



**HAL**  
open science

# The Application of Artificial Intelligence in Prostate Cancer Management-What Improvements Can Be Expected? A Systematic Review

Ronan Thenault, Kevin Kaulanjan, Thomas Darde, Nathalie Rioux-Leclercq, Karim Bensalah, Marie Mermier, Zine-Eddine Khene, Benoit Peyronnet, Shahrokh Shariat, Benjamin Pradere, et al.

## ► To cite this version:

Ronan Thenault, Kevin Kaulanjan, Thomas Darde, Nathalie Rioux-Leclercq, Karim Bensalah, et al.. The Application of Artificial Intelligence in Prostate Cancer Management-What Improvements Can Be Expected? A Systematic Review. Applied Sciences, 2020, 10 (18), 10.3390/app10186428 . hal-02996051

**HAL Id: hal-02996051**

**<https://univ-rennes.hal.science/hal-02996051>**

Submitted on 9 Nov 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.




L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Review

# The Application of Artificial Intelligence in Prostate Cancer Management—What Improvements Can Be Expected? A Systematic Review

Ronan Thenault <sup>1,2,\*</sup> , Kevin Kaulanjan <sup>2</sup> , Thomas Darde <sup>3</sup>, Nathalie Rioux-Leclercq <sup>2</sup>, Karim Bensalah <sup>2</sup>, Marie Mermier <sup>2</sup>, Zine-eddine Khene <sup>2</sup>, Benoit Peyronnet <sup>2</sup>, Shahrokh Shariat <sup>4,5,6,7,8,9,10</sup> Benjamin Pradère <sup>4,11,\*</sup>  and Romain Mathieu <sup>1,2</sup>

<sup>1</sup> IRSET, 35000 Rennes, France; romain.mathieu@chu-rennes.fr

<sup>2</sup> Department of Urology, Service d'urologie, Rennes University Hospital, Hôpital Pontchaillou, 2, rue Henri Le Guilloux, 35000 Rennes, France; kevin.kaulanjan@chu-rennes.fr (K.K.); nathalie.rioux-leclercq@chu-rennes.fr (N.R.-L.); karim.bensalah@chu-rennes.fr (K.B.); marie.mermier@chu-rennes.fr (M.M.); zine-eddine.khene@chu-rennes.fr (Z.-e.K.); benoit.peyronnet@chu-rennes.fr (B.P.)

<sup>3</sup> SciLicum, 5 la hurbinais, 35850 Gévezé, France; thomas.darde@scilicum.com

<sup>4</sup> Department of Urology, Medical University Vienna, General Hospital, 1090 Vienna, Austria; sfshariat@gmail.com

<sup>5</sup> Department of Urology, University of Texas Southwestern Medical Center at Dallas, Dallas, TX 75390, USA

<sup>6</sup> Department of Urology, Weill Cornell Medical College, New York, NY 10065, USA

<sup>7</sup> Karl Landsteiner Institute, 1090 Vienna, Austria

<sup>8</sup> Division of Urology, Department of Special Surgery, Jordan University Hospital, The University of Jordan, 11942 Amman, Jordan

<sup>9</sup> Institute for Urology and Reproductive Health, Sechenov University, 119991 Moscow, Russia

<sup>10</sup> Department of Urology, Second Faculty of Medicine, Charles University, 15006 Prague, Czech Republic

<sup>11</sup> Department of Urology, CHU de Tours, 2, boulevard Tonnellé, 37000 Tours, France

\* Correspondence: ronan.thenault@chu-rennes.fr (R.T.); Benjamin.pradere@chu-tours.fr (B.P.); Tel.: +33-299-284-321 (R.T.)

Received: 15 August 2020; Accepted: 7 September 2020; Published: 15 September 2020



**Abstract:** Artificial Intelligence (AI) is progressively remodeling our daily life. A large amount of information from “big data” now enables machines to perform predictions and improve our healthcare system. AI has the potential to reshape prostate cancer (PCa) management thanks to growing applications in the field. The purpose of this review is to provide a global overview of AI in PCa for urologists, pathologists, radiotherapists, and oncologists to consider future changes in their daily practice. A systematic review was performed, based on PubMed MEDLINE, Google Scholar, and DBLP databases for original studies published in English from January 2009 to January 2019 relevant to PCa, AI, Machine Learning, Artificial Neural Networks, Convolutional Neural Networks, and Natural-Language Processing. Only articles with full text accessible were considered. A total of 1008 articles were reviewed, and 48 articles were included. AI has potential applications in all fields of PCa management: analysis of genetic predispositions, diagnosis in imaging, and pathology to detect PCa or to differentiate between significant and non-significant PCa. AI also applies to PCa treatment, whether surgical intervention or radiotherapy, skills training, or assessment, to improve treatment modalities and outcome prediction. AI in PCa management has the potential to provide a useful role by predicting PCa more accurately, using a multiomic approach and risk-stratifying patients to provide personalized medicine.

**Keywords:** artificial intelligence; machine learning; deep learning artificial neural network; natural-language processing; prostate cancer; computer-aided diagnosis; prediction performance; individualized medicine; healthcare improvement; outcomes; diagnosis; prognosis; treatment

---

## 1. Introduction

Artificial intelligence (AI) now plays a major place in our daily life, and has spread to every field. AI is a sub-field of cognitive sciences and uses computational neurobiology, mathematical logic, and computer science. AI may be used for problem-solving with high logic or algorithmic complexity by mimicking human cognitive function [1]. Using training data, AI systems can achieve defined purposes such as analyzing, automating, optimizing, or predicting. In healthcare, AI has led to improvements in different fields of cancer management. AI has already shown interesting potential in diagnosis, prognosis, treatment, and outcome prediction for a few years [2]. Over the past decade, the number of scientific studies based on AI has significantly increased. However, it may be difficult to fully apprehend the terminology [3]. AI application in healthcare mainly refers to the concept of machine learning (ML) [4,5].

The purpose of this review is to group the current knowledge about AI and prostate cancer (PCa) management, from diagnosis to outcome. Thus, we present a brief overview of ML in this field, explain the metrics used in ML and how to interpret them, and explain some of the technical jargon associated with the field so that readers with a medical background and a basic knowledge of statistics can feel more comfortable when examining ML applications.

## 2. Materials and Methods

### 2.1. Ethical Consideration

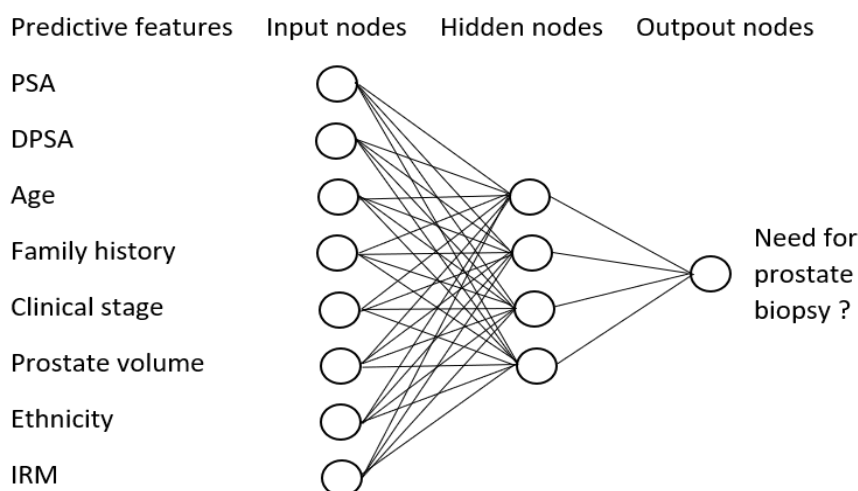
No ethics committee approval was required for this systematic review.

### 2.2. Study Selection

A detailed and systematic review of the literature was performed to identify publications that assessed the application of AI in PCa, using PubMed/MEDLINE, Google Scholar, IEEE Explore, and DBLP databases, which is a computer science database, for scientific journals from January 2009 to November 2019. We used the search terms prostatic neoplasm combined with one of the following terms: Artificial Intelligence, Artificial Neural Networks, Machine Learning, Natural-Language Processing, and computer vision.

Machine Learning is based on a mathematical and statistical approach to give computers the ability to “learn” from data and to identify intricate patterns. There are usually two phases in the process of ML: the first is called the “learning” phase and the second one the “verification” phase. The learning phase is performed to create a model to solve the requested task. According to the type of information that is provided during the learning phase, ML algorithms can be classified into “supervised”, “unsupervised”, and “reinforcement” learning. Random Forests (RF), Support Vector Machine (SVM) [6], or Naïve Bayes classification are examples of supervised ML models [7,8].

Artificial Neural Networks (ANNs) are one of the most important supervised ML approaches. ANNs is a system inspired schematically on the functioning of biological neurons [9]. As a biologic neural network, ANNs also contain layers of simple points of information, called nodes, which form an interconnected processing network. Commonly, to have a functional ANNs, a training, validation, and verification steps are necessary with non-overlapping sets of data. The most used ANN in medicine research is the multilayer perceptron, with 3 neuronal layers [9]. Between the first and the last layer, there is an interconnecting layer (hidden nodes), forward feeding neurons (Figure 1).

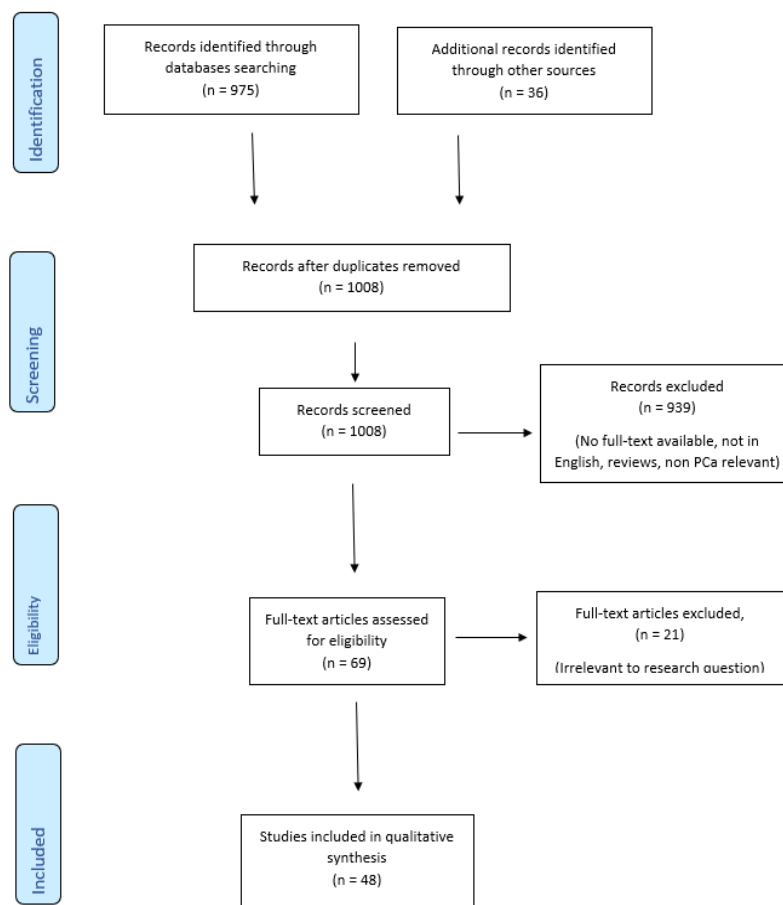


**Figure 1.** Artificial Neural Network model based on prostate cancer biopsies.

There is a subtype of ANN that showed its effectiveness for segmenting, classifying, and recognizing complex patterns (e.g., animals, objects, cells) on digitized images that is called Convolutional Neural Network (CNN) [10]. Identification of the prostate on MRI images [11] or classification of skin cancer from digital images are examples of CNNs abilities [12]. Over the past decade, a growing amount of research in the field of AI has been applied to prostate cancer (PCa). AI has been proposed to help clinicians in the management of PCa and it may be difficult to have a global view of the actual situation [13–15].

### 2.3. Search Strategy

The inclusion criteria for this systematic review were published full articles, written in English, concerning humans, and related to AI and PCa. Non-original research (review, meta-analysis, editorial, conference abstract, response, or opinion to the editor), *ex vivo* studies, and articles published before 2009 were excluded. Two authors (R.T. et K.K.) independently reviewed the titles and abstracts. A third review (by R.M.) and a full-text screening were required in case of different findings and discrepancies during this step. Additional informative publications were selected by cross-referencing the bibliography of previously selected articles. All the selected studies then underwent a full-text screening. Agreement was completed by consensus between the two reviewers. In accordance with PRISMA guidelines [16], Figure 2 summarizes our articles selection process. This study was registered in the International Protective Register of Systematic Reviews (PROSPERO, ID 158454). Information on first author, journal, publication year, AI method, size of the population, aim of the study, and prediction accuracy was collected.



