in Health Care. JAMA 2018;319:1317. [Crossref]
17. Logvinenko T, Chow J, Nelson C. Predictive value of specific ultrasound findings when used as a screening test for abnormalities on VCUG. J Pediatr Urol 2015;11:176.e1-176.e7. [Crossref]
renal masses: Differentiation of angiomyolipoma without visible fat from renal cell carcinoma. Eur Radiol 2017;28:1625-33. [Crossref]


30. Li P, Ren H, Zhang Y, Zhou Z. Fifteen-gene expression based model predicts the survival of clear cell renal cell carcinoma. Medicine 2018;97:e11839. [Crossref]

31. Kocak B, Durmaz E, Ates E, Baykara Ulusan M. Radiogenomics in Clear Cell Renal Cell Carcinoma: Machine Learning-Based High-Dimensional Quantitative CT Texture Analysis in Predicting PBRM1 Mutation Status. AJR Am J Roentgenol 2019;212:W55-W63. [Crossref]


43. Wildeboer R, van Sloun R, Wijkstra H, Mischi M. Artificial intelligence in multiparametric prostate cancer imaging with focus on deep-learning methods. Comput Methods Programs Biomed 2020;189:105316. [Crossref]


46. Deng K, Li H, Guan Y. Treatment Stratification of Patients with Metastatic Castration-Resistant Prostate Cancer by Machine Learning. iScience 2020;23:100804. [Crossref]


