Artificial Intelligence in Urological Oncology

Andrew Brodie, Hertfordshire and Bedfordshire Urological Cancer Centre, Department of Urology, Lister Hospital, Stevenage, UK

Dai Dairu, Luton and Dunstable University Hospital, UK

Jeremy Yuen-Chun Teoh, S.H. Ho Urology Centre, Department of Surgery, The Chinese University of Hong Kong, Hong Kong, China

Karel Decaestecker, Department of Urology, Ghent University Hospital, Belgium

Prokar Dasgupta, Faculty of Life Sciences and Medicine, King's College London, London, UK

Nikhil Vasdev, Hertfordshire and Bedfordshire Urological Cancer Centre, Department of Urology, Lister Hospital, Stevenage, UK / School of Medicine and Life Sciences, University of Hertfordshire, Hatfield, UK

Email - karel.decaestecker@gmail.com

prokar.dasgupta@kcl.ac.uk

nikhil.vasdev@nhs.net

Correspondence – nikhil.vasdev@nhs.net

Commented [NV1]: Abstract and Key words

INTRODUCTION

The nascence of the term ‘Artificial Intelligence’ (AI) originates from an ambitious research project led by a 20-strong team of Dartmouth College mathematicians in the 1950s (1). 60-years later and the early aspirations of this group of optimistic pioneers has come to fruition with the development of thinking machines capable of human-like intelligence. Multiple definitions exist for the innovative computer science encompassed under the umbrella term of artificial intelligence (AI) but broadly speaking it describes the capability of a machine to mimic human cognition through perception of external stimulus and determination of an optimal strategy to achieve a desired outcome. Whilst this definition may sound far-fetched, AI is already well-established in today’s society with commonplace applications such as smartphone speech recognition, targeted advertising, spam-filtering, search engines and fraud detection in the banking sector(1).

The speed of progress to date has been staggering and such is the undeniable potential of AI that it is hardly surprising that the UK government has earmarked AI as a ‘Grand Challenge’ for its future industrial strategy(2). With the government pledge to

put the UK at the ‘forefront of AI innovation’ medicine and surgery can take centre stage as a model of how sectors may benefit from adopting this new technology. Furthermore, The Royal College of Surgeon’s independent report on the Future of Surgery provides a glimpse into the future and highlights predictive analytics, radiology, pathology, genomics and robotic surgery as key areas where AI will benefit patients most (3).

Urology as a forward-thinking surgical speciality with an exemplary track-record in early adoption of new technologies is uniquely placed to take full advantage of the benefits of embracing AI. This review provides an outline on the current and future applications of AI in the context of urological oncology. Accomplishing this task requires a broad appreciation of the key concepts and esoteric definitions underpinning AI and thus the first section of the review will concentrate on the fundamentals of AI.

1. BASICS OF ARTIFICIAL INTELLIGENCE

1A). Machine learning

Machine learning (ML) combines computer science, mathematics and statistical analysis to generate algorithms assessing variables (referred to as ‘features’) to predict outcomes (referred to as ‘labels’). There are similarities with conventional statistical methods. However, ML focuses on outcome prediction over inference of relationships between variables, and whilst classical statistics tests hypotheses through probability (p-value), significance levels are rarely cited in ML research. A key feature of ML algorithms is the ability over time to autonomously adapt their own programming to improve performance towards pre-determined outcomes. This ‘learning’ process is dependent upon the quality and volume of data introduced to the algorithm. The ML

methodology (figure 1) necessitates processing of raw data into ‘training’, ‘validation’ and ‘test’ data sets. Training data annotated by experts; for example, lesions on histopathological slides may be pre-designated as benign or malignant by pathologists in order for ML to create an algorithm capable of formulating this classification. The validation process provides feedback to optimise algorithm performance and the testing phase, ideally with external data, assesses final performance of the complete algorithm. Classically the explicit features or variables that a ML algorithm assesses are pre-determined based on expert knowledge, an example would include the number of glands per unit measured in prostate histopathological slides.

1B). Deep learning

Deep learning (DL) is a contemporary subsection of ML receiving increased attention due to its impressive performance made possible through recent improvements in computational power and expanding data sets. DL techniques include artificial neural networks (ANN), interesting models simulating the organisational structure of the human brain with separate computational units (artificial neurons) connected to each other through ‘synapses’. Multiple layers of increasingly complex interconnected units add ‘depth’ to the system. For example, images of animals may be inputted into an ANN, the first layer recognises image edges, the second layer colours and so on with increasing complexity until the system gives an output in the form of the classifying an animal species. Multiple DL models exist, one often described class is a convolutional neural network (CNN). By mirroring the architecture of the human visual cortex CNNs are proficient in image analysis and been applied to facial recognition in social media platforms. In contrast to traditional ML technique’s prerequisite for explicit feature identification by experts prior to training, a key component of DL techniques is the ability to autonomously identify pertinent features from raw data during the training

phas. This is advantageous because DL techniques can identify novel features within medical images imperceptible to the human eye and such previously unidentified features may be strongly predictive of clinical outcomes, vastly improving the analytical power and clinical utility in medical diagnostics. Figure 2 provides a pictorial comparison of the described methods.

1C). Big data

Big data refers to increasingly complex data sets so large they exceed the processing and analytical capacity of conventional software. To put big data in context the International Data Corporation estimates that the volume of stored data globally will rise from 130 exabytes (EB) in 2005 to 40,000EB by 2020(4)(1 EB = 1 billion gigabytes - GB). A similar trend is noted in medicine with a predicted 48% annual increase in healthcare related data storage (5). Increasing volumes of unstructured data are generated from a variety of sources including electronic health records, clinical data e.g. vital signs , radiological imaging, 'omics' data, patient 'wearables' and technical data from surgical robotic systems(6). AI is the perfect tool to analyse large volumes of heterogenous data and proposed benefits including reduced healthcare costs, detection of novel disease patterns, earlier disease prediction, reduced healthcare fraud, increased clinician efficiency and novel drug design and delivery. Big Data projects are currently underway in the UK; for example the world-first '100,000 Genomes Project' launched in 2012 has sequenced the DNA of 100,000 NHS patients with rare diseases and cancers linking genomics with electronic health records(7) and enhancing our understanding of these conditions. The role of AI in harnessing the

benefits of Big data will be vital in realising the goal of early disease detection and personalised medicine for all patients including those with urological malignancies.

2. APPLICATIONS OF AI IN UROLOGICAL ONCOLOGY

2A). AI IN PROSTATE CANCER

Prostate cancer is an ideal candidate to benefit from AI implementation due to the already established data-rich diagnostic approaches including multi-parametric magnetic resonance imaging (mpMRI), extensive mapping prostate biopsies, analytics from robotic surgical systems and mature genomic sequencing. The next section of the review focuses on differing aspects of AI in prostate cancer diagnostics, management and prediction. Table 1 provides a summary of key AI prostate cancer studies.

2A1). AI in prostate cancer imaging

The benefits of combining AI techniques with radiological imaging in cancer patients has received significant media attention recently due to publication of a landmark study in breast mammography claiming AI 'outperformed' radiologists in identifying malignant lesions (8). Similar AI techniques are being applied to prostate imaging such as mpMRI with remarkable results. mpMRI of the prostate has become a central component in the diagnostic pathway of prostate cancer and its use is now widespread in the western world. The process of mpMRI acquisition generates large volumes of quantitative data reconstructed into 2D images for subjective interpretation by radiologists. As a consequence, there is a surplus of latent quantitative data stored within MRI datasets imperceptible to the human eye. The study of this data extracted from medical images has spawned an entirely new field of radiology entitled

'radiomics'(9). ML algorithms can rapidly process millions of radiomic features, wheras radiological reporting system such as the Prostate Imaging-Reporting and Data System (PI-RADS) are constrained by human cognition in the number of assessable features. Such limitations leads to the poor reproducibility and wide inter-user variability in sensitivity, specificity and prostate cancer detection rates demonstrated in current reporting prostate mpMRI(10). These criticisms make mpMRI ideal for analysis and interpretation by AI algorithms.