**Figure 2.** Summary of study selection process.

### 3. Results

#### 3.1. Diagnosis

PCa diagnosis is based on the pathological evaluation of prostate biopsies. Historical factors that guide the decision to propose biopsies are blood levels of Prostate Specific Antigen (PSA), age of the patient, co-morbidities, and presentation of the prostate at digital rectal examination. The use of PSA remains controversial because of its lack of specificity. Therefore, new serum, urine, and tissue-based biomarkers have emerged but are not yet widespread in daily practice [17]. Family history of PCa and genetic predisposition also plays an important role in the decision to propose PCa early detection. Magnetic Resonance Imaging (MRI) is now part of the PCa diagnosis pathway and is performed before biopsies. However, despite the constant improvement of its abilities to detect PCa, MRI remains a dependent operator and a time-consuming assessment. The same applies to pathology at the time of biopsy interpretation. Therefore, AI may play a crucial role at several steps of the diagnostic pathway.

##### 3.1.1. Genomics

In an era of personalized medicine, genomics is one of the most promising paths. With “DEPTH”, a ML algorithm, MacInnis et al. [18] identified several PCa risk-associated regions through 541,129 germline Single Nucleotide Polymorphism (SNP) from 1854 PCa patients and 1894 controls. Only 14 of them were already known, and 112 novel putative susceptibility areas were proposed. Besides the identification of genomic regions of PCa susceptibility, AI also permitted the discovery of new genes involved in PCa. Hou et al. [19] collected public data of 466 PCa patients from the Gene Expression Omnibus and The Cancer Genome Atlas dataset and proposed a genetic algorithm with ANN to

establish a diagnostic and prognostic prediction model. The authors identified C1QTNF3 as a good predictor for PCa diagnosis (GSE6956: Area Under the Curve (AUC) = 0.791; GSE8218: AUC = 0.868; GSE26910: AUC = 0.972) and established a 15-gene signature. The prediction model based on this 15-gene signature also showed an AUC of 0.953 for PCa diagnosis and an AUC of 0.808 for prognosis.

### 3.1.2. Imaging and Radiomics

Radiomics is a method that extracts many features from medical imaging using data-characterization algorithms, with the potential to reveal disease characteristics that are impossible to appreciate by the naked-eye. Many studies assessed the potential benefits of AI in PCa imaging, especially with multi-parametric MRI. These studies are summarized in Table 1. The two main outcomes regarding imaging PCa diagnosis that addressed these studies are detection of malignant lesion and differentiation of significant and non-significant cancers.

From MRI images, ML software can be trained with labeled, anatomical, or morphological data. Ishioka et al. [20] demonstrated that CNN algorithms can automatically diagnose PCa by analyzing thousands of MR images in a matter of seconds. From the prostatic MR imaging, several authors [21–24] have shown that the use of SVM models allows a PCa diagnosis with accuracies that ranged from 0.85 to 0.91. With another AI approach, two studies [25,26] used CNN algorithms to bring complementary results which echo those from an SVM approach, with similar performance. Matulewicz et al. [27] showed that anatomical differentiation between prostatic areas (peripheral zone vs central zone) could significantly increase the classification capabilities of ANN models, highlighting the important role of anatomic segmentation. The authors used two ANN algorithms, one trained with information on zonal anatomy of the prostate and the other without to demonstrate that prostate anatomic segmentation improve the performance of PCa diagnosis significantly (96.8% vs. 94.9%;  $p = 0.03$ ).

MRI provides a large amount of quantitative data through textural, morphological analysis. These data are available for analysis after the definition of a region of interest, a segmentation, and an extraction. The most frequently used descriptors are the average and median of grey levels, variance, skewness, or kurtosis. Recent studies included radiomics based on the integration of quantitative and texture features analysis, to help PCa detection on imaging. The work of Zhao et al. [28] led to the identification of radiomic features significantly associated with PCa in PZ. Betrouni et al. [29] showed the incorporation of these radiomic features may be a useful tool in PCa diagnosis if the lesion was more than 0.5cc. Nevertheless, some studies did not report any benefit from ML model trained with radiomic parameters. Bonekamp et al. [30] found that radiomic machine learning had similar but not superior performance than mean Apparent Diffusion Coefficient (ADC) assessment in distinguishing malignant from benign prostate lesion. Considering all these different approaches, Viswanath et al. [31] empirically compared different machine classifiers. Within the framework of multicentric data from different MRI, with different acquisition protocols or resolution, boosted Quadratic Discriminant Analysis (QDA) was the more robust and efficient among the models tested for PCa detection, even if it was not the more accurate.

Several studies also reported the ability of AI to help radiologists distinguish significant lesions. Le et al. [32] used CNN architecture to reach this objective with 91.5% accuracy. Li et al. [33] trained SVM algorithms with MRI variables to classify NS (non-significant) PCa and CS (clinically significant) PCa into two groups in the central gland (CG). Wang et al. [34] demonstrated that adjunction of radiomic features in a ML algorithm could significantly improve (AUC = 98% vs. 88%) the performance of Prostate Imaging-Reporting and Data System (PI-RADS), a score related to clinically significant cancer [35]. In addition, a ML model trained with radiomic features was better than PI-RADS to distinguish transitional zone (TZ) from PCa tumor. Fehr et al. [36] also used a trained ML algorithm with radiomic features to differentiate among NS PCa and CS PCa. The performance of the created model was significantly enhanced by combining T2-W and ADC MRI-based texture features, whether the tumor was in the PZ or in the TZ. Similarly, Min et al. [37] found consistent results by combining T2-W, ADC, and DWI MRI-based texture features.

**Table 1.** Summary of the studies reporting on Artificial Intelligence (AI) in the imaging prostate cancer diagnosis.

Author	AI Model	Patients (Training/Validation/Testing)	Parameters	Predicted Outcomes
Ishioka et al. BJUI 2018 [20]	CNN algorithm (U-net and ResNet50)	335 PCa patients (301/34)	T2-weighted (T2-w) MR images labeled as ‘cancer’ or ‘no cancer’	Detection of PCa on MRI with CAD algorithm required 5.5 h to learn to analyze 2 million images, operating at approximately 30 ms/ image for evaluation and showed AUC of 0.645 and 0.636. 16/17 and 7/17 patients mistakenly diagnosed as having PCa in two validation datasets.
Vos et al. Phys Med Biol 2010 [21]	SVM	34 PCa patients	T2-w and dynamic contrast-enhanced (DCE) T1-w MR images labeled as malignant, benign, or normal regions	T2 values alone achieved a diagnostic AUC of 0.85 (0.77–0.92) with a significantly improved discriminating performance of 0.89 (0.81–0.95), when combined with DCE T1-w features.
Niaf et al. Phys. Med. Biol. 2012 [22]	SVM (MATLAB)	30 RARP	42 cancer ROIs, 49 suspicious ROIs, and 124 benign ROIs	For the discrimination of the malignant versus nonmalignant tissues AUC was 0.89 (0.81–0.94) and 0.82 (0.73–0.90) for the discrimination between malignant versus suspicious tissues.
Giannini et al. Comput Med Imaging Graph, 2015 [23]	Computer Vision and SVM	56 patients with 65 PCa lesions	MR images with anatomical and pharmacokinetic features	For the discrimination of the malignant versus nonmalignant tissues, AUC was 0.91 (1st–3rd quartile; 0.83–0.98), with a median sensitivity and specificity of 0.84 (1st–3rd quartile; 0.77–0.93) and 0.86 (1st–3rd quartile; 0.76–0.95) respectively. Detection of 62/65 lesions (per-lesion sensitivity increased to 98%, if CS PCa considered only) 2/3 false negative lesions were NS PCa.
Rampun et al. Int J Numer Method Biomed Eng 2016 [24]	Unsupervised ML algorithm	37 PCa patients	4 images features on T2-W MRI	CAD detection and localization of PCa within the peripheral zone with 86% of accuracy, 87% and 86% of sensitivity and specificity respectively.
Wang et al. Sci Rep 2017 [25]	DCNN vs. a non-DL with SIFT image feature and bag-of-word (BoW)	172 PCa patients	Clinicopathological and MR imaging Datasets with 2602 morphologic images	Statistically higher area under the receiver operating characteristics curve (AUC) for DCNN than non-DL ( $p = 0.0007 < 0.001$ ). The AUCs were respectively 0.84 (95% CI 0.78–0.89) for DL method and 0.70 (95% CI 0.63–0.77) for non-deep learning method.
Song et al. J MagnReson Imaging 2018 [26]	DCNN (Python scikit-learn and R software)	195 (159/17/19) patients with localized PCa	MR images labeled as ‘cancer’ or ‘no cancer’	Strong diagnostic performance for DCNN in distinguishing between PCa and non-cancerous tissues: AUC of 0.944 (95% CI: 0.876–0.994), 87.0% and 90.6% of sensitivity and specificity respectively, PPV of 87.0%, and NPV of 90.6%. Joining PI-RADS and DCNN provided additional net benefits rather than in particular.
Matulewicz et al. J MagnReson Imaging 2014 [27]	2 ANN models (STATISTICA)	18 PCa patients 5308 voxels (3716/796/796)	Model 1: 256 MRI variables Model 2: 256 MRI variables + 4 anatomical variables (percentage of PZ, TZ, U, and O).	By adding training with anatomic segmentation PCa, performance of model 2 was significantly higher than model 1 with respectively AUC 0.968 vs. 0.949 ( $p = 0.03$ ), sensitivity of 62.5% vs. 50% and specificity of 99.0% vs. 98.7%.
Zhao et al. Sci China Life Sci 2015 [28]	ANN (in-house software)	71 patients (35 with PCa and 36 without PCa)	12 radiomic features on T2-W	In the PZ, 10/12 radiomic features had significant difference ( $p < 0.01$ ) between PCa and non-PCa. In CG, 5/12 radiomic features (sum average, minimum value, standard deviation, 10th percentile, and entropy) had significant difference.

Table 1. Cont.

Author	AI Model	Patients (Training/Validation/Testing)	Parameters	Predicted Outcomes
Betrouni et al. Int J Comput Assist RadiolSurg. 2015 [29]	Unsupervised ML algorithm	15 RARP	Radiomic features extracted from MRI T2-w, T1-w DCE, and ADC images	The mean sensitivity and specificity were 65% (0–100%) and 81% (50–100%), respectively. The poorer scores were for lesions with volumes less than 0.5 cc (value actually use to define CS PCa on MR images). With only tumors with a significant size, the mean sensitivity and specificity grew to 70% and 88%, respectively.
Bonekamp et al. EurRadiol 2018 [30]	ML algorithm (Python scikit-learn)	316 men with MRI-TRUS fusion biopsy (183/133)	Lesion segmentation and 846 radiomic features analysis	Radiomic characterization of prostate lesions with MRI: ML vs. ADC comparison AUC for the mean ADC (AUC global = 0.84; AUC zone-specific $\leq 0.87$ ) vs. the ML (AUC global = 0.88, $p = 0.176$ ; AUC zone-specific $\leq 0.89$ , $p \geq 0.493$ ) showed no significantly different performance.
Viswanath et al. BMC Med Imaging 2019 [31]	Comparing 12 radiomic classifier derived from 4 ML families (MATLAB)	85 RARP	116 MRI features of 4 different types of texture	The boosted QDA classifier was identified as the most accurate, fast and robust for voxel-wise detection of PCa (AUCs of 0.735, 0.683, 0.768 across the 3 sites) but results therefore suggest that simpler classifiers may be more robust, accurate, and efficient for prostate cancer especially in multi-site validation.
Le et al. Phys Med Biol 2017 [32]	CNN algorithm (ResNet)	364 patients with a total of 463 PCa lesions and 450 identified non-cancerous image	MRI quantitative features	In distinguishing cancer from non-cancerous tissues: sensitivity of 89.85% and a specificity of 95.83% and 100% with 91.46% accuracy and a specificity of 76.92% and accuracy 88.75% for distinguishing NS PCa from CS PCa.
Li et al. Eur J Radiol 2018 [33]	SVM algorithms (MATLAB)	152 CG cancerous ROIs from 63 PCa patients	6 mpMRI significant variables (selected by feature-selection and variation test) from 55 variables	For classification between NS PCa vs. CS PCa in CG, the trained SMV algorithm in dataset A2 and validated in B2 had AUC of 0.91 (95% CI: 0.85, 0.95), with an accuracy of 84.1% and a sensitivity and specificity of 82.3% and 87.5% respectively. When the data sets were reversed, the trained SMV algorithm in dataset B2 and validated in A2 had AUC of 0.90 (95% CI: 0.85, 0.95), with an accuracy of 81.1% and a sensitivity and specificity of 90.5% and 79.7% respectively.
Wang et al. EurRadiol 2017 [34]	SVM algorithm	54 PCa patients	PI-RADS v2 and 8 radiomic features	For PCa vs. normal TZ, the ML algorithm trained with radiomic features had a significantly higher AUC (0.955 [95% CI 0.923–0.976]) than PI-RADS (0.878 [0.834–0.914], $p < 0.001$ ) but not vs. normal PZ (0.972 [0.945–0.988] vs. 0.940 [0.905–0.965], $p = 0.097$ ). Performance of PI-RADS was significantly improved for PCa vs. PZ AUC 0.983 [0.960–0.995]) and PCa vs. TZ (0.968 [0.940–0.985]) with radiomic features addition.



Table 1. Cont.