2A2). AI prostate cancer detection

Early application of ML focused on low-level registration and segmentation of prostate anatomy from radiological imaging (1). Recently, more sophisticated DL algorithms focus on clinically useful identification of index lesions or prediction of Gleason grade/tumour aggressiveness from regions of interest; Hambrock et al.(11) developed an in-house computer assisted diagnostic (CAD) tool for prostate mpMRI capable of overlaying likelihood of malignancy graphic representations on to MRI to assist radiologists in reporting. The study noted significant improvements in the identification of prostate cancer on mpMRI by less experienced radiologists, with pre-CAD area under the curve (AUC)=0.81 improving to AUC=0.91 following implementation of the CAD tool. The performance of CAD-augmented novice radiologists was similar to that of more experienced radiologists; AUC=0.93. Karimi et al.(12) combined a CNN with other forms of machine learning to classify MRI lesions into benign or malignant. The training dataset included 232 lesions and the testing dataset 98 lesions. By combining ML techniques, the AUC achieved was 0.87, higher than when either ML technique was used in isolation. Wang et al(13) adopted a similar combined approach but merged ML analysis of MRI radiomic data with radiologist review using PI-RADS version 2. The combined ML approach performed better than PI-RADS alone with a significant improvement in detection of prostate cancer from AUC=0.88 to 0.98 in the

peripheral zone and AUC=0.94 to 0.97 in the transitional zone. A multi-institute study by Gaur et al.(14) noted not only an improvement in specificity when combining CAD with radiologist reporting of mpMRI but also an increase in efficiency with mean time of mpMRI reporting reducing from 4.6 minutes to 3.4 minutes ($p<0.001$).

2A3). AI prostate cancer grading

The biological behaviour of prostate cancer is heterogeneous and a significant proportion of prostate cancer patients die with the disease rather than because of it. Thus, the ability to accurately identify clinically significant prostate cancer from insignificant disease is vitally important for risk-stratification and decision-making. Increasingly advanced ML algorithms have the capability to assess aggressiveness of prostate cancer lesions based on mpMRI radiomic features. A 2019 study by Zhong et al.(15) employed a DL algorithm to predict Gleason grade ≥ 7 from the mpMRIs of 140 patients undergoing robotic prostatectomy. The study noted an AUC=0.73 comparable to radiologists using PI-RADS (AUC=0.71). Varghese et al.(16) compared 7 ML techniques' ability to predict the National Comprehensive Cancer Network (NCCN) prostate cancer risk-group of 68 patients undergoing mpMRI followed by transrectal ultrasound-MRI fusion biopsy. The optimal ML technique was a support vector machine (SVM), which assessed 110 radiomic features and reached comparable AUC (0.71) to PI-RADS (0.73). Similarly, Cao et al.(17) reported the results of 'FocalNet' a CNN used for identifying clinically significant PCa from the mpMRI of an impressive cohort of 417 prostatectomy patients. Comparison was made with expert radiologists with over 10-years' experience and reading of over 1000 prostate mpMRIs. There was no statistical difference in the sensitivity of FocalNet in identifying clinically significant lesions (Gleason score ≥ 7) compared to expert radiologists, 79.2% and 80.7% respectively.

The majority of ML studies in prostate mpMRI aim to identify lesions or Gleason grade. However, other applications have been described including local staging, prediction of extraprostatic extension at robotic prostatectomy, radiotherapy planning and prediction of biochemical recurrence after radical treatment(18). Furthermore, whilst ML research primarily focuses on mpMRI, other imaging modalities such as transrectal ultrasound (TRUS) have been investigated, notably in MRI/US fusion technology for biopsy(19).

2A4). PROSTATEx Challenge

One obstacle to widespread adoption of AI in prostate cancer imaging is the lack of high-quality annotated training data. The PROSTATEx challenge is worthy of mention because it represents an interesting solution to this issue by providing a publicly available training database of mpMRI images. AI research groups are pitted against each other to enable head to head comparison of ML techniques using validated performance metrics in the form of 'challenges'. The PROSTATEx training data set contains over 300,000 images from mpMRIs with 412 lesions for ML lesion identification in the first challenge and prediction of Gleason grade group in the second challenge. Armato et al.(20) provide a thorough analysis of the challenges methodology and results demonstrating the importance of high-quality large training data sets in enabling translation of AI prostate imaging methods from the laboratory to clinical practice.

2A5). AI in prostate histopathology

Gleason grading has been a vital tool in prognosticating risk in PCa patients since its inception in the 1960's. As our understanding of the molecular and genetic basis of prostate cancer has advanced the Gleason grading system has been refined. The

current iteration based on the 2014 International Society for Urological Pathology (ISUP) grading system recommends classification into 5 grade groups(21). Similar to criticisms of mpMRI interpretation there is comparable inter-observer variability in Gleason grading(22). This issue is compounded by a nationwide pathologist workforce crisis and increased amounts prostate pathology specimens due to adoption of template prostate biopsies and a move away from 12-core transrectal ultrasound guided biopsy (TRUS). Technological advancements may hold the key to reproducible standardised grading. For example, increasing digitisation of pathology slides and recent FDA approval for whole-slide imaging (WSI) technology has transferred pathologist review from the microscope to the computer monitor enabling telepathology, increased collaboration, expedient second-reading and facilitation of automatic grading with AI techniques(23). Specifically, the role of AI in prostate cancer histopathology will be discussed further.

Nagpal et al.(24) describe the largest study incorporating a DL CNN algorithm to identify Gleason grade from WSIs of prostatectomy specimens and provide quantitation of grade groups. The algorithm was trained on an impressive 112 million images from 912 annotated slides and validated on a reference set of 331 slides. The concordance of the DL algorithm with the reference set, provided by 3 specialist genitourinary pathologists, was compared to concordance from a cohort of 29 general pathologists. The DL algorithm outperformed general pathologist review with an accuracy of 0.7 (95% confidence interval (CI) 0.65–0.75) compared to 0.61 (95% CI: 0.56–0.66) from the pathologist cohort (p=0.002). Furthermore, the DL approach had a 4-6% lower mean absolute error for quantitation of patterns 3 and 4 compared to the pathologist cohort. Perhaps one of the most striking features of the paper is the requirement for 900 pathologist hours for annotation of the training data set. Similar work by Nir et al(25) evaluated an automatic grading DL CAD tool trained on 331 tissue microarray cores from 231 radical prostatectomies. The validation set included 230

whole mount slides from 56 patients. The CAD tool achieved an overall grading agreement of unweighted kappa 0.51, in line with the agreement between individual pathologists; 0.45 – 0.62. The study employed a novel technique to circumnavigate the issue of arduous training data set labelling by developing an android app (Pathmarker) for expedient tablet and stylus annotation of digital slides by expert pathologists. Lucas et al.(26) also adopted a CNN but instead automatically graded prostate needle biopsy specimens (n=34) and prostatectomy WSI. The CNN produced probability maps and was able to differentiate between benign and malignant tissue (Gleason grade ≥ 3) with an impressive accuracy of 92%. Further differentiation between GS≥4 and GS≤3 demonstrated an accuracy of 90% and there was substantial agreement with expert pathologist reporting (Kappa=0.7). This study was limited by small patient numbers and limited expert pathologist review of the training data set. Multiple other published studies have reported on CAD tools for automatic grading of prostate histopathological specimens with comparable results (27–29).

Besides automatic Gleason grading novel uses of AI in histopathology have been described; Chen et al.(30) demonstrate the fusion of AI with an optical microscope with a real-time DL cancer detection graphic superimposed onto an augmented microscope visual display. The results are preliminary but the concept may gain traction as pathology labs are not required to purchase expensive whole slide digital scanners, currently a prerequisite for implementation of AI enhanced pathology.

2A6). AI in prostate cancer genomics

Currently prostate cancer decision-making relies on risk stratification from clinico-pathological features such as PSA, Gleason grade and clinical staging. Deeper

understanding of prostate cancer at a molecular/genetic level enabled by novel high-throughput genome sequencing technologies has led to increased interest in prostate cancer genomics. AI can analyse large volumes of genomic data to identify genes responsible for disease characteristics such as metastatic potential or castrate resistance. Furthermore, these novel biomarkers can be incorporated into ML predictive algorithms for superior risk-stratification models.

Decipher® is a commercially available 22-RNA feature genomic ML algorithm which predicts early prostate metastasis. The initial validation study(31) included 256 prostatectomy patients and noted an AUC of 0.79 for prediction of 5-year metastasis, significantly outperforming clinico-pathological based prediction techniques. Decipher® is also validated for use with biopsy specimens to predict metastasis and prostate cancer specific mortality following radical treatment(32). Meta-analysis including 5 studies further confirmed the benefit of Decipher® and concluded the tool adds prognostic benefit and should be incorporated into the decision-making process in addition to conventional risk stratification(33). Other clinically available genomic tests utilising machine learning include the 31-gene Prolaris® and the 17-gene Oncotype DX® panels, both capable of predicting recurrence from biopsy diagnosis or post prostatectomy(34). Due to the intra-tumour and inter-patient heterogeneity of prostate cancer traditional methods of risk stratification utilising clinico-pathological features may be blunt instruments. The addition of genomic biomarkers enables personalised precision medicine to accurately predict tumour behaviour and response to treatment. AI is a vital tool in both the identification of novel genomic biomarkers and the process of risk modelling.

2A7). Critical analysis of AI in prostate cancer

Prostate cancer, as the most prevalent urological malignancy, receives considerable attention from AI researchers. While meaningful advances have been made in the application of AI to automatically detect malignancy, predict Gleason grade, reduce inter-observer variability in pathological grading and identify genomic biomarkers, significant barriers to widespread implementation of AI in prostate cancer need to be overcome. For example, key concerns in ML prostate imaging include heterogeneity in research methodology and a lack of robust training data sets making comparison of ML-approaches very difficult. To illustrate the point a recent review by Wildeboer et al.(35) provide an exhaustive list of 83 separate ML/CAD tools for prostate imaging published over a 10-year period. In order for a ML algorithm to 'learn' an annotated training data set is required, this process is a hugely laborious task requiring significant resources from pathologist and radiologist for accurate high-quality labelling. Whilst annotated whole-mount prostatectomy specimens provide excellent training maps for ML algorithms, study populations are inherently biased with a lack of benign conditions and subsequent algorithms are not validated for conditions such as prostatitis or benign prostatic hyperplasia. Furthermore, those algorithms trained on biopsy specimens alone will be contaminated by the sampling error inherent in prostate biopsy techniques, leading to as much as 30% under-diagnosis(36). Additionally, it has been suggested that ML algorithms requires 1000 – 10,000 data pieces for reliable training(37), few of the published studies have sample sizes anywhere near these figures.

One major barrier to realising fully automated AI prostate histopathology is the prerequisite for digital pathology with its associated costly scanning and reporting equipment. A national UK pathologist survey(38) noted only 31% of respondents used digital pathology services for primary diagnosis highlighting the need for greater uptake. A common limitation of pathological studies described is the relative lack of highest-grade disease in the training sets due to its inherent rareness leading to the

introduction of classification bias into the AI algorithms. To overcome this shortcoming augmentation of data sets through artificial synthesis of highest Gleason grade has been described(1) but the technique requires further validation. Once these obstacles are overcome AI will become commonplace because the potential is undeniable and AI will soon be augmenting pathologists to streamline services, adapt to pathologist shortages, improve consensus in grading, facilitate expedient second opinion and support smaller resource-scarce institutions.

3). AI IN RENAL CANCER

Over the last few decades there's been a continual rise in the detection of renal masses owing to increased utilisation of cross-sectional imaging. Commonly those lesions identified are asymptomatic small renal masses (SRMs), up to 30% are found to be benign(39) on histopathology. Renal cancer exhibits heterogeneous histological subtypes affecting prognosis and there are a number of benign lesions such as oncocytomas and fat poor-AMLs that radiologically mimic malignant tumours posing a diagnostic challenge for radiologists. Renal biopsy increases the diagnostic yield in indeterminate lesions but requires an invasive procedure with associated complications that may not be suitable for the obese or co-morbid patient. Combining radiomics with novel AI algorithms has shown promise in improving the diagnostic and prognostic accuracy of cross-sectional renal imaging and is described further. Table 2 outlines the latest AI studies in renal cancer.

3A1). AI in renal cancer imaging

One of the earliest descriptions of AI in renal cancer imaging was the 2015 Yan et al(40). study describing an ANN classifier discriminating fat poor angiomyolipoma (AML), clear cell renal cell cancer (ccRCC), and papillary renal cell cancer (pRCC) with triple phase computed topography (CT). The study reported a classification accuracy of 90.7-100% and adopted a novel radiomics technique, texture analysis, for mathematical exploration of pixel level spatial and temporal heterogeneity within imaged tumours. Numerous further studies(41–43) have applied ML algorithms and texture analysis for classification of renal tumours to differentiate RCC subtypes and discriminate benign from malignant tumours, reporting AUCs approaching 0.90-0.99(44). A large contemporaneous study(45) included 179 patients and differentiated ccRCC from oncoctyoma with an accuracy of 74.4%, sensitivity of 85.8%, and positive predictive value of 80.1%. Perhaps the most striking feature of this study is the relatively small patient number, despite being one of the larger studies, calling in to question the reliability of the training data set.

Beyond subtype classification multiple groups have adopted ML approaches to predict the grade of renal cancer using radiomic features. Higher Fuhrman grades (III-IV) carry a substantially greater metastatic potential than lower grades (I-II)(46). Equipped with such knowledge clinicians can accurately risk assess patients enabling informed shared decision-making in regards of treatment strategies such as nephrectomy, nephron-sparing surgery, focal therapies and surveillance. A small retrospective study by Bektas et al.(47) including 53 patients adopted ML-based enhanced CT texture analysis for classification of high grade versus low grade ccRCC. The group compared different ML strategies for classification, the best performing algorithm used a SVM and achieved an AUC of 0.86 comparable to concordance of renal biopsy grade

prediction with final specimen grade. Kocak et al.(48) elaborated further on this work by comparing the prediction of high/low-grade ccRCC by an ANN compared to conventional binary logistic regression. They took the novel approach of including only unenhanced CT images for texture analysis hypothesising that nuclear grade is not directly associated with tumour vasculature, furthermore their image database included multi-institute CT images with heterogeneous contrast administration protocols. The ANN outperformed multivariate logistic regression achieving diagnostic accuracy of 81.5% and AUC of 0.71 compared to 75.3% and AUC 0.66. Conversely, Lin et al.(49) applied ML texture analysis to contrast enhanced triple-phase CT comparing analysis across the three phases (pre-contrast, corticomedullary and nephrogenic phase). The three-phase combined analysis produced superior performance to single-phase analysis with accuracy of 74%, positive predictive value of 91% and AUC of 0.87. Recently, several studies(50,51) have adopted the international society of pathologist (ISUP) grading system for RCC, replacing the potentially unreliable Fuhrman grading system, and achieving comparable diagnostic accuracy with ML-grade prediction. Understandably, the vast majority of research to date focuses on CT imaging reflecting the widespread use of this modality in renal cancer diagnostics. To the authors knowledge there is currently only one study(52) assessing MRI in grade prediction and this notes comparable diagnostic accuracy between multi-phase CT and mpMRI. The study is noteworthy because unlike many of those previously described, the trained ML algorithm was validated on external data sets providing robust performance data and more accurately reflecting real life practice.

3A2). AI in renal cell cancer histopathology

The merits of digitised histopathology have been well described in the prostate section of this review. Histopathological Fuhrman grading of RCC, similar to prostate cancer Gleason grading, is plagued by inter-observer heterogeneity and poor reproducibility. Inter-pathologist concordance rates have been reported as low as 24%(53). Despite these shortcomings there is a paucity of research in AI RCC histopathological analysis; 4 studies in total(54–57). The earliest work dates back as far as 2014(57) and describes a simple SVM classifier automatically differentiating high from low grade ccRCC based solely on one feature, nucleus size. An impressive AUC of 0.97 was described, albeit, in a small study (n=39) and no impact on patient survival was noted. The largest and most recent study(54) included 277 ccRCC WSIs from the publicly available Cancer Genomic Atlas (TGCA) database. The study assessed 26 histiomic features reporting an AUC of 0.84 for high versus low-grade classification with a LASSO ML predictive model. The studies described have several limitations including the use of Fuhrman grading soon to be replaced with ISUP grading, small sample size, inclusion of ccRCC only and, specifically in relation to the last study described, the need for manual identification by pathologists of regions of interest within the digital slides.

3A3). AI in renal cancer genomics

A number of clinico-pathological risk-stratification prediction models are described for renal cancer in the current iteration of the EAU guidelines. Unfortunately, some early-stage tumours exhibit aggressive pathophysiology underappreciated by current risk modelling. Multiple RCC genetic biomarkers(58) have been identified from publicly available databases such as TGCA. RCC genomic studies may further improve accuracy of risk-stratification, providing individualised medicine and informed decision-making. Li et al.(59) describe a ML centred 15-gene expression risk score model

stratifying ccRCC into high or low-risk categories and noting a significant association with overall and recurrence-free survival. Similar work by Park et al.(60) concentrated solely on T1 tumours and aimed to identify early-stage tumours associated with a poor outcome. Utilising an ANN prediction model assessing expression of FOXC2, PBRM, and BAP1 genes the group predicted synchronous metastasis, recurrence and cancer-specific death. The DNN model outperformed conventional logistic regression observing superior accuracy and AUC 0.85 and 0.80 versus 0.76 and 0.76, respectively. A South Korean prospective-trial by the same study group is currently underway to assess genomic biomarkers of aggressiveness in T1 stage ccRCC (ClinicalTrials.gov Identifier: NCT03694912).