Author	AI Model	Patients (Training/Validation/Testing)	Parameters	Predicted Outcomes
Fehr et al. Proc. Natl. Acad. Sci. USA 2015 [36]	ML algorithm (in-house software in MATLAB)	147 RARP	Haralick features on T2-w and ADC	By combining T2-W and ADC, ML model accuracy was 93% for lesions developing in both PZ and TZ and 92% in the PZ alone for stratification between NS and CS PCa. Accuracy of 92% for distinction between GS 7(3+4) and 7(4 + 3) for cancers occurring in both the PZ and TZ and 93% in the PZ alone. To compare, a model using only the ADC mean achieved an accuracy of 58% and 63% for distinguishing between NS and CS PCa for cancers occurring in PZ and TZ and in PZ alone respectively. The same model performed an accuracy of 59% for stratification between GS 7(3 + 4) from 7(4 + 3) appearing in the PZ and TZ and 60% in PZ alone.
Min et al. Eur J Radiol 2019 [37]	ML algorithm (SMOTE)	280 (187/93) PCa patients	9 radiomic features among 918 extracted on T2-w, ADC, and DWI	Significant differences in the radiomic features existed between the CS PCa and NS PCa groups in the training and test cohorts ( $p < 0.01$ for both). For radiomic signature in the training cohort, the AUC was 0.872 and the sensitivity and specificity were 0.883 and 0.753, respectively. In the testing cohort, the AUC was 0.823, and 0.841 and 0.727 for the sensibility and specificity, respectively.
Azizi et al. IEEE 2018 [38]	ANN (Python, Keras, or Tensorflow)	255 (84/171) prostate biopsy cores of 157 patients	Temporal Enhanced Ultrasound image analysis 172 cores labeled as benign and 83 labeled as cancerous	Temporal Enhanced US using ANN improve significantly cancer detection and achieve an AUC = 0.96, with 76% of sensitivity, 98% of specificity and an accuracy of 93%.
Koizumi et al. Ann Nucl Med 2017 [39]	ANN algorithm (BONENAVI)	226 PCa patients (124 with skeletal metastasis and 101 without)	Bone Scintigraphy (BS) images	BONENAVI showed for metastasis detection 82% and 83% of sensitivity and specificity, respectively. AUC was 0.888 (95% CI 0.843–0.932). The mean ANN values were 0.78 (SD = 0.29) for the skeletal metastasis patients and 0.22 (SD = 0.30) for the non-skeletal metastasis patients. False negative had often a solitary lesion in the pelvis.
Acar et al. Br J Radiol, 2019 [40]	3 ML algorithms (MATLAB)	75 PCa metastatic patients	35 radiomic features on 68Ga-PSMA PET/CT images	The Weighted KNN (ML algorithm) succeeded to differentiate sclerotic lesion from metastasis or completely responded lesions with 0.76 AUC, sensitivity and specificity were 73.5% and 73.7%, respectively. PPV and NPV were 86.9% and 53.8%, respectively. GLZLM_SZHGE and histogram-based kurtosis were found to be the most important parameters, but 28/35 radiomic features were significant in differentiating metastatic from completely responded sclerotic lesions.

AI artificial intelligence, CNN convolutional neural network, DCNN deep convolutional neural network, PCa prostate cancer, AUC area under the curve, mpMRI multi-parametric magnetic resonance imaging, CAD computer-aided diagnosis, SVM support vector machine, DCE dynamic contrast-enhanced, RARP robot-assisted radical prostatectomy, ROI region of interest, CS PCa clinically significant PCa, NS PCa non-significant PCa, DCNN deep convolutional neural network, PI-RADS prostate imaging reporting and data system, PZ peripheral zone, TZ transition zone, U periurethral region, O outside prostate, CG central gland, ANN artificial neural network, ML machine learning, DL deep learning, ADC apparent diffusion coefficient, TRUS transrectal ultrasound, GS Gleason score, DWI diffusion weighted images, QDA quadratic discriminant analysis, PET-PSMA positron emission tomography-prostate specific membrane antigen ligand, KNN K-nearest neighbors.

Moreover, AI demonstrated promising results with other imaging used for PCa management. Azizi et al. demonstrated that ANN can be used with Temporal Enhanced US for PCa detection (AUC = 96%) [38]. Regarding detection of PCa metastases, Koizumi et al. reported the ability of an ANN algorithm, to identify skeletal metastases with an accuracy of 80% on bone scintigraphy. However, regarding the detection of non-skeletal metastasis, the performance of the model decreased to 22%, [39]. After treatment, metastatic lesions stand sclerotic and it can be difficult to identify recurrence. Acar et al. [40] trained three ML algorithms with radiomic features based on PMSA PET and CT scanner images to help clinicians distinguish active lesions among sclerotic lesions after treatment. This study identified important textural parameters to allow such detection reaching an accuracy of 76%.

### 3.1.3. Biopsies

In 2011, Lawrentschuk et al. [41] compared the performance of ANN, trained with clinical and pathological data from 3025 patients, and Logistic Regression (LR) to predict Trans Rectal Ultrasound (TRUS) biopsy outcomes. No difference was observed between both models (55 vs. 57%, respectively), to discriminate histological groups (LR-ANN respectively): benign outcomes were correctly identified (86–88%) and CS PCa was well classified in 65–66%. However, more recently, an ANN outperformed a model based on LR to predict PCa at biopsies [42]. Using this model, clinicians may avoid biopsies for one in two patients. However, the ANN missed 16% of all PCa of which 6% were clinically significant.

ANN has also wide applications in Natural-Language Processing (NLP), to extract data from reports. Kim et al. [43] have shown how this technological breakthrough could be useful with pathological reports. In their study, using 100 RARP pathology reports, extraction of pathological or staging information was feasible and reached an accuracy of nearly 100% for all characteristics: histologic subtype (99.0%), perineural invasion (98.9%), TNM stage (98.0%), surgical margin (97.0%), and dominant tumor size (95.7%). The execution time for NLP was under 1 s, while manual reviewers took more than 3 min per report.

## 3.2. Pathology

Several studies assessed the ability of AI to help pathologists in PCa identification and evaluation, with a gain in time and reduction of inter-observer variability (Table 2). Nir et al. [44] showed that a ML algorithm reported good accuracy in diagnosing (90.5%) or classifying (79.2%) PCa compared to pathologists. Working speed of the machine allowed the analysis of hundreds of slides in a few hours. With different AI models, [45–47] found identical results with a good concordance between pathologist and a CNN model, with high accuracy, sensibility, and specificity in diagnosing and differentiating low and high-grade PCa. Arvaniti et al. [48] provided additional information. Once again, kappa score between CNN and pathologists was identical to that between pathologists. However, interestingly, when there was a disagreement with pathologists, the stratification made by the CNN model improved prediction regarding disease-specific survival outcomes. Kim et al. [49] compared the pre-operative prediction of advanced PCa by using and comparing an ANN algorithm and a SVM model. Both have good results in predicting the pre-operative probability for > pT3a PCa, but SVM slightly outperformed ANN in this study. Partin tables are used to estimate the distribution of the risk of pathological stages. The Partin tables use clinical features, such as Gleason score, PSA, and clinical stage, to predict whether the tumor will be confined to the organ. Tsao et al. [50] compared the performance of these nomogram tables to an ANN model. Their ANN algorithm, STATISTICA, trained with total serum PSA, TNM stage, and biopsy Gleason score, had a significantly higher efficiency than Partin tables. Two clinical variables were significantly associated with a non-organ-confined disease: PSA and BMI. Wang et al. [51] optimized Partin tables by combining conventional parameters with those from a SMV model. This model significantly outperformed conventional Partin tables for organ-confined PCa (AUC = 0.891 vs. 0.730).

**Table 2.** Summary of the characteristics of studies reporting on Artificial Intelligence (AI) for prostate cancer pathological diagnosis.

Author	AI Model	Patients (Training/Validation/Testing)	Parameters	Predicted Outcomes
Nir et al. Med Image Anal 2018 [44]	Supervised ML algorithm (U-Net)	287 RARP (231/56) with 563 (333/230) TMA cores	23 weighted histological features on TMA cores	<p>The inter-annotator agreement between ML model and pathologist was <math>k = 0.51</math>, while the overall agreements between the 6 pathologists was <math>0.45</math> to <math>0.62</math> (<math>0.56 \pm 0.07</math>).</p> <p>The classifier's accuracy in detecting cancer was 90.5%. For the distinction between NS and SC PCa, the accuracy was 79.2%, with 79.2% and 79.1% of sensitivity and specificity, respectively.</p> <p>The manual grading by pathologists leads to inter-observer mean accuracy of 97.2%, with 98.4% and 86.2% of sensitivity and specificity respectively for cancer detection, and 78.8%, 79.2% and 79.3%, respectively, for NS vs. SC PCa. ML's working time was 18.5h for the training and validation steps.</p>
Gorelick et al. IEEE Trans Med Imaging 2013 [45]	ML algorithm (MATLAB 7.12.0)	15 RARP with 991 sub-images extracted from digital pathology images of 50 whole-mount tissue sections	7 morphometric features, 100 geometric features, and 9 tissue component labels	Difference between cancer versus non-cancer and high-grade versus low-grade was made with accuracies of 90% and 85% respectively with 12% and 5% of FN for cancer detection and classification, respectively.
Nguyen et al. J BiomedOpt 2017 [46]	Unsupervised ML algorithm (RF on MATLAB)	368 TMA cores with PCa consensus diagnosis	Combining morphological features with quantitative information from TMA cores	<p>To separate GS 3 to GS 4, the AUC value was 0.82, which is in the range of human error when inter-observer variability is considered.</p> <p>The AUC reduced from 0.98 to 0.87, inversely correlated with GS (6 to 10).</p>
Lucas et al. Virchows Arch 2019 [47]	CNN (Inception-v3, MATLAB)	96 TRUS biopsies from 38 patients	Digitized slides annotated with 4 groups using histological and simplify Gleason Pattern (GP)	<p>Differentiation between benign and PCa (<math>GP \geq 3</math>) areas resulted in an accuracy of 92%, with 90 and 93% of sensitivity and specificity, respectively.</p> <p>The differentiation between NS PCa and CS PCa was accurate for 90%, with 77 and 94% of sensitivity and specificity, respectively. Concordance between CNN and genito-urinary pathologist was achieved in 65% (<math>\kappa = 0.70</math>).</p>
Arvaniti et al. Sci Rep 2018 [48]	CNN (Python3, Keras, or Tensorflow)	886 PCa patients (641/245)	GP annotation on TMA cores	<p>The inter-annotator agreements between CNN and pathologists were 0.75 and 0.71, which was comparable with the inter-pathologist agreement (<math>k = 0.71</math>) for automated GS grading.</p> <p>On disease-specific survival, the stratification made by the CNN algorithm into prognostically distinct groups was more significant than the one achieved by pathologists, achieving a pathologist-level survival stratification.</p>
Kim et al. Korean J Radiol 2011 [49]	ANN vs. SVM	532 PCa patients (300/232)	3 clinical and 5 biopsy variables	<p>SMV significantly outperformed ANN with an AUC of 0.805 and 0.719, respectively (<math>p = 0.020</math>).</p> <p>Pre-operative probability for <math>&gt; pT3a</math>: for SMV, the accuracy was 77%, and the sensitivity and specificity were respectively 67% and 79%. For ANN, the sensitivity, specificity, and accuracy were 63%, 81%, and 78%, respectively.</p>

Table 2. Cont.

Author	AI Model	Patients (Training/Validation/Testing)	Parameters	Predicted Outcomes
Tsao et al. J Chin Med Assoc 2014 [50]	ANN (STATISTICA, StatSoft Inc) vs. LR (Partin Tables)	299 ORP or RARP	7 clinical and pathological variables	To predict the pathological stage (T2 or T3) after surgery, ANN was significantly better than LR ( $0.795 \pm 0.023$ versus $0.746 \pm 0.025$ , $p = 0.016$ ), with a sensitivity and specificity of 83% and 56% respectively for ANN and 70% and 56% respectively for LR. PSA and BMI were significantly associated with T3 stages.
Wang et al. J MagnReson Imaging 2018 [51]	SVM vs. LR (Partin Tables) (R software v.3.3.4)	541 PCa patients	Clinical, pathological, and MRI variables	The created model PartinMR (Partin tables + mpMRI staging + SMV) outperformed other models of this study, with AUC of 0.891 (95% CI: 0.884–0.899, $p < 0.001$ ), 79.3% and 75.7% of sensibility and specificity respectively, 79% PPV, and 76.0% NPV for OC PCa. Using MR staging, mPartin table (AUC, 0.814, 95% CI: 0.779–0.846, $p = 0.001$ ) is significantly better than the Partin table (AUC, 0.730, 95% CI: 0.690–0.767). MR stage was the most influential factor of extracapsular extension (HR, 2.77, 95% CI: 1.54–3.33), then D-max (2.01, 95% CI: 1.31–2.68), biopsy GS (1.64, 95% CI: 1.35–2.12), and PI-RADS score (1.21, 95% CI: 1.01–1.98).

PCa prostate cancer, ML machine learning, RARP robot-assisted prostatectomy, TMA tissue micro array, NS PCa non-significant PCa, CS PCa clinically significant PCa, GS Gleason score, CNN convolutional neural network, TRUS transrectal ultrasound, ANN artificial neural network, SVM support vector machine, LR logistic regression, ORP open radical prostatectomy, PSA prostate specific antigene, BMI body mass index, AUC area under the curve, NPV negative predictive value, PI-RADS prostate imaging reporting and data system.

### 3.3. Treatment

Similar to its development in PCa diagnosis, AI may also play an important role in the treatment pathway (Table 3). After PCa diagnosis, AI can help patients and physicians in treatment decisions by predicting oncological outcomes, complications, or biochemical recurrence. It also supports the practitioner in the improvement and optimization of his practices.

#### 3.3.1. Decision Making

Auffenberg et al. [52] used a ML algorithm, trained with clinical and pathological variables such as age, medical history, PSA, Gleason scores, or number of positive cores, to guide patients in their choice of treatment and to help them understand their disease. This application also allowed urologists to guide their patients towards treatment or further examination by calculating the probability of organ-confined disease with a good accuracy (81%).

#### 3.3.2. Surgery

In the field of surgery, ML algorithms can be used to improve surgical capacities. Ukimura et al. [53] used Virtual Reality 3D surgical navigation to perform robot-assisted radical prostatectomy (RARP) among 10 patients. This technology allowed the surgeon to identify anatomic structures while performing surgery. In three works, Hung et al. [54–56] developed an objective method to assess surgical performance, and used parameters from this method as training data for ML algorithms. In their first work [54], the authors created “dVLogger” to record automated performance metrics (APM) for expert and novice urological surgeons. Experts significantly outperformed novices for all criteria of bimanual dexterity and most of the movements efficiency criteria. The authors illustrated this higher efficiency with a 3D instrument-path tracing that demonstrates lesser instrument-path length with higher velocity on dominant and non-dominant instrument in the group of experts. Based on these APM data, a ML model was built to predict clinical outcomes with over 85% accuracy, such as length of hospital stay (LoS) and Foley catheter duration [55] or urinary continence [56], after RARP. Ranasinghe et al. [57] used DeepSurv, a NLP algorithm, to compare RARP and open radical prostatectomy (ORP) outcomes from individualized patient-reported information. Based on this model, the authors reported more negative emotions until 3 months after ORP due to discomfort and pain than patients who underwent RARP. The anxiety of PSA failure and sexual side effects were also higher at 9 months with ORP.

#### 3.3.3. Radiotherapy

As in surgery, radiotherapists used AI algorithms to optimize their treatment. Thus, Nicolae et al. [58] demonstrated that compared to brachytherapists, ML helped to reduce radiotherapy planning time from a mean time of 18 min to shorter than 1 min with no significant difference regarding irradiated prostate volume. Kajikawa et al. [59] built two CNN algorithms for optimization radiotherapy planning. With a CNN model trained with anatomical structure label dataset, the algorithm led to more contrasting results, with 70% accuracy for predicting the dosimetric eligibility of patients treated with intensity-modulated radiation therapy (IMRT).

Finally, some authors aimed at predicting patient susceptibility to side effects after radiotherapy by associating AI with genomics. Based on genome-wide association studies (GWAS), Lee et al. [60] and Oh et al. [61] tried to predict lower urinary tract symptoms (LUTS) and rectal bleeding or erectile dysfunction after radiation therapy, respectively. Both ML algorithms reached 70% accuracy. Pella et al. [62] compared SVM and ANN outcomes prediction after conformal radiotherapy. Both demonstrated similar prediction performance of 70% for gastro-intestinal and genito-urinary toxicity after radiotherapy.

**Table 3.** Summary of the characteristics of studies reporting on Artificial Intelligence (AI) in the prostate cancer treatment and outcomes.

	Treatment Modality	AI Model	Patients (Training/Validation/Testing)	Parameters	Predicted Outcomes
Surgery	Auffenberg et al. EurUrol 2019 [52]	ML algorithm (askMUSIC)	7543 PCa patients	Clinical and pathological variables	3413 (45%) have chosen RP, 2289 (30%) active surveillance, 1280 (17%) RT, 422 (5.6%) ADT, and 139 (1.8%) watchful waiting. Personalized prediction with ML algorithm to inform patients of PCa treatments was highly accurate (AUC = 0.81). Age, number of positive cores, and GS were the most influential parameters influencing patient treatment choice.
Surgery	Ukimura et al. J. Endourol.2014 [53]	Virtual Reality 3D surgical navigation (TilePro system)	10 RARP	5 anatomic structures (prostate, image-visible biopsy-proven “index” cancer lesion, neurovascular bundles, urethra, and recorded biopsy trajectories)	The 3D model allowed careful surgical dissection in the proximity of the biopsy-proven index lesion. Negative surgical margins were achieved in 90%. No intraoperative complications and no additional time for surgical procedures. There was a good correlation between geographic location of the index lesion on the final pathology report with the created 3D model. At 3 months, PSA level were undetectable (<0.03 ng/mL) in 100%. TRUS underestimated index lesion volume compared to MR-based modeling. The addition of multi-parametric MR and digital data of biopsy-proven improved to 90% of actual pathologic volume.
Surgery	Hung et al. J. Urol. 2018 [54]	Surgeon performance metrics recorder (dVLogger)	20 urological surgeons (10 experts and 10 novices)	Automated performance metrics (APM)	Mean number cases for novices was 35 (5–80) and 810 (100–2000) for experts. Experts were more efficient in their moves by completing operative steps faster ( $p < 0.001$ ) with lesser instrument-travel distance ( $p < 0.01$ ), lesser aggregate instrument idle time ( $p < 0.001$ ), shorter camera-path length ( $p < 0.001$ ) and more frequent camera movements ( $p < 0.03$ ). Experts had greater ratio of dominant: non-dominant instrument-path distance for all steps ( $p < 0.04$ ), except for anterior vesicourethral anastomosis.
Surgery	Hung et al. J Endourol 2018 [55]	3 ML algorithms (Python scikit-learn 0.19.0 package)	78 RARP	Hospital length of stay (LoS) and APM	ML algorithm predicted length of stay with 87.2% accuracy. Trained with APM, its accuracy was improved to 88.5%. Metrics related with camera use were main predictors for surgery time, LoS, and foley duration. Patient outcomes predicted by the algorithm had significant association with ‘ground truth’ in surgery time, length of stay, and Foley duration.
Surgery	Hung et al. BJUI 2019 [56]	ML algorithm (DeepSurv)	100 RARP cases	APM and patient clinicopathological and continence data	Urinary continence was achieved in 79% after a median of 126 days. ML model performed C-index of 0.6 and mean absolute error of 85.9 in predicting continence. 3 APM top-ranked features were taken during vesico-urethral anastomosis and 1 during prostatic apical dissection. APM were ranked higher by the ML model than clinicopathological features for prediction of urinary continence recovery after RARP.

Table 3. Cont.

	Treatment Modality	AI Model	Patients (Training/Validation/Testing)	Parameters	Predicted Outcomes
Surgery	Ranasinghe et al. Urol. Oncol. 2018 [57]	NLP algorithm (PRIME-2)	5157 RARP and 579 ORP	Pre- and post-operative clinical data	Surgeon experience and erectile function preservation ( $p < 0.01$ ) were important factors in treatment choice. There were no significant differences in urinary, sexual, or bowel symptoms between RARP and ORP during the 12-month follow-up period. Emotions expressed by patients who underwent RARP were more positive while ORP expressed more negative emotions immediately and 3 months post-surgery ( $p < 0.05$ ), due to pain and discomfort, and during 9 months due to fear and anxiety of pending PSA tests and sexual side effects.
Radiotherapy	Nicolae et al. Int. J. Radiat. Oncol. Biol. Phys. 2017 [58]	ML algorithm	100 PCa patients treated by Low Dose Rate (LDR) brachytherapy	6 key measures of clinical quality of LDR brachytherapy	The average planning time for the ML algorithm was significantly shorter with $0.84 \pm 0.57$ min compared to $17.88 \pm 8.76$ min for the expert planner ( $p = 0.020$ ). The average prostate V150% was 4% higher for brachytherapists than ML, but not clinically significant. Only 2 brachytherapists were able to distinguish ML's work from an expert.
Radiotherapy	Kajikawa et al. Radiol Phys Technol 2018 [59]	2 CNN algorithm (AlexNet)	60 PCa patients treated by IMRT	Planning CT images and structure labels	For the ANN model with structure labels, the accuracy was 70%, with sensitivity and specificity of 94.6% and 31% respectively. The ANN with planning CT had an accuracy of 56.7%, with 70% and 11.3% of sensitivity and specificity, respectively. These models had moderate performance to predict dosimetric eligibility.
Radiotherapy	Lee et al. Int J Radiat Oncol Biol Phys. 2018 [60]	ML (PRFR model)	324 PCa patients treated with RT were genotyped for 606,563 germline Single Nucleotide Polymorphism (SNP)	14 clinical variables and genome-wide association studies (GWAS) data	The predictive accuracy of the ML model differed across the urinary symptoms. Only for the weak stream endpoint did it achieve a significant AUC of 0.70 (95% CI 0.54–0.86; $p = 0.01$ ). 7 interconnected proteins were highlighted by gene ontology analysis and were already known to be associated with LUTS.
Radiotherapy	Oh et al. Sci Rep 2017 [61]	ML (PRFR model)	368 PCa patients treated with RT were genotyped for 606,571 germline SNP	9 clinical variables and GWAS data	For rectal bleeding and erectile dysfunction, AUC were 0.70 and 0.69, respectively. With chi-square test $p$ -values of 0.95 and 0.93 for rectal bleeding and erectile dysfunction, respectively, there was a good correlation between the real incidence and the predicted one. 10 biological processes/genes were highly associated with both complications and the estimated false discovery rate (FDR) was less than 0.05.
Radiotherapy	Pella et al. Med Phys 2011 [62]	SVM and ANN (Matlab)	321 PCa patients treated by prostate conformal RT	13 clinical and RT variables	ANN and SVM-based models showed similar prediction accuracy for gastro-intestinal and genito-urinary toxicity after RT. For SVM model, the AUC index was 0.717, with 84.6% and 58.8% of sensitivity and specificity respectively, and 70% of accuracy. For ANN, the AUC index was 0.697 and 69% of accuracy.

Table 3. Cont.

	Treatment Modality	AI Model	Patients (Training/Validation/Testing)	Parameters	Predicted Outcomes
Medication	Heintzelman et al. J Am Med Inform Assoc 2013 [63]	NLP (ClinREAD, Lockheed Martin, Bethesda, Maryland, USA))	33 metastatic PCa patients and 4409 clinical encounters	4 pain severity groups with 42 pain severity contextual rules, 16 semantic types	NLP identified 6387 pain and 13 827 drug mentions in medical records. For all patient died from metastatic PCa, except for 2, the level of pain increased drastically in the last 2 years of life. Severe pain was associated with opioids prescription (OR = 6.6, $p < 0.0001$ ) and palliative radiation (OR = 3.4, $p = 0.0002$ ). 5 factors were significantly associated with severe pain: receipt of chemotherapy, opioids, or palliative radiotherapy, being in the last year of life and the number of medical appointments. The 5 African American patients clustered at the high end of the pain index spectrum, but non-significant.
Surgery (outcomes)	Zhang et al. Oncotarget. 2016 [64]	SVM vs. LR analysis	205 PCa patients treated by surgery	Clinicopathologic and MR imaging datasets	To predict the probability of PCa BCR, SVM had significantly higher AUC (0.959 vs. 0.886; $p = 0.007$ ), sensitivity (93.3% vs. 83.3%; $p = 0.025$ ), specificity (91.7% vs. 77.2%; $p = 0.009$ ) and accuracy (92.2% vs. 79.0%; $p = 0.006$ ) than LR analysis. By adding MRI-derived variables, D'AMICO classification performance was effectively improved (AUC: 0.970 vs. 0.859, $p < 0.001$ ; sensitivity: 91.7% vs. 86.7%, $p = 0.031$ ; specificity: 94.5% vs. 78.6%, $p = 0.001$ ; and accuracy: 93.7% vs. 81.0%, $p = 0.007$ ). ADC (HR = 0.149, $p = 0.035$ ) was the only imaging predictor of time until PSA failure. Others were GS (HR = 1.560, $p = 0.008$ ), surgical-T3b (HR = 4.525, $p < 0.001$ ) and positive surgical margin (HR = 1.314, $p = 0.007$ ).
Surgery (outcomes)	Wong et al. BJUI 2019 [65]	3 supervised ML algorithms vs. LR analysis	338 RARP	19 different variables (demographic, clinical, imaging, and operative data)	To predict patient with biochemical recurrence at 1 year, the 3 models were K-nearest neighbor, random forest tree, and logistic regression with an accuracy prediction scores of 0.976, 0.953 and 0.976, respectively. All 3 ML models were better than conventional statistical regression model AUC 0.865, vs. 0.903, 0.924 and 0.940, respectively, to predict early biochemical recurrence after RARP.
Radiotherapy (outcomes)	Abdollahi et al. Radiol Med 2019 [66]	ML algorithm (in-house Python codes)	33 PCa patients treated by intense-modulated radiation therapy (IMRT)	Radiomic features	For GS prediction, T2-w radiomic models, with a mean AUC of 0.739 had better efficiency. To predict stages, ADC models were more effective with an AUC of 0.675. For treatment response after IMRT, 22 radiomic features were strongly correlated, with a wide range achievement from 0.55 to 0.78.