Several studies have investigated the role of radiomic features as surrogates for genomic biomarkers by demonstrating the radiomic association with gene expression related to poor outcomes including BAP1(61), PBMR1(62) and expression of VEGF(63).

3A4). Critical analysis of AI in renal cancer

Early results are promising for the application of non-invasive radiomic features to differentiate RCC subtypes, identify malignant lesions and predict pathological grade. Importantly, some of the most recent research highlights that radiomic biomarkers may be comparable to percutaneous biopsy and thus may represent a viable alternative. Automatic AI Fuhrman grading reduces inter-observer variability and genomic risk modelling may provide an improvement on current clinico-pathological risk-stratification methods. Nevertheless, there are significant limitations to clinical applicability of AI in renal cancer diagnosis and management. Many of the general

limitations have been described in the prostate cancer section. However, specific to renal cancer AI research is almost all studies are retrospective in nature. Additionally, even the largest study has a relatively small sample size, particularly in respect of AI in radiomics, and this limits the quality of ML algorithm training and validity of conclusions drawn. Furthermore, research into non-clear cell RCC is rare and there is scant research on renal cancer histopathology, with those studies conducted often assessing Fuhrman grade, which may soon be replaced by ISUP pathological grading. Large multi-institute prospective studies validated on external data will enable AI to play a major role in the future landscape of renal cancer diagnosis and may identify those patients with aggressive SRMs better suited to radical treatments without the need for biopsy.

4). AI IN BLADDER CANCER

Management of bladder cancer is unique compared to the other already described urological malignancies due to the reliance on endoscopic diagnosis, propensity for recurrence and the widespread practice of intra-vesical therapy and neo-adjuvant chemotherapy. These idiosyncrasies require inventive applications of AI, current research is described in the follow section of the review and detailed further in table 3.

4A1). AI in bladder cancer cystoscopy

Numerous techniques have been developed to improve the efficiency and accuracy of cystoscopy including photo-dynamic diagnosis, narrow-band imaging and optical coherence tomography(64). The application of AI to cystoscopy is a relatively new

concept and early studies focus on assessment of still images taken during cystoscopy. Hashemi et al.(65) developed a multilayer perceptron ML algorithm to detect malignancy from 540 cystoscopic images. The group reported a relatively low accuracy in classification of only 49.3%. Similar work by Lorencin et al.(66), also utilising a multilayer perceptron algorithm, reporting an improved AUC of 0.99 when categorising benign versus malignant lesions from 2983 cystoscopic images. Likewise, Ikeda et al.(67) describes a CNN achieving sensitivity and specificity of 89.7% and 94% in the identification of bladder malignancy from a cohort of 2102 cystoscopy images. These aforementioned preliminary studies demonstrate the strength of ANNs in image recognition but the clinical applicability is limited because cystoscopic diagnosis is a dynamic process not formulated on still images alone. ML systems operating in real-time or used to asses endoscopy videos are well described in other endoscopic procedures such as colonoscopy(68). To date there is only one small study in urology assessing augmentation of cystoscopy using real time ML technology; Shkolyar et al.(69) describe 'CystoNet' a deep learning algorithm trained on 95 cystoscopy videos achieving a malignancy detection sensitivity and specificity of 91% and 99%, respectively, albeit the testing cohort was relatively small with just 54 patients. As this technology advances further, it will facilitate improved training, standardisation in detection rates and democratise high-quality diagnostics to resource-poor countries.

4A2). AI in bladder cancer imaging

Research into AI and bladder cancer imaging is limited compared to prostate and renal cancer explained somewhat by the predominant role of endoscopy in initial diagnosis. Several researchers have combined radiomic features and AI to make the all-important distinction between muscle and non-muscle invasive disease, recognising that the

current reliance on biopsy can lead up to as much as 50% under-staging(70). Garapati et al.(71) developed a ML classifier utilising morphological and textural radiomic features from contrast enhanced CT to categorise bladder tumours into muscle invasive or non-muscle invasive. The data set was small (76 CT scans) and utilised a number of classifiers noting a reasonable AUC of 0.77 – 0.95. MRI provides enhanced soft-tissue contrast and more accurately differentiates muscle invasion compared to CT (72). Zheng et al.(73) describe a ML normogram assessing 23 radiomic features from pre-operative bladder mpMRI able to predict muscle invasion with an AUC of 0.88 on the validation set. Interestingly, the group separately assessed different areas of interest within the tumour itself noting improved performance when assessing just the basal section of the tumour compared analysis of the whole tumour. Furthermore, in this study of 31 patients under-staged on initial TURBT, 28 (90.3%) were correctly staged by the MRI ML normogram. This highlights a potential role of non-invasive ML for pre-operative staging in identifying patients requiring expedient re-resection to avoid under-staging and residual tumour.

Grading of bladder cancer traditionally relies on tumour resection and has important ramifications for risk stratification and future management. Zhang et al.(74) found textural features extracted from bladder mpMRI could accurately differentiate high versus low-grade bladder cancer with an AUC of 0.86, although this was a single-centre retrospective study limiting generalisability. More recently, in a larger study, Wang et al.(75) utilised a LASSO algorithm with bladder mpMRI and observed an AUC of 0.93 for classification of high versus low-grade bladder tumours. Beyond local staging and grading of the bladder tumours, knowledge of lymph node status is vitally important as the presence of lymph node metastasis is a poor prognostic sign. Current staging relies exclusively on CT/MRI size criteria and normally sized metastatic nodes may be overlooked, hence the poor sensitivity of CT and MRI for lymph nodes metastasis described (31-45%)(76). Wu et al.(76) assessed 103 post cystectomy +

extended pelvic lymph node dissection (PLND) patients adopting a LASSO ML algorithm to identify and validate radiomic features associated with lymph node metastasis and achieved an AUC of 0.84 for prediction of lymph node metastasis. Armed with such knowledge pelvic oncologists will be able to better identify patients suitable for extended dissection templates and those more likely to benefit from neo-adjuvant systemic therapy.

Neoadjuvant cisplatin-based chemotherapy leads to an 8% absolute improvement in in 5-year survival(77). However, not all patients respond to chemotherapy and thus may progress during systemic treatment whilst also being exposed to potential toxic side-effects. Several researchers have applied AI methods to non-invasively predict and assess therapeutic response; Wu et al.(78) compared a number of ANNs to assess pre-and post-treatment CT scans for tumour chemotherapy response. The best performing ML algorithm achieved similar performance outcomes to 2 trained radiologists. Similar work by Cha et al.(79) combining radiomic features with ML algorithms noting equivalent performance in assessing tumour response and concluded that such CAD tools can augment radiologists by providing an expedient second review. Assessment of response during treatment through non-invasive imaging will enable clinicians to alter chemotherapy or offer early cystectomy in non-responders reducing morbidity and mortality.

4A3). AI in bladder cancer histopathology

Similar to prostate and renal cancer AI has been applied to histopathological analysis of bladder cancer to automate the process, improving reproducibility and efficiency, and to discover novel prognostic biomarkers. An excellent study by Zhang et al.(80) applied a ML CAD tool to automate analysis of 913 bladder cancer WSIs and

compared the tool to review by 17 expert pathologists. The system outperformed or at least matched individual pathologists on the reporting of high and low-grade cancer with an AUC of 0.97. There was inter-pathologist disagreement of 23.8% highlighting the strength of the automated system in improving reproducibility. The group adopted a novel approach by utilising natural language descriptions of microscopic findings (tumour appearance, cell morphology etc.) and visual annotations of analysed slides allowing easy and interpretable second review by human pathologists. Beyond WSI interpretation researchers have utilised AI to analyse novel pathological features. For example, tumour budding (TB), is a new concept describing the clustering of cancer cells with infiltrative growth patterns. Increased density of these cells is a bad prognostic factor in non-urological cancers and may represent an early step towards metastasis(81,82). Automated ML quantification of TB in bladder WSIs reveals that increased TB is associated with worse disease-specific survival and the prognostic accuracy outperforms standard clinico-pathological risk-stratification(83). A unique application of AI to histopathology is described by Glaser et al.(84), who recognised that up to 20% of pathology reports on transurethral resection do not adequately provide sufficient information on bladder cancer stage, largely owing to their descriptive free-text nature of reports. The group applied a natural language processing algorithm to scrutinise the written text of 1638 resection pathology reports identifying and collating information such as muscle inclusion, stage and grade. This allowed individual surgeons to monitor the quality of their resection through assessment of detrusor muscle inclusion and acted as quality control for pathologist reporting. This study highlights the capability of AI to automatically extract categorical data from free-text, such as medical notes or reports. This application will no doubt become commonplace in a variety of contexts as paper medical notes are increasingly superseded by electronic health records.

4A4). AI in bladder cancer genomics

The European Organisation for Research and Treatment of Cancer (EORTC) Genito-urinary cancer group provide a risk of recurrence and progression prediction model(85) based on 6 clinico-pathological features; number of tumours, tumour diameter, prior recurrence rate, stage, grade and presence of carcinoma-in-situ (CIS). Further models have been created to assess risk in those patients receiving intravesival therapies such as mitomycin-C and Bacillus Calmette-Guerin(86,87). A number of oncogenic genes are associated with pathogenesis in bladder cancer and combining AI with genomic analysis will enable more accurate cancer risk-stratification.

Bartsch et al.(88) developed a ML classifier identifying 21 genes predictive of 5-year recurrence from 112 frozen initial TURBT pathology specimens. The study included non-muscle invasive specimens only and was able to accurately predict 5-year recurrence with a sensitivity and specificity of 71% and 67% respectively. The study is noteworthy due to the long follow-up period, particularly when considering AI prediction is a relatively new phenomenon. Furthermore, the group identified novel genomic biomarkers associated with recurrence not previously investigated. However, due to the low progression rate (17%) no genomic markers of progression could be reported, it is also important to note the classifier was not tested on external data and thus its reproducibility is questionable.

As previously described, pre-operative assessment of lymph node metastasis is important when deciding on neo-adjuvant chemotherapy and the extent of PLND. Radiomic features associated with prediction have been detailed earlier in this review. Wu et al.(89) adopted an alternate approach identifying 5 lymph node status related genes from 424 patients undergoing cystectomy and PLND. The group incorporated clinico-pathological metrics with a ML genomic classifier to produce a lymph node

status predictive normogram. Notably, the combined normogram was validated on two external data sets and in terms of predictive ability achieved an AUC of 0.89 outperforming either the individual genomic classifier (AUC=0.73) or the clinico-pathological model (AUC=0.82). The study improves upon previously reported lymph-node status genomic classifiers(90,91) in two respects; first the combination of clinical and genomic features for AI predictive purposes is a unique approach rarely seen elsewhere in urological AI literature(88) and second, the normogram is validated on external data, verifying its real-world clinical utility.

4A5). Critical analysis of AI in bladder cancer

The combination of AI with cystoscopy is an interesting concept and whilst full automation of the process is unlikely to transpire anytime soon, augmentation in a hybrid approach could add diagnostic value. Current research in this area is limited to 3 studies only and all, except 1, involve static cystoscopic images, limiting the clinical applicability. However, progress in other endoscopy procedures such as colonoscopy and oesophago-gastro-duodenoscopy is encouraging and proves that accurate automatic real-time detection of malignancy is currently achievable. Given the current paradigm expects an often novice cystoscopist to operate, detect, analyse and instigate treatment independently, AI is well placed to offer tangible benefits.

AI in bladder cancer imaging shows promise in providing more accurate staging when compared to traditional biopsy methods alone, the importance of which cannot be understated given the rapidity of progression in high-grade disease. Furthermore, accurate non-surgical diagnostic methods are vitally important in an often highly co-morbid elderly patient population. Assessing chemotherapy response during treatment

itself will aid in identifying those patients better suited to early surgery and prediction of lymph node metastasis will further guide the treatment decision-making process. Advantages of automated pathological reporting have already been well described earlier and genomics in bladder cancer will improve upon current risk stratification models. However, AI bladder cancer research has not reached sufficient maturity to provide real clinical benefit, studies are small, retrospective and often lack external validation. Once these shortcomings are overcome AI in bladder cancer will make great strides in promoting bladder cancer from it's often quoted 'Cinderella cancer' status providing substantial benefits to clinicians and patients.

5). AI IN ROBOTIC SURGERY

Current robotic systems operate a master-slave relationship; the surgeon inputs operative instructions through a control interface enacted by a robotic system through end-effectors such as laparoscopic instruments. The benefits of robotic surgery are well established and include improved dexterity, tremor removal, optical magnification and fatigue-free surgery(92). Combining these advantages with AI's ability to analyse realms of robotic kinematic data will enable rapid dissemination and standardisation of surgical technique though shared 'cloud' ML. A complete paradigm shift towards cognisant autonomous surgical robots may seem implausible but nevertheless, with the advances in AI diagnostics already nearing clinical applicability, the opportunity in the operating theatre is exciting. Table 4 outlines the latest AI research in robotic surgery.

5A1). Robotic automation

Attempts have been made to automate certain steps of surgical procedures; Shademan et al.(93) demonstrated a Smart Tissue Autonomous Robot (STAR) that outperformed expert surgeons in relation to suture spacing and leak testing in an ex-vivo porcine intestinal anastomosis model. The STAR robot has since been applied to a squamous cell carcinoma resection model achieving an oncologically safe resection margin(94). Remarkably, the only foray of automated robotics in urology harks back to the 1980s with John Wickam's automated transurethral resection of prostate robot entitled the PROBOT(95). This system never reached mass production but proved automation was feasible with acceptable outcomes.

5A2). Robotic skill assessment

An area where AI shows early promise is in the assessment of surgical proficiency. ML algorithms have been used to asses a range of robotic kinematic metrics such as speed, smoothness, dexterity and camera manipulation to ascertain the expertise of robotic surgeons within only a few seconds of task initiation and with accuracy approaching 90%(96,97). Ershad et al.(98) describe the novel approach of attaching movement tracking devices to the shoulders, elbows and wrists of 14 surgeons undertaking robotic simulation exercises. This approach augmented the kinematic data provided from the robotic system with skill level classification improving by 69%. The clinical implication of these skill classification studies beyond merely grading individual surgeons is the recognition that skill and experience can be broken down into data allowing future automated systems to acquire and rapidly disseminate technical skills.

5A3). AI Robotic outcome prediction

The logical progression of AI robotic skill evaluation is prediction of post-operative clinical outcome; the DL algorithm developed by Hung et al.(99) predicted 3- and 6-month urinary incontinence rates following robotic assisted laparoscopic prostatectomy (RALP) based on surgeon performance features including operative time, instrument kinematics, camera movement, system event and Endo-wrist articulation metrics. The features having the greatest impact on incontinence rates were those assessed during apical dissection and vesico-urethral anastomosis. Prediction accuracy was greater when combining the DL assessment of performance metrics with clinico-pathological features compared to either used in isolation (Concordance index 59.9% vs 56.2%, respectively). The same group have demonstrated ML prediction of other surgical outcomes based on kinematic data including length of stay, duration of catheter post RALP and operative time(100). This data is useful for optimisation of theatre and efficiencies here may offset some of the often-criticised increased cost of robotic surgery. Research in this area is preliminary but provides interesting insight into truly objective assessment of surgical performance beyond the current paradigm of peer-assessment and prior case load numbers. Interestingly, Hung et al's study group aim to apply the described principles to recovery of erectile function and risk of biochemical recurrence, representing a complete AI trifecta assessment of RALP.

6). AI UROLOGICAL ONCOLOGY CHALLENGES

Whilst AI in urological oncology has rightfully received much attention for its unparalleled potential, it is prudent to remain wary of such panaceas and consider the significant challenges to widespread implementation in clinical practice. Several cancer-specific limitations have been outlined earlier in the relevant sections.