AI artificial intelligence, ML machine learning, PCa prostate cancer, ADT androgen deprivation therapy, AUC area under the curve, RARP robot-assisted radical prostatectomy, SVM support vector machine, LR logistic regression, RP radical prostatectomy, RT radiotherapy, APM automated performance metrics, LoS length of stay, NLP natural-language processing, LDR low dose rate, IMRT intensity-modulated radiotherapy, SNP single nucleotide polymorphism, GWAS genome-wide association studies, LUTS lower urinary tract symptoms, BCR biochemical recurrence, GS Gleason score, ANN artificial neural network.



### 3.3.4. Medication

Support by the NLP method, Heintzelman et al. [63] extracted clinical data related to pain and related drug prescription from medical reports of metastatic PCa patients. A correlation was observed between a significant increase in pain in the last two years of life, and a rising number of medical appointments, rising opioid prescription, and palliative radiotherapy.

### 3.4. Oncological Outcomes

Zhang et al. [64] compared the ability of an SVM model and LR analysis to predict PCa biochemical recurrence (BCR). Trained with clinicopathological data, SVM had a significantly higher accuracy. The model increases in performance when implemented with MRI-derived variables. However, the only significant imaging predictor of BCR was ADC. Expectedly, positive surgical margin, surgical non-OCD PCa, and GS were clinicopathological predictors of time to failure. More recently, Wong et al. [65], using training three ML models, also pointed out that ML techniques could outperform traditional Cox regression analysis. AUC of the ML models could reach more than 0.95, to predict early BCR after RARP.

Radiotherapists also used an AI algorithm to predict treatment response or outcomes. In 2019, Abdollahi et al. [66] built a ML model trained with several radiomic features extracted from MRI data before and after intensity-modulated radiotherapy (IMRT). The model could predict the early treatment response with credible performance, up to 78%, after IMRT based on pre-treatment MRI.

## 4. Discussion

AI is on the rise and seems to hold a predominant place in our daily lives. There is presently a competitive battle for innovation, involving countries and private companies in this field. We hereby confirm the extensive work regarding the development of AI in PCa. We demonstrated AI now applies to almost all aspects of PCa management.

We are now entering an era of personalized healthcare. The “one treatment fits for all” is no longer considered to be appropriate. This is largely due to the progress in genomics. Genomics can predict genetic predispositions to a pathology, such as diabetes or cancer [67,68]. The characterization of the individual genome is becoming more accessible, but also generates a massive amount of data. Therefore, ML is a resource perfectly adapted for the use of genomics [69,70]. AI has already helped discoveries in molecular [71] and genetic medicine. With an unsupervised ML algorithm, Theofilatos et al. had promising results in the prediction of over 5000 protein complexes and gene ontology function [72]. Moreover, there are new computational methodologies under development to detect DNA variants, as SNP, as predictive factors of diseases [73]. In PCa, the combination of these two technological advances have already led to the identification of new genetic susceptibility regions and gene signatures [18,19,60,61] hitherto unknown via conventional logistic regression.

In the past few years, the number of prostate biopsy sections has increased significantly [74,75]. With 12 to 16 biopsies, each procedure results in several slides. Each slide must be analyzed, even if it does not contain cancer. Automated diagnosis for prostate cancer could allow the exclusion of these normal sections to save time for pathologists [76]. Some CNN models have already outperformed pathologists in the Barrett’s esophagus inter-observer agreement [77] or in the detection of breast cancer metastasis in the lymph nodes with a time constraint [78]. In terms of time constraint, the contributions from Kim [43] and Heintzelman [63] have shown that the deployment of AI can enable the instantaneous tasks completion that are burdensome for humans with nearly perfect accuracy.

Radiomics appears to be another promising field of study in the scope of personalized healthcare. Radiomics can help through the entire continuum of healthcare, from the diagnosis to the prediction of therapy response. Based on image acquisition and segmentation from different modalities (MRI, CT, PET), radiomics allows the extraction of a great amount of radiomic imaging features. These features correlated with patient information, and may be clinically relevant to define cancer aggressiveness

and prognosis but also treatment outcomes [79–82]. New fields of research are now emerging from these developments such as radiogenomics, which study the correlation between gene expression and imaging or radiomic features [79]. With its large amount of numerical data, radiomics also benefits substantially from coupling with ML algorithms. Radiomics models and ML approaches now represent non-invasive methods to equip physicians to provide personalized healthcare through prognostic information. In radiation oncology, Lambin [83] already highlighted the importance of tailoring the radiotherapy on intrinsic tumor properties for treatment decisions and personalized healthcare [84]. However, the contributions of radiomics still need to be assessed prospectively and with a standardization of methods.

AI also provides us with insightful support in the treatment field. Robotic surgery is now widespread. To improve the time and safety of surgery, augmented reality seems to be promising. In RARP, Ukimura [53] and Porpiglia [85,86], have shown that the use of this feature was safely feasible. By 3D reconstructing a virtual model based on two-dimension imaging, anatomical structures and tumor location are highlighted to guide surgeons during the operation and could improve surgical and oncologic outcomes. This new technology is also usable in open surgery [87], in simulation and training [88]. In addition to urological surgery [89], it is also used in hepatobiliary [90–92], otolaryngology [93,94] or orthopedic surgery [95]. Hung [54–56] has developed an interesting tool combining robotics and AI. By recording the surgeon's movements while operating, AI provide automated performance metrics and determine global movement features [96–98]. ML now allows the extraction and exploitation of surgical skills. With these data, new opportunities are opening. It is now feasible to evaluate surgical skills, train them efficiently, establish as yet unknown correlations with post-operative or oncologic outcomes, and set a new standardization of practices [99].

These new discoveries are achieved by the ability to recognize complex patterns through nonlinear relationships, with high accuracy, even when data are missing or contaminated [100]. These methods raise some concerns and bring up issues. First, to be legitimate or acceptable, or even to be rejected as irrelevant, a ML algorithm's decisions must be understandable and therefore explainable. The number of micro-thought processes carried out by the system is beyond our understanding capacities. As previously stated, there is a hidden layer of neurons in ANN software. Therefore, calculations that result in outcomes are not visible for the programmer who created ANNs, responsible for opacity [3,101]. It is the black-box effect. The data goes in the black box and the prediction comes out, but without any accompanying justification [102]. It is the difference between prediction and explanation. Failure to understand the algorithm decision can lead to misinterpretation. One other concern is related to the training sample used to propose ML algorithm. To be functional, a ML algorithm needs a training step, using usually a training sample from the study population. However, this sample, which is often limited in size, may not represent the general population. In a systematic review of skin cancer classification using CNN, Brinker [16] highlights the lack of dark-skinned people in the different training samples to be representative and be applied on world population. For PCa, it is known that ethnic origin is source of some predisposition. PCa does not have the same characteristics and severity in an Afro-American or Asian population [103]. Yet, most of the previously reviewed studies are based on Asian populations [20,26,28,37,42,50,59] and are probably not applicable to an African American population. Heintzelman [63] emphasizes this difference by showing that Afro-American patients in their study were clustered at the upper end of the pain index spectrum, even if these results were not significant.

As already mentioned, a ML algorithm performance is related to the training step, cohort parameters, and optimization processes. Each model is therefore different from another, created and trained with another set of data. In addition, programmers and researchers often keep to themselves their code source code and data. There is an emergence of many AI models, which differs in the functioning, with complicated reproducibility to establish. Therefore, there is a critical need to propose external validation at these algorithms before any use in daily practice. In addition to this lack of external validation, AI algorithms are insufficiently compared to current references and comparing ML

software with human experts is not a common practice [104], even in PCa. In this review, only 6 studies have been compared to current reference prediction methods, such as Partin tables [105,106], or expert opinion [44,47,48]. Therefore, the comparison between AI and human experts is still underdeveloped in PCa while in other fields, AI has achieved a confirmed expert level such as in classification of skin cancer [107], retinal diseases [108,109], or colorectal polype classification [110,111].

A systematic review of the literature [112] and a methodology review [113] have been performed to assess the benefits of ML and ANN on logistic regression, respectively. They both found no evidence of superior performance of ML and ANN over logistic regression for clinical prediction modeling. However, they highlighted the necessity of improvements in methodology, such as model validation or corrections for imbalanced outcomes. However, they did not investigate which factors influencing the difference in performance, such as the sample size. In this article, some of the reviewed studies have small sample size. Even though the sample size affects the performance of the ML model created [114], it appears possible to design a robust ML algorithm with a limited sample size [115] by using “synthetic dataset”. Partin tables use clinical parameters of PCa to predict whether the tumor will be organ-confined [105,106]. In our review, NLP software [43,63] have outperformed the working and data-mining abilities of humans. Systematically, ML models had results comparable at expert levels. The overall agreements between pathologists or radiotherapists was equivalent to the AI. In [48], there was a slight difference in the groups made by the CNN model compared to pathologist, with a better disease-specific prediction. This further underlines the fact that an AI algorithm can reveal complex patterns.

This black-box effect is also responsible for a lack of reproducibility [70,116,117]. These reproducibility issues are also related to the lack of sharing of assets and numerous modular components. There is an emergence of many AI models, which differs in the functioning, with complicated reproducibility to establish. However, in all studies reviewed, a major part of them cited their AI software or the method they used.

Finally, AI models are also used in other cancers. In skin cancer classification, CNNs already display a high-performance skin lesion classification [16]. For breast cancer, interest in the application of AI models to breast MRI is growing worldwide [118]. Such applications are still in development and are not used in clinical practice, but computer-aided diagnosis could be useful for breast cancer [119,120]. As for PCa, AI has been used in many cancers with mixed results such as lung cancer [121], neuro-oncology [122,123], or esophagus cancer [124].

#### *Review Limitation*

There are limitations to our review. First, we decided to select only published studies from 2009. We considered that articles had to be recent to be relevant, since AI has undergone a major expansion over the past decade. Secondly, since most of us are urologists, and not computer scientists, we focused on the use of AI models and clinical results obtained, rather than detailing techniques or methodological parts of these same models. As we developed previously, this review highlights the retrospective and heterogeneous nature of these studies. Each study has a different ML algorithm. Larger validation datasets, with different populations and environments, are needed to achieve external validation and promise a universal tool usable by everyone. Until now, the clinical relevance of these algorithms has not been deeply explored from a clinical and a statistical point of view, leading to a slowdown of the applicability of these results.

### **5. Conclusions and Perspectives**

Worldwide, 1.4 million new cases of PCa were diagnosed in 2016 and nearly 400,000 people died, which is twice the rate of 1990 [125]. Predictive, preventive, and personalized medicine [126] has the potential to provide a useful role by predicting PCa more accurately, using a multiomic approach [127] and risk-stratifying patients to provide personalized medicine. In PCa, AI can already help through the entire continuum of healthcare, from the detection of predictive factors to the identification and

prediction of risks after treatment. AI models can be superior to conventional statistical techniques while emphasizing complex patterns and reliable prediction. To use these new models, external validation is critical and still poor at that time. To continue to expand and improve, AI models will require a large amount of well-annotated and representative data from various sources and many forms. The challenge for the future is therefore to collect and share data, develop and combine different approaches, and, above all, propose reliable validations of any model. Only in this way will AI promptly enter in the daily management of PCa.

**Author Contributions:** All authors researched data for the article, made substantial contributions to discussions of content and reviewed and edited the manuscript before submission. R.T. wrote the manuscript. The project was developed by R.T. and R.M. Literature review by R.T. and K.K., R.T., K.K., T.D., N.R.-L., K.B., M.M., Z.-e.K., B.P. (Benoit Peyronnet), S.S., B.P. (Benjamin Pradère), R.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Ethical Approval:** As this is a review of the literature, no ethical approval was required.

## References

1. Kaplan, A.M.; Haenlein, M. Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Bus. Horiz.* **2019**, *62*, 15–25. [CrossRef]
2. Tran, B.X.; Vu, G.T.; Ha, G.H.; Vuong, Q.H.; Ho, T.M.; Vuong, T.T.; La, V.P.; Nghiem, K.C.P.; Le, H.T.; Latkin, C.A.; et al. Global evolution of research in artificial intelligence in health and medicine: A bibliometric study. *J. Clin. Med.* **2019**, *8*, 360. [CrossRef] [PubMed]
3. Appenzeller, T. *The AI Revolution in Science*; Science AAAS: Washington, DC, USA, 2017. Available online: <https://www.sciencemag.org/news/2017/07/ai-revolution-science> (accessed on 9 January 2020).
4. Ghahramani, Z. Probabilistic machine learning and artificial intelligence. *Nature* **2015**, *521*, 452–459. [CrossRef] [PubMed]
5. Jordan, M.I.; Mitchell, T.M. Machine learning: Trends, perspectives, and prospects. *Science* **2015**, *349*, 255–260. [CrossRef] [PubMed]
6. Noble, W.S. What is a support vector machine? *Nat. Biotechnol.* **2006**, *24*, 1565–1567. [CrossRef] [PubMed]
7. Bishop, C.M. *Pattern Recognition and Machine Learning*; CERN Document Server: New York, NY, USA, 2006. Available online: <https://cds.cern.ch/record/998831> (accessed on 30 January 2020).
8. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
9. Anagnostou, T.; Remzi, M.; Lykourinas, M.; Djavan, B. Artificial neural networks for decision-making in urologic oncology. *Eur. Urol.* **2003**, *43*, 596–603. [CrossRef]
10. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]
11. Karimi, D.; Samei, G.; Kesch, C.; Nir, G.; Salcudean, T. Prostate segmentation in MRI using a convolutional neural network architecture and training strategy based on statistical shape models. *Int. J. Comput. Assist. Radiol. Surg.* **2018**, *13*, 1211–1219. [CrossRef] [PubMed]
12. Brinker, T.J.; Hekler, A.; Utikal, J.; Grabe, N.; Schadendorf, D.; Klode, J.; Berking, C.; Steeb, T.; Enk, A.H.; Von Kalle, C.; et al. Skin cancer classification using convolutional neural networks: Systematic review. *J. Med. Internet Res.* **2018**, *20*, e11936. [CrossRef] [PubMed]
13. Abbod, M.F.; Catto, J.; Linkens, D.A.; Hamdy, F.C. Application of artificial intelligence to the management of urological cancer. *J. Urol.* **2007**, *178*, 1150–1156. [CrossRef]
14. Goldenberg, S.L.; Nir, G.; Salcudean, S.E. A new era: Artificial intelligence and machine learning in prostate cancer. *Nat. Rev. Urol.* **2019**, *16*, 391–403. [CrossRef] [PubMed]
15. 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 J. Urol.* **2019**. [CrossRef] [PubMed]
16. Moher, D.; Shamseer, L.; Clarke, M.J.; Ghersi, D.; Liberati, A.; Petticrew, M.P.; Shekelle, P.G.; Stewart, L. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst. Rev.* **2015**, *4*, 1. [CrossRef] [PubMed]

17. Brönimann, S.; Pradere, B.; Karakiewicz, P.; Abufaraj, M.; Briganti, A.; Shariat, S.F. An overview of current and emerging diagnostic, staging and prognostic markers for prostate cancer. *Expert Rev. Mol. Diagn.* **2020**, 1–10. [CrossRef]
18. MacInnis, R.J.; Schmidt, D.F.; Makalic, E.; Severi, G.; Fitzgerald, L.M.; Reumann, M.; Kapuscinski, M.K.; Kowalczyk, A.; Zhou, Z.; Goudey, B.; et al. Use of a novel nonparametric version of DEPTH to identify genomic regions associated with prostate cancer risk. *Cancer Epidemiol. Biomark. Prev.* **2016**, *25*, 1619–1624. [CrossRef]
19. Hou, Q.; Bing, Z.T.; Hu, C.; Li, M.Y.; Yang, K.H.; Mo, Z.; Xie, X.W.; Liao, J.L.; Lu, Y.; Horie, S.; et al. RankProd combined with genetic algorithm optimized artificial neural network establishes a diagnostic and prognostic prediction model that revealed C1QTNF3 as a biomarker for prostate cancer. *EBioMedicine* **2018**, *32*, 234–244. Available online: <https://www.ncbi.nlm.nih.gov/pubmed/29861410> (accessed on 5 December 2019). [CrossRef]
20. Ishioka, J.; Matsuoka, Y.; Uehara, S.; Yasuda, Y.; Kijima, T.; Yoshida, S.; Yokoyama, M.; Saito, K.; Kihara, K.; Numao, N.; et al. Computer-aided diagnosis of prostate cancer on magnetic resonance imaging using a convolutional neural network algorithm. *BJU Int.* **2018**, *122*, 411–417. [CrossRef]
21. Vos, P.C.; Hambrock, T.; Barenstz, J.O.; Huisman, H.J. Computer-assisted analysis of peripheral zone prostate lesions using T2-weighted and dynamic contrast enhanced T1-weighted MRI. *Phys. Med. Biol.* **2010**, *55*, 1719–1734. [CrossRef]
22. Niaf, E.; Rouvière, O.; Mège-Lechevallier, F.; Bratan, F.; Lartizien, C. Computer-aided diagnosis of prostate cancer in the peripheral zone using multiparametric MRI. *Phys. Med. Biol.* **2012**, *57*, 3833–3851. [CrossRef]
23. Giannini, V.; Mazzetti, S.; Vignati, A.; Russo, F.; Bollito, E.; Porpiglia, F.; Stasi, M.; Regge, D. A fully automatic computer aided diagnosis system for peripheral zone prostate cancer detection using multi-parametric magnetic resonance imaging. *Comput. Med. Imaging Graph.* **2015**, *46*, 219–226. [CrossRef] [PubMed]
24. Rampun, A.; Chen, Z.; Malcolm, P.; Tiddeman, B.P.; Zwiggelaar, R. Computer-aided diagnosis: Detection and localization of prostate cancer within the peripheral zone. *Int. J. Numer. Method Biomed. Eng.* **2015**, *32*, e02745. [CrossRef]
25. Wang, X.; Yang, W.; Weinreb, J.; Han, J.; Li, Q.; Kong, X.; Yan, Y.; Ke, Z.; Luo, B.; Liu, T.; et al. Searching for prostate cancer by fully automated magnetic resonance imaging classification: Deep learning versus non-deep learning. *Sci. Rep.* **2017**, *7*, 15415. [CrossRef]
26. Song, Y.; Yan, X.; Liu, H.; Zhou, M.; Hu, B.; Yang, G.; Zhang, Y.-D. Computer-aided diagnosis of prostate cancer using a deep convolutional neural network from multiparametric MRI. *J. Magn. Reson. Imaging* **2018**, *48*, 1570–1577. [CrossRef]
27. Matulewicz, L.; Jansen, J.; Bokacheva, L.; Vargas, H.A.; Akin, O.; Fine, S.W.; Shukla-Dave, A.; Eastham, J.A.; Hricak, H.; Koutcher, J.A.; et al. Anatomic segmentation improves prostate cancer detection with artificial neural networks analysis of 1H magnetic resonance spectroscopic imaging. *J. Magn. Reson. Imaging* **2013**, *40*, 1414–1421. [CrossRef]
28. Zhao, K.; Wang, C.; Hu, J.; Yang, X.; Wang, H.; Li, F.Y.; Zhang, X.; Zhang, J.; Wang, X. Prostate cancer identification: Quantitative analysis of T2-weighted MR images based on a back propagation artificial neural network model. *Sci. Chin. Life Sci.* **2015**, *58*, 666–673. [CrossRef]
29. Betrouni, N.; Makni, N.; Lakroum, S.; Mordon, S.; Villers, A.; Puech, P. Computer-aided analysis of prostate multiparametric MR images: An unsupervised fusion-based approach. *Int. J. Comput. Assist. Radiol. Surg.* **2015**, *10*, 1515–1526. [CrossRef]
30. Bonekamp, D.; Kohl, S.; Wiesenfarth, M.; Schelb, P.; Radtke, J.P.; Götz, M.; Kickingereider, P.; Yaqubi, K.; Hitthaler, B.; Gählert, N.; et al. Radiomic machine learning for characterization of prostate lesions with MRI: Comparison to ADC values. *Radiology* **2018**, *289*, 128–137. [CrossRef]
31. Viswanath, S.; Chirra, P.V.; Yim, M.C.; Rofsky, N.M.; Purysko, A.; Rosen, M.; Bloch, B.N.; Madabhushi, A. Comparing radiomic classifiers and classifier ensembles for detection of peripheral zone prostate tumors on T2-weighted MRI: A multi-site study. *BMC Med. Imaging* **2019**, *19*, 22–28. [CrossRef]
32. Le, M.H.; Chen, J.; Wang, L.; Wang, Z.; Liu, W.; Cheng, K.T.; Yang, X. Automated diagnosis of prostate cancer in multi-parametric MRI based on multimodal convolutional neural networks. *Phys. Med. Biol.* **2017**, *62*, 6497–6514. [CrossRef]
33. Li, J.; Weng, Z.; Xu, H.; Zhang, Z.; Miao, H.; Chen, W.; Liu, Z.; Zhang, X.; Wang, M.; Xu, X.; et al. Support Vector Machines (SVM) classification of prostate cancer Gleason score in central gland using multiparametric magnetic resonance images: A cross-validated study. *Eur. J. Radiol.* **2018**, *98*, 61–67. [CrossRef] [PubMed]

34. Wang, J.; Wu, C.J.; Bao, M.L.; Zhang, J.; Wang, X.N.; Zhang, Y.D. 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**, *27*, 4082–4090. [[CrossRef](#)] [[PubMed](#)]
35. Weinreb, J.; Barentsz, J.O.; Choyke, P.L.; Cornud, F.; Haider, M.A.; Macura, K.J.; Margolis, D.; Schnall, M.D.; Shtern, F.; Tempany, C.M.; et al. PI-RADS prostate imaging—Reporting and data system: 2015, Version 2. *Eur. Urol.* **2015**, *69*, 16–40. [[CrossRef](#)]
36. Fehr, D.; Veeraraghavan, H.; Wibmer, A.; Gondo, T.; Matsumoto, K.; Vargas, H.A.; Sala, E.; Hricak, H.; Deasy, J. Automatic classification of prostate cancer Gleason scores from multiparametric magnetic resonance images. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E6265–E6273. [[CrossRef](#)]
37. Min, X.; Li, M.; Dong, D.; Feng, Z.; Zhang, P.; Ke, Z.; You, H.; Han, F.; Ma, H.; Tian, J.; et al. Multi-parametric MRI-based radiomics signature for discriminating between clinically significant and insignificant prostate cancer: Cross-validation of a machine learning method. *Eur. J. Radiol.* **2019**, *115*, 16–21. [[CrossRef](#)]
38. Azizi, S.; Bayat, S.; Yan, P.; Tahmasebi, A.; Kwak, J.T.; Xu, S.; Turkbey, B.; Choyke, P.; Pinto, P.; Wood, B.; et al. Deep recurrent neural networks for prostate cancer detection: Analysis of temporal enhanced ultrasound. *IEEE Trans. Med. Imaging* **2018**, *37*, 2695–2703. [[CrossRef](#)]
39. Koizumi, M.; Motegi, K.; Koyama, M.; Terauchi, T.; Yuasa, T.; Yonese, J. Diagnostic performance of a computer-assisted diagnosis system for bone scintigraphy of newly developed skeletal metastasis in prostate cancer patients: Search for low-sensitivity subgroups. *Ann. Nucl. Med.* **2017**, *31*, 521–528. [[CrossRef](#)]
40. Acar, E.; Leblebici, A.; Ellidokuz, B.E.; Başbınar, Y.; Kaya, G. Çapa Machine learning for differentiating metastatic and completely responded sclerotic bone lesion in prostate cancer: A retrospective radiomics study. *Br. J. Radiol.* **2019**, *92*, 20190286. [[CrossRef](#)]
41. Lawrentshuk, N.; Lockwood, G.; Davies, P.; Evans, A.; Sweet, J.; Toi, A.; Fleshner, N.E. Predicting prostate biopsy outcome: Artificial neural networks and polychotomous regression are equivalent models. *Int. Urol. Nephrol.* **2010**, *43*, 23–30. [[CrossRef](#)] [[PubMed](#)]
42. Takeuchi, T.; Hattori-Kato, M.; Okuno, Y.; Iwai, S.; Mikami, K. Prediction of prostate cancer by deep learning with multilayer artificial neural network. *Can. Urol. Assoc. J.* **2018**, *13*, E145–E150. [[CrossRef](#)]
43. Kim, B.J.; Merchant, M.; Zheng, C.; Thomas, A.A.; Contreras, R.; Jacobsen, S.J.; Chien, G.W. Second prize: A natural language processing program effectively extracts key pathologic findings from radical prostatectomy reports. *J. Endourol.* **2014**, *28*, 1474–1478. [[CrossRef](#)] [[PubMed](#)]
44. Nir, G.; Hor, S.; Karimi, D.; Fazli, L.; Skinnider, B.F.; Tavassoli, P.; Turbin, D.; Villamil, C.F.; Wang, G.; Wilson, R.S.; et al. Automatic grading of prostate cancer in digitized histopathology images: Learning from multiple experts. *Med. Image Anal.* **2018**, *50*, 167–180. [[CrossRef](#)] [[PubMed](#)]
45. Gorelick, L.; Veksler, O.; Gaed, M.; Gomez, J.A.; Moussa, M.; Bauman, G.S.; Fenster, A.; Ward, A.D. Prostate histopathology: Learning tissue component histograms for cancer detection and classification. *IEEE Trans. Med. Imaging* **2013**, *32*, 1804–1818. [[CrossRef](#)] [[PubMed](#)]
46. Nguyen, T.H.; Sridharan, S.; Macias, V.; Kajdacsy-Balla, A.; Melamed, J.; Do, M.N.; Popescu, G. Automatic Gleason grading of prostate cancer using quantitative phase imaging and machine learning. *J. Biomed. Opt.* **2017**, *22*, 36015. [[CrossRef](#)]
47. Lucas, M.; Jansen, I.; Savci-Heijink, C.D.; Meijer, S.L.; De Boer, O.J.; Van Leeuwen, T.G.; De Bruin, D.M.; Marquering, H.A. Deep learning for automatic Gleason pattern classification for grade group determination of prostate biopsies. *Virchows Archiv.* **2019**, *475*, 77–83. [[CrossRef](#)]
48. Arvaniti, E.; Fricker, K.S.; Moret, M.; Rupp, N.; Hermanns, T.; Fankhauser, C.D.; Wey, N.; Wild, P.J.; Rueschoff, J.H.; Claassen, M. Automated Gleason grading of prostate cancer tissue microarrays via deep learning. *Sci. Rep.* **2018**, *8*, 12054. [[CrossRef](#)]
49. Kim, S.Y.; Moon, S.K.; Jung, D.C.; Hwang, S.I.; Sung, C.K.; Cho, J.Y.; Kim, S.H.; Lee, J.; Lee, H.J. Pre-operative prediction of advanced prostatic cancer using clinical decision support systems: Accuracy comparison between support vector machine and artificial neural network. *Korean J. Radiol.* **2011**, *12*, 588–594. [[CrossRef](#)]
50. Tsao, C.W.; Liu, C.Y.; Cha, T.L.; Wu, S.T.; Sun, G.H.; Yu, D.S.; Chen, H.I.; Chang, S.Y.; Chen, S.C.; Hsu, M.H. Artificial neural network for predicting pathological stage of clinically localized prostate cancer in a Taiwanese population. *J. Chin. Med. Assoc.* **2014**, *77*, 513–518. [[CrossRef](#)]
51. Wang, J.; Wu, C.J.; Bao, M.L.; Zhang, J.; Shi, H.B.; Zhang, Y.D. Using support vector machine analysis to assess PartinMR: A new prediction model for organ-confined prostate cancer. *J. Magn. Reson. Imaging* **2018**, *48*, 499–506. [[CrossRef](#)]