However, there are general shortcomings of AI research worthy of mention; First, there is considerable heterogeneity in research methodology, AI algorithm design and definition of measured outcomes. This multiplicity makes meaningful comparison and quantitative analysis difficult. Second, significant concerns exist over the generalisability and robustness of results because the majority AI algorithms described are trained, validated and tested on the same dataset leading to a statistical phenomenon known as 'overfitting', whereby models perform favourably on their own data but poorly when applied to novel data. This point is compounded by a scarcity of multi-institute prospective studies, an inevitable skew towards large training datasets and small validation/testing sets and a process of data annotation so laborious that many study sample sizes are too small to draw meaningful conclusions. Third, the methods used to asses model/algorithm's predictive performance can be misleading. Those papers quoting accuracy only (correct predictions / total number of predictions) may be subject to the "accuracy paradox". For example, if the dataset is skewed towards high-grade tumours, a high accuracy displayed by the AI to detect high-grade tumours may just be reflective of the underlying skew in distribution within the dataset. AUC analysis provides a more useful analysis of performance, particularly with imbalanced data sets. Fourth, unsurprisingly there is a research bias towards the most common urological cancers with penile and testicular cancer largely overlooked. The provided review is guilty of such an omission and further research focus is required on the less common malignancies as well as benign urological conditions. Fifth, the inner-workings of many ML algorithms, particularly ANNs, are so complex they become un-interpretable leading to the often-cited 'black box' criticism of AI. The proprietary nature of most AI algorithms produces an extra barrier to rigorous testing of the most promising AI algorithms by other researchers on external datasets. Medicine as a scientific discipline founded in rationalisation will inevitably distrust such a lack of transparency and rightfully so as inherent unforeseen biases within ML processes could lead to disastrous consequences in patient care, particularly when considering

the example of AI automation in robotic surgery. Finally, there are logistical and ethical issues requiring consideration prior to implementation such as limitations on computational power, storage issues associated with huge data sets, concerns over access to patient sensitive data and culpability dilemmas in relation to decision-making and automation in surgery.

7). FUTURE DIRECTIONS

A main strength of AI in urological cancer is the ability to interrogate large volumes of data for powerful predictive modelling. Future research will need to focus on the establishment of multi-institute open access databases to allow improved ML training and validation. Furthermore, the unchecked proliferation of unstructured data due to increased electronic health record utilisation, wide spread multi-parametric imaging, advanced genomic profiling and increase in ‘omics’ fields requires integration and organisation in order for AI to maximise its potential and achieve clinical applicability. Fused analysis of the various data streams is currently an under-researched area and will likely hold the key to powerful prognostication, early prediction of disease such as urological cancer and truly personalised healthcare. Further research into integration of robotics and AI is an intriguing field and may lead to automation of certain basic repetitive surgical processes freeing up the surgeon for more complex steps in surgical procedures. Furthermore, the deconstruction of surgical skill into transferrable data is an exciting concept and will lead to global democratisation of surgery through rapid skill dissemination, especially in the context of anticipated low-latency 5G networks. Improvements in ML automatic radiological and histopathological interpretation is an obvious extension of AI and current research has already reached expert level diagnostic accuracy. Clinical application will enable more efficient reproducible results easing the workload on radiologists and pathologists alike. As with all surgical research

there is a requirement for high quality prospective randomised multi-institute studies to ensure the benefits of AI in urological oncology are confirmed with proper scientific rigour.

CONCLUSION

This review provides a foundation in basic AI principles, critically appraises the impact of AI in urological oncology and delivers a wider commentary on the future direction and challenges facing AI implementation in healthcare. In order to overcome such challenges computer scientists, pathologists, radiologists, surgeons and policy makers will need to collaborate and appreciate that human expertise will always be necessary in the training and development of AI systems. Furthermore, human qualities such as intuition, compassion, reasoning and experience will never be replaced by AI but rather augmented and in the near future this combination will deliver an immeasurably superior healthcare service for all patients, not just those with urological malignancies.

1. Goldenberg SL, Nir G, Salcudean SE. A new era: artificial intelligence and machine learning in prostate cancer. Nature Reviews Urology. 2019.
2. Department of Business Energy and Industrial Strategy. Policy Paper: The Grand Challenges. 2019.
3. The Royal College of Surgeons of England. Future of Surgery Report. 2018;
4. Sarkar BK. Big data for secure healthcare system: a conceptual design. Complex Intell Syst. 2017;
5. Bhandari M, Zeffiro T, Reddiboina M. Artificial intelligence and robotic surgery: Current perspective and future directions. Current Opinion in Urology. 2020.
6. El Aboudi N, Benhlima L. Big Data Management for Healthcare Systems: Architecture, Requirements, and Implementation. Advances in Bioinformatics. 2018.
7. Genomics England. The National Genomics Research and Healthcare Knowledgebase v5 [Internet]. The National Genomics Research and Healthcare Knowledgebase v5. 2019 [cited 2020 Jul 27]. p. 1–91. Available from: https://www.genomicsengland.co.uk/about-genomics-england/the-100000-genomes-project/
8. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast cancer screening. Nature. 2020;
9. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images are more than pictures, they are data. Radiology. 2016;

10. Harmon SA, Tuncer S, Sanford T, Choyke PL, Türkbey B. Artificial intelligence at the intersection of pathology and radiology in prostate cancer. Diagnostic Interv Radiol. 2019;
11. Hambrock T, Vos PC, Hulsbergen-Van De Kaa CA, Barentsz JO, Huisman HJ. Prostate cancer: Computer-aided diagnosis with multiparametric 3-T MR imaging - Effect on observer performance. Radiology. 2013;
12. Karimi D, Ruan D. Synergistic combination of learned and hand-crafted features for prostate lesion classification in multiparametric magnetic resonance imaging. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2017.
13. Wang J, Wu CJ, Bao ML, Zhang J, Wang XN, Zhang YD. Machine learning-based analysis of MR radiomics can help to improve the diagnostic performance of PI-RADS v2 in clinically relevant prostate cancer. Eur Radiol. 2017;
14. Gaur S, Lay N, Harmon SA, Doddakashi S, Mehralivand S, Argun B, et al. Can computer-aided diagnosis assist in the identification of prostate cancer on prostate MRI? A multi-center, multi-reader investigation. Oncotarget. 2018;
15. Zhong X, Cao R, Shakeri S, Scalzo F, Lee Y, Enzmann DR, et al. Deep transfer learning-based prostate cancer classification using 3 Tesla multi-parametric MRI. Abdom Radiol. 2019;
16. Varghese B, Chen F, Hwang D, Palmer SL, De Castro Abreu AL, Ukimura O, et al. Objective risk stratification of prostate cancer using machine learning and radiomics applied to multiparametric magnetic resonance images. Sci

Rep. 2019;

17. Cao R, Mohammadian Bajgiran A, Afshari Mirak S, Shakeri S, Zhong X, Enzmann D, et al. Joint Prostate Cancer Detection and Gleason Score Prediction in mp-MRI via FocalNet. IEEE Trans Med Imaging. 2019;
18. Cuocolo R, Cipullo MB, Stanzione A, Ugga L, Romeo V, Radice L, et al. Machine learning applications in prostate cancer magnetic resonance imaging. European Radiology Experimental. 2019.
19. Anas EMA, Mousavi P, Abolmaesumi P. A deep learning approach for real time prostate segmentation in freehand ultrasound guided biopsy. Med Image Anal. 2018;
20. Armato SG, Huisman H, Drukker K, Hadjiiski L, Kirby JS, Petrick N, et al. PROSTATEx Challenges for computerized classification of prostate lesions from multiparametric magnetic resonance images. J Med Imaging. 2018;
21. Epstein JI, Egevad L, Amin MB, Delahunt B, Srigley JR, Humphrey PA. The 2014 international society of urological pathology (ISUP) consensus conference on gleason grading of prostatic carcinoma definition of grading patterns and proposal for a new grading system. Am J Surg Pathol. 2016;
22. Egevad L, Ahmad AS, Algaba F, Berney DM, Boccon-Gibod L, Compérat E, et al. Standardization of Gleason grading among 337 European pathologists. Histopathology. 2013;
23. Zarella MD, Bowman D, Aeffner F, Farahani N, Xthona A, Absar SF, et al. A practical guide to whole slide imaging a white paper from the digital pathology association. Archives of Pathology and Laboratory Medicine. 2019.
24. Nagpal K, Foote D, Liu Y, Chen P-HC, Wulczyn E, Tan F, et al. Development

and validation of a deep learning algorithm for improving Gleason scoring of prostate cancer. npj Digit Med. 2019;

25. Nir G, Hor S, Karimi D, Fazli L, Skinnider BF, Tavassoli P, et al. Automatic grading of prostate cancer in digitized histopathology images: Learning from multiple experts. Med Image Anal. 2018;
26. Lucas M, Jansen I, Savci-Heijink CD, Meijer SL, de Boer OJ, van Leeuwen TG, et al. Deep learning for automatic Gleason pattern classification for grade group determination of prostate biopsies. Virchows Arch. 2019;
27. Arevalo J, Cruz-roa A. Histopathology Image Representation for Automatic Analysis: a State-of-the-Art Review. Rev Med. 2014;
28. Mosquera-Lopez C, Agaian S, Velez-Hoyos A, Thompson I. Computer-Aided Prostate Cancer Diagnosis from Digitized Histopathology: A Review on Texture-Based Systems. IEEE Rev Biomed Eng. 2015;
29. Angel Arul Jothi J, Mary Anita Rajam V. A survey on automated cancer diagnosis from histopathology images. Artif Intell Rev. 2017;
30. Chen PHC, Gadepalli K, MacDonald R, Liu Y, Kadowaki S, Nagpal K, et al. An augmented reality microscope with real-time artificial intelligence integration for cancer diagnosis. Nat Med. 2019;
31. Jeffrey Karnes R, Bergstralh EJ, Davicioni E, Ghadessi M, Buerki C, Mitra AP, et al. Validation of a genomic classifier that predicts metastasis following radical prostatectomy in an at risk Patient population. J Urol. 2013;
32. Nguyen PL, Haddad Z, Ross AE, Martin NE, Deheshi S, Lam LLC, et al. Ability of a Genomic Classifier to Predict Metastasis and Prostate Cancer-specific Mortality after Radiation or Surgery based on Needle Biopsy Specimens. Eur Urol. 2017;