52. Auffenberg, G.B.; Ghani, K.R.; Ramani, S.; Usoro, E.; Denton, B.; Rogers, C.; Stockton, B.; Miller, D.C.; Singh, K. askMUSIC: Leveraging a clinical registry to develop a new machine learning model to inform patients of prostate cancer treatments chosen by similar men. *Eur. Urol.* **2019**, *75*, 901–907. [[CrossRef](#)]
53. Ukimura, O.; Aroni, M.; Nakamoto, M.; Shoji, S.; Abreu, A.L.; Matsugasumi, T.; Berger, A.; Desai, M.; Gill, I.S. Three-dimensional surgical navigation model with TilePro display during robot-assisted radical prostatectomy. *J. Endourol.* **2014**, *28*, 625–630. [[CrossRef](#)] [[PubMed](#)]
54. Hung, A.J.; Chen, J.; Jarc, A.; Hatcher, D.; Djaladat, H.; Gill, I.S. Development and validation of objective performance metrics for robot-assisted radical prostatectomy: A pilot study. *J. Urol.* **2018**, *199*, 296–304. [[CrossRef](#)] [[PubMed](#)]
55. Hung, A.J.; Chen, J.; Che, Z.; Nilanon, T.; Jarc, A.; Titus, M.; Oh, P.; Gill, I.S.; Liu, Y. Utilizing machine learning and automated performance metrics to evaluate robot-assisted radical prostatectomy performance and predict outcomes. *J. Endourol.* **2018**, *32*, 438–444. [[CrossRef](#)] [[PubMed](#)]
56. Hung, A.J.; Chen, J.; Ghodoussipour, S.; Oh, P.J.; Liu, Z.; Nguyen, J.; Purushotham, S.; Gill, I.S.; Liu, Y. A deep-learning model using automated performance metrics and clinical features to predict urinary continence recovery after robot-assisted radical prostatectomy. *BJU Int.* **2019**, *124*, 487–495. [[CrossRef](#)]
57. Ranasinghe, W.K.B.; De Silva, D.; Bandaragoda, T.; Adikari, A.; Alahakoon, D.; Persad, R.; Lawrentshuk, N.; Bolton, D. Robotic-assisted vs. open radical prostatectomy: A machine learning framework for intelligent analysis of patient-reported outcomes from online cancer support groups. *Urol. Oncol. Semin. Orig. Investig.* **2018**, *36*, 529.e1–529.e9. [[CrossRef](#)]
58. Nicolae, A.; Morton, G.; Chung, H.; Loblaw, A.; Jain, S.; Mitchell, D.; Lu, L.; Helou, J.; Al-Hanaqta, M.; Heath, E.; et al. Evaluation of a machine-learning algorithm for treatment planning in prostate low-dose-rate brachytherapy. *Int. J. Radiat. Oncol.* **2017**, *97*, 822–829. [[CrossRef](#)]
59. Kajikawa, T.; Kadoya, N.; Ito, K.; Takayama, Y.; Chiba, T.; Tomori, S.; Takeda, K.; Jingu, K. Automated prediction of dosimetric eligibility of patients with prostate cancer undergoing intensity-modulated radiation therapy using a convolutional neural network. *Radiol. Phys. Technol.* **2018**, *11*, 320–327. [[CrossRef](#)]
60. Lee, S.; Kerns, S.; Ostrer, H.; Rosenstein, B.; Deasy, J.; Oh, J.H. Machine learning on a genome-wide association study to predict late genitourinary toxicity after prostate radiation therapy. *Int. J. Radiat. Oncol.* **2018**, *101*, 128–135. [[CrossRef](#)]
61. Oh, J.H.; Kerns, S.; Ostrer, H.; Powell, S.N.; Rosenstein, B.; Deasy, J. Computational methods using genome-wide association studies to predict radiotherapy complications and to identify correlative molecular processes. *Sci. Rep.* **2017**, *7*, 43381. [[CrossRef](#)]
62. Pella, A.; Cambria, R.; Riboldi, M.; Jerezek-Fossa, B.A.; Fodor, C.; Zerini, D.; Torshabi, A.E.; Cattani, F.; Garibaldi, C.; Pedroli, G.; et al. Use of machine learning methods for prediction of acute toxicity in organs at risk following prostate radiotherapy. *Med. Phys.* **2011**, *38*, 2859–2867. [[CrossRef](#)]
63. Heintzelman, N.H.; Taylor, R.J.; Simonsen, L.; Lustig, R.; Anderko, D.; Haythornthwaite, J.A.; Childs, L.C.; Bova, G.S. Longitudinal analysis of pain in patients with metastatic prostate cancer using natural language processing of medical record text. *J. Am. Med. Inform. Assoc.* **2013**, *20*, 898–905. [[CrossRef](#)] [[PubMed](#)]
64. Zhang, Y.D.; Wang, J.; Wu, C.J.; Bao, M.L.; Li, H.; Wang, X.N.; Tao, J.; Shi, H.B. An imaging-based approach predicts clinical outcomes in prostate cancer through a novel support vector machine classification. *Oncotarget* **2016**, *7*, 78140–78151. [[CrossRef](#)] [[PubMed](#)]
65. Wong, N.C.; Lam, C.J.; Patterson, L.; Shayegan, B. Use of machine learning to predict early biochemical recurrence after robot-assisted prostatectomy. *BJU Int.* **2018**, *123*, 51–57. [[CrossRef](#)] [[PubMed](#)]
66. Abdollahi, H.; Mofid, B.; Shiri, I.; Razzaghdoust, A.; Saadipoor, A.; Mahdavi, A.; Galandooz, H.M.; Mahdavi, S.R.; Moid, B. Machine learning-based radiomic models to predict intensity-modulated radiation therapy response, Gleason score and stage in prostate cancer. *Radiol. Med.* **2019**, *124*, 555–567. [[CrossRef](#)]
67. Golub, T.; Slonim, D.K.; Tamayo, P.; Huard, C.; Gaasenbeek, M.; Mesirov, J.P.; Coller, H.; Loh, M.L.; Downing, J.R.; Caligiuri, M.A.; et al. Molecular classification of cancer: Class discovery and class prediction by gene expression monitoring. *Science* **1999**, *286*, 531–537. [[CrossRef](#)]
68. Fogel, A.L.; Kvedar, J.C. Artificial intelligence powers digital medicine. *NPJ Digit. Med.* **2018**, *1*, 1–4. [[CrossRef](#)]
69. Libbrecht, M.W.; Noble, W.S. Machine learning applications in genetics and genomics. *Nat. Rev. Genet.* **2015**, *16*, 321–332. [[CrossRef](#)]

70. Xu, J.; Yang, P.; Xue, S.; Sharma, B.; Sanchez-Martin, M.; Wang, F.; Beaty, K.A.; Dehan, E.; Parikh, B. Translating cancer genomics into precision medicine with artificial intelligence: Applications, challenges and future perspectives. *Qual. Life Res.* **2019**, *138*, 109–124. [[CrossRef](#)]
71. Golkov, V.; Skwark, M.J.; Golkov, A.; Dosovitskiy, A.; Brox, T.; Meiler, J.; Cremers, D. Protein contact prediction from amino acid co-evolution using convolutional networks for graph-valued images. In *Advances in Neural Information Processing Systems, Proceedings of the 30th Conference on Neural Information Processing Systems (NIPS 2016)*; Lee, D.D., Sugiyama, M., Luxburg, U.V., Guyon, I., Garnett, R., Eds.; Curran Associates Inc.: Barcelona, Spain, 2016; pp. 4222–4230.
72. Theofilatos, K.; Pavlopoulou, N.; Papasavvas, C.A.; Likiothanassis, S.; Dimitrakopoulos, C.; Georgopoulos, E.; Moschopoulos, C.; Mavroudi, S. Predicting protein complexes from weighted protein–protein interaction graphs with a novel unsupervised methodology: Evolutionary enhanced Markov clustering. *Artif. Intell. Med.* **2015**, *63*, 181–189. [[CrossRef](#)]
73. Rapakoulia, T.; Theofilatos, K.; Klefogiannis, D.; Likiothanassis, S.; Tsakalidis, A.; Mavroudi, S. EnsembleGASVR: A novel ensemble method for classifying missense single nucleotide polymorphisms. *Bioinformatics* **2014**, *30*, 2324–2333. [[CrossRef](#)]
74. Fairweather, M.; Raut, C.P. To biopsy, or not to biopsy: Is there really a question? *Ann. Surg. Oncol.* **2019**, *26*, 4182–4184. [[CrossRef](#)] [[PubMed](#)]
75. Heidenreich, A.; Bellmunt, J.; Bolla, M.; Joniau, S.; Mason, M.; Matveev, V.; Mottet, N.; Schmid, H.-P.; Van Der Kwast, T.; Wiegel, T.; et al. EAU guidelines on prostate cancer. Part 1: Screening, diagnosis, and treatment of clinically localised disease. *Eur. Urol.* **2011**, *59*, 61–71. [[CrossRef](#)] [[PubMed](#)]
76. Doyle, S.; Hwang, M.; Shah, K.; Madabhushi, A.; Feldman, M.; Tomaszewski, J. Automated grading of prostate cancer using architectural and textural image features. In *Proceedings of the 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, Arlington, VA, USA, 12–15 April 2007; pp. 1284–1287. [[CrossRef](#)]
77. Vennalaganti, P.R.; Kanakadandi, V.N.; Gross, S.A.; Parasa, S.; Wang, K.K.; Gupta, N.; Sharma, P. Inter-observer agreement among pathologists using wide-area transepithelial sampling with computer-assisted analysis in patients with Barrett’s esophagus. *Am. J. Gastroenterol.* **2015**, *110*, 1257–1260. [[CrossRef](#)]
78. Bejnordi, B.E.; Veta, M.; Van Diest, P.J.; Van Ginneken, B.; Karssemeijer, N.; Litjens, G.; Van Der Laak, J.A.W.M. The CAMELYON16 Consortium Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *JAMA* **2017**, *318*, 2199–2210. [[CrossRef](#)] [[PubMed](#)]
79. Kumar, V.; Gu, Y.; Basu, S.; Berglund, A.; Eschrich, S.A.; Schabath, M.B.; Forster, K.; Aerts, H.J.; Dekker, A.; Fenstermacher, D.; et al. Radiomics: The process and the challenges. *Magn. Reson. Imaging* **2012**, *30*, 1234–1248. [[CrossRef](#)] [[PubMed](#)]
80. Peeken, J.C.; Bernhofer, M.; Wiestler, B.; Goldberg, T.; Cremers, D.; Rost, B.; Wilkens, J.J.; Combs, S.E.; Nüsslin, F. Radiomics in radiooncology—Challenging the medical physicist. *Phys. Med.* **2018**, *48*, 27–36. [[CrossRef](#)]
81. Gillies, R.J.; Kinahan, P.E.; Hricak, H. Radiomics: Images are more than pictures, they are data. *Radiology* **2016**, *278*, 563–577. [[CrossRef](#)]
82. Stoyanova, R.; Takhar, M.; Tschudi, Y.; Ford, J.C.; Solórzano, G.; Erho, N.; Balagurunathan, Y.; Punnen, S.; Davicioni, E.; Gillies, R.J.; et al. Prostate cancer radiomics and the promise of radiogenomics. *Transl. Cancer Res.* **2016**, *5*, 432–447. [[CrossRef](#)]
83. Lambin, P.; Van Stiphout, R.G.P.M.; Starmans, M.H.W.; Rios-Velazquez, E.; Nalbantov, G.; Aerts, H.J.W.L.; Roelofs, E.; Van Elmpt, W.; Boutros, P.C.; Granone, P.; et al. Predicting outcomes in radiation oncology—Multifactorial decision support systems. *Nat. Rev. Clin. Oncol.* **2012**, *10*, 27–40. [[CrossRef](#)]
84. Fave, X.; Zhang, L.; Yang, J.; Mackin, D.; Balter, P.; Gomez, D.; Followill, D.; Jones, A.K.; Stingo, F.C.; Liao, Z.; et al. Delta-radiomics features for the prediction of patient outcomes in non-small cell lung cancer. *Sci. Rep.* **2017**, *7*, 1–11. [[CrossRef](#)]
85. Porpiglia, F.; Fiori, C.; Checcucci, E.; Amparore, D.; Bertolo, R. Augmented reality robot-assisted radical prostatectomy: Preliminary experience. *Urology* **2018**, *115*, 184. [[CrossRef](#)] [[PubMed](#)]
86. Porpiglia, F.; Checcucci, E.; Amparore, D.; Autorino, R.; Piana, A.; Bellin, A.; Piazzolla, P.; Massa, F.; Bollito, E.; Gned, D.; et al. Augmented-reality robot-assisted radical prostatectomy using hyper-accuracy three-dimensional reconstruction (HA3D™) technology: A radiological and pathological study. *BJU Int.* **2018**, *123*, 834–845. [[CrossRef](#)] [[PubMed](#)]