33. Spratt DE, Yousefi K, Deheshi S, Ross AE, Den RB, Schaeffer EM, et al. Individual patient-level meta-Analysis of the performance of the decipher genomic classifier in high-risk men after prostatectomy to predict development of metastatic disease. J Clin Oncol. 2017;
34. Loeb S, Ross AE. Genomic testing for localized prostate cancer: where do we go from here? Current Opinion in Urology. 2017.
35. Wildeboer RR, van Sloun RJG, Wijkstra H, Mischi M. Artificial intelligence in multiparametric prostate cancer imaging with focus on deep-learning methods. Computer Methods and Programs in Biomedicine. 2020.
36. Ukimura O, De Castro Abreu AL, Gill IS, Shoji S, Hung AJ, Bahn D. Image visibility of cancer to enhance targeting precision and spatial mapping biopsy for focal therapy of prostate cancer. BJU Int. 2013;
37. Choyke PL. Quantitative MRI or machine learning for prostate MRI: Which should you use? Radiology. 2018.
38. Williams BJ, Lee J, Oien KA, Treanor D. Digital pathology access and usage in the UK: Results from a national survey on behalf of the National Cancer Research Institute's CM-Path initiative. J Clin Pathol. 2018;
39. Patel HD, Johnson MH, Pierorazio PM, Sozio SM, Sharma R, Iyoha E, et al. Diagnostic accuracy and risks of biopsy in the diagnosis of a renal mass suspicious for localized renal cell carcinoma: Systematic review of the literature. J Urol. 2016;
40. Yan L, Liu Z, Wang G, Huang Y, Liu Y, Yu Y, et al. Angiomyolipoma with Minimal Fat: Differentiation From Clear Cell Renal Cell Carcinoma and Papillary Renal Cell Carcinoma by Texture Analysis on CT Images. Acad Radiol. 2015;

41. Yu HS, Scalera J, Khalid M, Touret AS, Bloch N, Li B, et al. Texture analysis as a radiomic marker for differentiating renal tumors. Abdom Radiol. 2017;
42. Feng Z, Rong P, Cao P, Zhou Q, Zhu W, Yan Z, et al. Machine learning-based quantitative texture analysis of CT images of small renal masses: Differentiation of angiomyolipoma without visible fat from renal cell carcinoma. Eur Radiol. 2018;
43. Cui EM, Lin F, Li Q, Li RG, Chen XM, Liu ZS, et al. Differentiation of renal angiomyolipoma without visible fat from renal cell carcinoma by machine learning based on whole-tumor computed tomography texture features. Acta radiol. 2019;
44. Suarez-Ibarrola R, Hein S, Reis G, Gratzke C, Miernik A. 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. World Journal of Urology. 2019.
45. Coy H, Hsieh K, Wu W, Nagarajan MB, Young JR, Douek ML, et al. Deep learning and radiomics: the utility of Google TensorFlow™ Inception in classifying clear cell renal cell carcinoma and oncocytoma on multiphasic CT. Abdom Radiol. 2019;
46. Gumundsson E, Hellborg H, Lundstam S, Erikson S, Ljungberg B. Metastatic potential in renal cell carcinomas ≤7 cm: Swedish kidney cancer quality register data. Eur Urol. 2011;
47. Bektas CT, Kocak B, Yardimci AH, Turkcanoglu MH, Yucetas U, Koca SB, et al. Clear Cell Renal Cell Carcinoma: Machine Learning-Based Quantitative Computed Tomography Texture Analysis for Prediction of Fuhrman

Nuclear Grade. Eur Radiol. 2019;

48. Kocak B, Durmaz ES, Ates E, Kaya OK, Kilickesmez O. Unenhanced CT texture analysis of clear cell renal cell carcinomas: A machine learning-based study for predicting histopathologic nuclear grade. Am J Roentgenol. 2019;
49. Lin F, Cui EM, Lei Y, Luo L ping. CT-based machine learning model to predict the Fuhrman nuclear grade of clear cell renal cell carcinoma. Abdom Radiol. 2019;
50. Shu J, Wen D, Xi Y, Xia Y, Cai Z, Xu W, et al. Clear cell renal cell carcinoma: Machine learning-based computed tomography radiomics analysis for the prediction of WHO/ISUP grade. Eur J Radiol [Internet]. 2019;121:108738. Available from: http://www.sciencedirect.com/science/article/pii/S0720048X19303882
51. Sun X, Liu L, Xu K, Li W, Huo Z, Liu H, et al. Prediction of ISUP grading of clear cell renal cell carcinoma using support vector machine model based on CT images. Medicine (Baltimore). 2019;
52. Cui E, Li Z, Ma C, Li Q, Lei Y, Lan Y, et al. Predicting the ISUP grade of clear cell renal cell carcinoma with multiparametric MR and multiphase CT radiomics. Eur Radiol. 2020;
53. Lang H, Lindner V, De Fromont M, Molinié V, Letourneux H, Meyer N, et al. Multicenter determination of optimal interobserver agreement using the Fuhrman grading system for renal cell carcinoma: Assessment of 241 patients with > 15-year follow-up. Cancer. 2005;
54. Tian K, Rubadue CA, Lin DI, Veta M, Pyle ME, Irshad H, et al. Automated clear cell renal carcinoma grade classification with prognostic significance.

PLoS One. 2019;

55. Kruk M, Kurek J, Osowski S, Koktysz R, Swiderski B, Markiewicz T. Ensemble of classifiers and wavelet transformation for improved recognition of Fuhrman grading in clear-cell renal carcinoma. Biocybern Biomed Eng. 2017;
56. Holdbrook DA, Singh M, Choudhury Y, Kalaw EM, Koh V, Tan HS, et al. Automated Renal Cancer Grading Using Nuclear Pleomorphic Patterns. JCO Clin Cancer Informatics. 2018;
57. Ho C, Yeh F-C, Parwani A, Pantanowitz L. Automated grading of renal cell carcinoma using whole slide imaging. J Pathol Inform. 2014;
58. Mitchell TJ, Rossi SH, Klatte T, Stewart GD. Genomics and clinical correlates of renal cell carcinoma. World J Urol. 2018;
59. Li P, Ren H, Zhang Y, Zhou Z. Fifteen-gene expression based model predicts the survival of clear cell renal cell carcinoma. Med (United States). 2018;
60. Park JS, Lee HJ, Cho NH, Kim J, Jang WS, Heo JE, et al. Risk Prediction Tool for Aggressive Tumors in Clinical T1 Stage Clear Cell Renal Cell Carcinoma Using Molecular Biomarkers. Comput Struct Biotechnol J. 2019;
61. Ghosh P, Tamboli P, Vikram R, Rao A. Imaging-genomic pipeline for identifying gene mutations using three-dimensional intra-tumor heterogeneity features. J Med Imaging. 2015;
62. Kocak B, Durmaz ES, Ates E, Ulusan MB. Radiogenomics in clear cell renal cell carcinoma: Machine learning–based high-dimensional quantitative CT texture analysis in predicting PBRM1 mutation status. Am J Roentgenol. 2019;
63. Yin Q, Hung S-C, Wang L, Lin W, Fielding JR, Rathmell WK, et al.

Associations between Tumor Vascularity, Vascular Endothelial Growth Factor Expression and PET/MRI Radiomic Signatures in Primary Clear-Cell-Renal-Cell-Carcinoma: Proof-of-Concept Study. Sci Rep. 2017 Mar;7:43356.

64. Jeong BC. Chapter 10 - Recent Technological Advances in Cystoscopy for the Detection of Bladder Cancer. In: Ku JHBT-BC, editor. Academic Press; 2018. p. 135–44. Available from: http://www.sciencedirect.com/science/article/pii/B9780128099391000102

65. Hashemi SMR, Hassanpour H, Kozegar E, Tan T. Cystoscopic Image Classification Based on Combining MLP and GA. Int J Nonlinear Anal Appl [Internet]. 2020;11(1):93–105. Available from: https://ijnaa.semnan.ac.ir/article_4232.html

66. Lorencin I, Anđelić N, Španjol J, Car Z. Using multi-layer perceptron with Laplacian edge detector for bladder cancer diagnosis. Artif Intell Med. 2020;

67. Ikeda A, Nosato H, Kochi Y, Kojima T, Kawai K, Sakanashi H, et al. Support System of Cystoscopic Diagnosis for Bladder Cancer Based on Artificial Intelligence. J Endourol. 2020;

68. Urban G, Tripathi P, Alkayali T, Mittal M, Jalali F, Karnes W, et al. Deep Learning Localizes and Identifies Polyps in Real Time With 96% Accuracy in Screening Colonoscopy. Gastroenterology. 2018;

69. Shkolyar E, Jia X, Chang TC, Trivedi D, Mach KE, Meng MQH, et al. Augmented Bladder Tumor Detection Using Deep Learning. Eur Urol. 2019;

70. Schwaibold HE, Sivalingam S, May F, Hartung R. The value of a second transurethral resection for T1 bladder cancer. BJU Int. 2006;
71. Garapati SS, Hadjiiski L, Cha KH, Chan HP, Caoili EM, Cohan RH, et al. Urinary bladder cancer staging in CT urography using machine learning. Med Phys. 2017;
72. McKibben MJ, Woods ME. Preoperative Imaging for Staging Bladder Cancer. Current Urology Reports. 2015.
73. Zheng J, Kong J, Wu S, Li Y, Cai J, Yu H, et al. Development of a noninvasive tool to preoperatively evaluate the muscular invasiveness of bladder cancer using a radiomics approach. Cancer. 2019;
74. Zhang X, Xu X, Tian Q, Li B, Wu Y, Yang Z, et al. Radiomics assessment of bladder cancer grade using texture features from diffusion-weighted imaging. J Magn Reson Imaging. 2017;
75. Wang H, Hu D, Yao H, Chen M, Li S, Chen H, et al. Radiomics analysis of multiparametric MRI for the preoperative evaluation of pathological grade in bladder cancer tumors. Eur Radiol. 2019;
76. Wu S, Zheng J, Li Y, Wu Z, Shi S, Huang M, et al. Development and Validation of an MRI-Based Radiomics Signature for the Preoperative Prediction of Lymph Node Metastasis in Bladder Cancer. EBioMedicine. 2018 Aug;34:76–84.
77. Yin M, Joshi M, Meijer RP, Glantz M, Holder S, Harvey HA, et al. Neoadjuvant Chemotherapy for Muscle-Invasive Bladder Cancer: A Systematic Review and Two-Step Meta-Analysis. Oncologist. 2016;
78. Wu E, Hadjiiski LM, Samala RK, Chan HP, Cha KH, Richter C, et al. Deep Learning Approach for Assessment of Bladder Cancer Treatment

Response. Tomogr (Ann Arbor, Mich). 2019;

79. Cha KH, Hadjiiski, PhD LM, Cohan, MD RH, Chan, PhD HP, Caoili, MD EM, Davenport, MD M, et al. Diagnostic Accuracy of CT for Prediction of Bladder Cancer Treatment Response with and without Computerized Decision Support. Acad Radiol. 2019;
80. Zhang Z, Chen P, McGough M, Xing F, Wang C, Bui M, et al. Pathologist-level interpretable whole-slide cancer diagnosis with deep learning. Nat Mach Intell. 2019;
81. Karamitopoulou E, Zlobec I, Born D, Kondi-Pafiti A, Lykoudis P, Mellou A, et al. Tumour budding is a strong and independent prognostic factor in pancreatic cancer. Eur J Cancer. 2013;
82. Van Wyk HC, Park JH, Edwards J, Horgan PG, McMillan DC, Going JJ. The relationship between tumour budding, the tumour microenvironment and survival in patients with primary operable colorectal cancer. Br J Cancer. 2016;
83. Brieu N, Gavriel CG, Nearchou IP, Harrison DJ, Schmidt G, Caie PD. Automated tumour budding quantification by machine learning augments TNM staging in muscle-invasive bladder cancer prognosis. Sci Rep. 2019;
84. Glaser AP, Jordan BJ, Cohen J, Desai A, Silberman P, Meeks JJ. Automated Extraction of Grade, Stage, and Quality Information From Transurethral Resection of Bladder Tumor Pathology Reports Using Natural Language Processing. JCO Clin Cancer Informatics. 2018;
85. Sylvester RJ, Van Der Meijden APM, Oosterlinck W, Witjes JA, Bouffioux C, Denis L, et al. Predicting recurrence and progression in individual patients with stage Ta T1 bladder cancer using EORTC risk tables: A combined

analysis of 2596 patients from seven EORTC trials. Eur Urol. 2006;

86. Lammers RJM, Hendriks JCM, Rodriguez Faba ORF, Witjes WPJ, Palou J, Witjes JA. Prediction model for recurrence probabilities after intravesical chemotherapy in patients with intermediate-risk non-muscle-invasive bladder cancer, including external validation. World J Urol. 2016;
87. Fernandez-Gomez J, Madero R, Solsona E, Unda M, Martinez-Piñeiro L, Gonzalez M, et al. Predicting Nonmuscle Invasive Bladder Cancer Recurrence and Progression in Patients Treated With Bacillus Calmette-Guerin: The CUETO Scoring Model. J Urol. 2009;
88. Bartsch G, Mitra AP, Mitra SA, Almal AA, Steven KE, Skinner DG, et al. Use of Artificial Intelligence and Machine Learning Algorithms with Gene Expression Profiling to Predict Recurrent Nonmuscle Invasive Urothelial Carcinoma of the Bladder. J Urol. 2016;
89. Wu SX, Huang J, Liu ZW, Chen HG, Guo P, Cai QQ, et al. A Genomic-clinicopathologic Nomogram for the Preoperative Prediction of Lymph Node Metastasis in Bladder Cancer. EBioMedicine. 2018;
90. Smith SC, Baras AS, Dancik G, Ru Y, Ding KF, Moskaluk CA, et al. A 20-gene model for molecular nodal staging of bladder cancer: Development and prospective assessment. Lancet Oncol. 2011;
91. Seiler R, Lam LL, Erho N, Takhar M, Mitra AP, Buerki C, et al. Prediction of Lymph Node Metastasis in Patients with Bladder Cancer Using Whole Transcriptome Gene Expression Signatures. J Urol. 2016;
92. McGuinness LA, Prasad Rai B. Robotics in urology. Ann R Coll Surg Engl [Internet]. 2018 May 1;100(6_sup):45–54. Available from: https://doi.org/10.1308/rcsann.supp1.38

93. Shademan A, Decker RS, Opfermann JD, Leonard S, Krieger A, Kim PCW. Supervised autonomous robotic soft tissue surgery. Sci Transl Med. 2016;
94. Opfermann JD, Leonard S, Decker RS, Uebele NA, Bayne CE, Joshi AS, et al. Semi-autonomous electrosurgery for tumor resection using a multi-degree of freedom electrosurgical tool and visual servoing. In: IEEE International Conference on Intelligent Robots and Systems. 2017.
95. Challacombe BJ, Khan MS, Murphy D, Dasgupta P. The history of robotics in urology. World J Urol. 2006;
96. Fard MJ, Ameri S, Darin Ellis R, Chinnam RB, Pandya AK, Klein MD. Automated robot-assisted surgical skill evaluation: Predictive analytics approach. Int J Med Robot. 2018 Feb;14(1).
97. Wang Z, Majewicz Fey A. Deep learning with convolutional neural network for objective skill evaluation in robot-assisted surgery. Int J Comput Assist Radiol Surg. 2018;
98. Ershad M, Rege R, Majewicz Fey A. Automatic and near real-time stylistic behavior assessment in robotic surgery. Int J Comput Assist Radiol Surg. 2019;
99. Hung AJ, Chen J, Ghodoussipour S, Oh PJ, Liu Z, Nguyen J, et al. A deep-learning model using automated performance metrics and clinical features to predict urinary continence recovery after robot-assisted radical prostatectomy. BJU Int. 2019;
100. Hung AJ, Chen J, Che Z, Nilanon T, Jarc A, Titus M, et al. Utilizing Machine Learning and Automated Performance Metrics to Evaluate Robot-Assisted Radical Prostatectomy Performance and Predict Outcomes. J Endourol. 2018;