87. Fida, B.; Cutolo, F.; Di Franco, G.; Ferrari, M.; Ferrari, V. Augmented reality in open surgery. *Updat. Surg.* **2018**, *70*, 389–400. [[CrossRef](#)] [[PubMed](#)]
88. Feußner, H.; Ostler, D.; Wilhelm, D. Robotics and augmented reality: Current state of development and future perspectives. *Chirurg* **2018**, *89*, 760–768. [[CrossRef](#)]
89. Wake, N.; Bjurlin, M.A.; Rostami, P.; Chandarana, H.; Huang, W.C. Three-dimensional printing and augmented reality: Enhanced precision for robotic assisted partial nephrectomy. *Urology* **2018**, *116*, 227–228. [[CrossRef](#)]
90. Pessaux, P.; Diana, M.; Soler, L.; Piardi, T.; Mutter, D.; Marescaux, J. Robotic duodenopancreatectomy assisted with augmented reality and real-time fluorescence guidance. *Surg. Endosc.* **2014**, *28*, 2493–2498. [[CrossRef](#)]
91. Pessaux, P.; Diana, M.; Soler, L.; Piardi, T.; Mutter, D.; Marescaux, J. Towards cybernetic surgery: Robotic and augmented reality-assisted liver segmentectomy. *Langenbecks Arch. Surg.* **2014**, *400*, 381–385. [[CrossRef](#)]
92. Tang, R.; Ma, L.F.; Rong, Z.X.; Li, M.D.; Zeng, J.P.; Wang, X.D.; Liao, H.; Dong, J. Augmented reality technology for preoperative planning and intraoperative navigation during hepatobiliary surgery: A review of current methods. *HBPD Int.* **2018**, *17*, 101–112. [[CrossRef](#)]
93. Lin, L.; Shi, Y.; Tan, A.; Bogari, M.; Zhu, M.; Xin, Y.; Xu, H.; Zhang, Y.; Xie, L.; Chai, G. Mandibular angle split osteotomy based on a novel augmented reality navigation using specialized robot-assisted arms—A feasibility study. *J. Cranio Maxillofac. Surg.* **2016**, *44*, 215–223. [[CrossRef](#)]
94. Pratt, P.; Arora, A. Transoral robotic surgery: Image guidance and augmented reality. *ORL J.* **2018**, *80*, 204–212. [[CrossRef](#)]
95. Ewurum, C.H.; Guo, Y.; Pagnha, S.; Feng, Z.; Luo, X. Surgical navigation in orthopedics: Workflow and system review. *Adv. Exp. Med. Biol.* **2018**, *1093*, 47–63. [[CrossRef](#)] [[PubMed](#)]
96. Fard, M.J.; Ameri, S.; Ellis, R.D.; Chinnam, R.B.; Pandya, A.K.; Klein, M.D. Automated robot-assisted surgical skill evaluation: Predictive analytics approach. *Int. J. Med. Robot. Comput. Assist. Surg.* **2018**, *14*, e1850. [[CrossRef](#)] [[PubMed](#)]
97. Wang, Z.; Fey, A.M. Deep learning with convolutional neural network for objective skill evaluation in robot-assisted surgery. *Int. J. Comput. Assist. Radiol. Surg.* **2018**, *13*, 1959–1970. [[CrossRef](#)]
98. Fawaz, H.I.; Forestier, G.; Weber, J.; Idoumghar, L.; Muller, P.A. Accurate and interpretable evaluation of surgical skills from kinematic data using fully convolutional neural networks. *Int. J. Comput. Assist. Radiol. Surg.* **2019**, *14*, 1611–1617. [[CrossRef](#)] [[PubMed](#)]
99. Kassahun, Y.K.; Yu, B.; Tibebu, A.T.; Stoyanov, D.; Giannarou, S.; Metzen, J.H.; Poorten, E.V. Surgical robotics beyond enhanced dexterity instrumentation: A survey of machine learning techniques and their role in intelligent and autonomous surgical actions. *Int. J. Comput. Assist. Radiol. Surg.* **2015**, *11*, 553–568. [[CrossRef](#)]
100. Hutson, M. *Missing Data Hinder Replication of Artificial Intelligence Studies*; Science AAAS: Washington, DC, USA, 2018. Available online: <https://www.sciencemag.org/news/2018/02/missing-data-hinder-replication-artificial-intelligence-studies> (accessed on 9 January 2020).
101. Lipton, Z.C. The doctor just won't accept that! *arXiv* **2017**, arXiv:1711.08037.
102. Watson, D.S.; Krutzinna, J.; Bruce, I.N.; Griffiths, C.E.; McInnes, I.B.; Barnes, M.R.; Floridi, L. Clinical applications of machine learning algorithms: Beyond the black box. *BMJ* **2019**, *364*, l886. [[CrossRef](#)]
103. Cussenot, O.; Valeri, A.; Berthon, P.; Fournier, G.; Mangin, P. Hereditary prostate cancer and other genetic predispositions to prostate cancer. *Urol. Int.* **1998**, *60*, 30–34. [[CrossRef](#)]
104. Keane, P.; Topol, E.J. With an eye to AI and autonomous diagnosis. *NPJ Digit. Med.* **2018**, *1*, 40. [[CrossRef](#)]
105. Eifler, J.B.; Feng, Z.; Lin, B.M.; Partin, M.T.; Humphreys, E.B.; Han, M.; Epstein, J.I.; Walsh, P.C.; Trock, B.J.; Partin, A.W. An updated prostate cancer staging nomogram (Partin tables) based on cases from 2006 to 2011. *BJU Int.* **2012**, *111*, 22–29. [[CrossRef](#)]
106. Leyh-Bannurah, S.R.; Gazdovich, S.; Budäus, L.; Zaffuto, E.; Dell'Oglio, P.; Briganti, A.; Abdollah, F.; Montorsi, F.; Schifmann, J.; Menon, M.; et al. Population-based external validation of the updated 2012 partin tables in contemporary North American prostate cancer patients. *Prostate* **2016**, *77*, 105–113. [[CrossRef](#)] [[PubMed](#)]
107. Esteva, A.; Kuprel, B.; Novoa, R.A.; Ko, J.; Swetter, S.M.; Blau, H.M.; Thrun, S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **2017**, *542*, 115–118. [[CrossRef](#)]
108. Gulshan, V.; Peng, L.; Coram, M.; Stumpe, M.C.; Wu, D.; Narayanaswamy, A.; Venugopalan, S.; Widner, K.; Madams, T.; Cuadros, J.; et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* **2016**, *316*, 2402–2410. [[CrossRef](#)] [[PubMed](#)]

109. De Fauw, J.; Ledsam, J.R.; Romera-Paredes, B.; Nikolov, S.; Tomasev, N.; Blackwell, S.; Askham, H.; Glorot, X.; O'Donoghue, B.; Visentin, D.; et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat. Med.* **2018**, *24*, 1342–1350. [[CrossRef](#)]
110. Korbar, B.; Olofson, A.M.; Mirafior, A.P.; Nicka, C.M.; Suriawinata, M.A.; Torresani, L.; Suriawinata, A.A.; Hassanpour, S. Deep learning for classification of colorectal polyps on whole-slide images. *J. Pathol. Inform.* **2017**, *8*, 30. [[CrossRef](#)]
111. Komeda, Y.; Handa, H.; Watanabe, T.; Nomura, T.; Kitahashi, M.; Sakurai, T.; Okamoto, A.; Minami, T.; Kono, M.; Arizumi, T.; et al. Computer-aided diagnosis based on convolutional neural network system for colorectal polyp classification: Preliminary experience. *Oncology* **2017**, *93*, 30–34. [[CrossRef](#)]
112. Christodoulou, E.; Ma, J.; Collins, G.S.; Steyerberg, E.W.; Verbakel, J.Y.; Van Calster, B.; Evangelia, C.; Jie, M. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J. Clin. Epidemiol.* **2019**, *110*, 12–22. [[CrossRef](#)]
113. Dreiseitl, S.; Ohno-Machado, L. Logistic regression and artificial neural network classification models: A methodology review. *J. Biomed. Inform.* **2003**, *35*, 352–359. [[CrossRef](#)]
114. Cui, Z.; Gong, G. The effect of machine learning regression algorithms and sample size on individualized behavioral prediction with functional connectivity features. *NeuroImage* **2018**, *178*, 622–637. [[CrossRef](#)]
115. Shaikhina, T.; Lowe, D.; Daga, S.; Briggs, D.C.; Higgins, R.; Khovanova, N. Machine Learning for predictive modelling based on small data in biomedical engineering. *IFAC PapersOnLine* **2015**, *48*, 469–474. [[CrossRef](#)]
116. Hutson, M. Artificial intelligence faces reproducibility crisis. *Science* **2018**, *359*, 725–726. [[CrossRef](#)] [[PubMed](#)]
117. Bizzego, A.; Bussola, N.; Chierici, M.; Maggio, V.; Francescato, M.; Cima, L.; Cristoforetti, M.; Jurman, G.; Furlanello, C. Evaluating reproducibility of AI algorithms in digital pathology with DAPPER. *PLoS Comput. Biol.* **2019**, *15*, e1006269. [[CrossRef](#)] [[PubMed](#)]
118. Codari, M.; Schiaffino, S.; Sardanelli, F.; Trimboli, R.M. Artificial intelligence for breast MRI in 2008–2018: A systematic mapping review. *Am. J. Roentgenol.* **2019**, *212*, 280–292. [[CrossRef](#)] [[PubMed](#)]
119. Yassin, N.I.; Omran, S.; El Houby, E.M.F.; Allam, H. Machine learning techniques for breast cancer computer aided diagnosis using different image modalities: A systematic review. *Comput. Methods Programs Biomed.* **2018**, *156*, 25–45. [[CrossRef](#)] [[PubMed](#)]
120. Nindrea, R.D.; Aryandono, T.; Lazuardi, L.; Dwiprahasto, I. Diagnostic accuracy of different machine learning algorithms for breast cancer risk calculation: A meta-analysis. *Asian Pac. J. Cancer Prev.* **2018**, *19*, 1747–1752.
121. Murphy, A.; Skalski, M.R.; Gaillard, F. The utilisation of convolutional neural networks in detecting pulmonary nodules: A review. *Br. J. Radiol.* **2018**, *91*, 20180028. [[CrossRef](#)]
122. Rudie, J.D.; Rauschecker, A.M.; Bryan, R.N.; Davatzikos, C.; Mohan, S. Emerging applications of artificial intelligence in neuro-oncology. *Radiology* **2019**, *290*, 607–618. [[CrossRef](#)]
123. Nguyen, A.; Blears, E.E.; Ross, E.; Lall, R.R.; Ortega-Barnett, J. Machine learning applications for the differentiation of primary central nervous system lymphoma from glioblastoma on imaging: A systematic review and meta-analysis. *Neurosurg. Focus* **2018**, *45*, E5. [[CrossRef](#)]
124. De Souza, L.A.; Palm, C.; Mendel, R.; Hook, C.; Ebigbo, A.; Probst, A.; Messmann, H.; Weber, S.A.T.; Papa, J.P. A survey on Barrett's esophagus analysis using machine learning. *Comput. Biol. Med.* **2018**, *96*, 203–213. [[CrossRef](#)]
125. Sharma, R. The burden of prostate cancer is associated with human development index: Evidence from 87 countries, 1990–2016. *EPMA J.* **2019**, *10*, 137–152. [[CrossRef](#)]
126. Golubnitschaja, O.; Baban, B.; Boniolo, G.; Wang, W.; Bubnov, R.; Kapalla, M.; Krapfenbauer, K.; Mozaffari, M.S.; Costigliola, V. Medicine in the early twenty-first century: Paradigm and anticipation—EPMA position paper 2016. *EPMA J.* **2016**, *7*, 23. [[CrossRef](#)] [[PubMed](#)]
127. Lu, M.; Zhan, X. The crucial role of multiomic approach in cancer research and clinically relevant outcomes. *EPMA J.* **2018**, *9*, 77–102. [[CrossRef](#)] [[PubMed](#)]

